De la robustesse des modèles hydrologiques face à des conditions climatiques variables

Paul Royer-Gaspard

▶ **To cite this version:**

Paul Royer-Gaspard. De la robustesse des modèles hydrologiques face à des conditions climatiques variables. Hydrologie. Sorbonne Université UPMC, 2021. Français. NNT : . tel-03649353v1

HAL Id: tel-03649353

https://hal.inrae.fr/tel-03649353v1

Submitted on 22 Apr 2022 (v1), last revised 18 Jul 2022 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sorbonne Université

Ecole doctorale Géosciences Ressources Naturelles et Environnement (ED 398)

INRAE – UR HYCAR (Antony)

De la robustesse des modèles hydrologiques face à des conditions climatiques variables

Par Paul Royer-Gaspard

Thèse de doctorat en hydrologie

Dirigée par Vazken Andréassian et Guillaume Thirel

Présentée et soutenue publiquement le 09/04/2021

Devant un jury composé de :

Gilles DROGUE, Professeur, Université de Lorraine, LOTERR	Rapporteur
Denis RUELLAND, Ingénieur de recherche, CNRS, HydroSciences Montpellier	Rapporteur
Agnès DUCHARNE, Directrice de recherche, CNRS, METIS	Examinatrice
Anne-Catherine FAVRE, Professeure, Université Grenoble-Alpes, IGE	Examinatrice
Joël GAILHARD, Ingénieur expert, EDF-DTG	Invité
Vazken ANDRÉASSIAN, Ingénieur en chef IPEF, INRAE, HYCAR	Directeur de thèse
Guillaume THIREL, Chargé de recherche, INRAE, HYCAR	Encadrant

Résumé

Il est crucial de disposer d'outils fiables et robustes pour l'évaluation des risques hydrologiques et de la ressource en eau dans les bassins versants sujets à des changements de leurs conditions climatiques, afin que les politiques de gestion et de prévention soient efficaces. Le changement climatique d'origine anthropique en cours devrait particulièrement bouleverser le cycle de l'eau et ses dynamiques. Dans ce contexte, l'un des défis pour les modèles hydrologiques utilisés pour simuler ces effets sur le régime des cours d'eau est d'être capable de représenter adéquatement les processus hydrologiques dans de larges gammes de conditions climatiques. Les modèles actuels éprouvent cependant des difficultés à être performants dans ce registre. De nombreux travaux montrent de fait une difficulté générale à simuler les débits des cours d'eau en dehors des conditions rencontrées lors du calage de leurs paramètres. Les causes de ce manque de robustesse face à des conditions climatiques variables sont encore incertaines et peuvent être multiples : méthodes de calage des paramètres, erreurs dans les données, structures inadéquates... Dans cette thèse, une étude de ces causes est proposée afin d'identifier des solutions pour l'amélioration des modèles hydrologiques, en traitant les questions suivantes :

- Comment améliorer la robustesse des modèles hydrologiques sans modifier leurs équations et leur structure ?
- Peut-on identifier au sein d'un modèle hydrologique simple des défauts structurels limitant sa robustesse et peut-on les corriger ?

Des réponses à la première question ont été recherchées par l'analyse des performances de modèles hydrologiques conceptuels dans le cadre d'une procédure systématique de calage-contrôle sur un échantillon de 558 bassins versants français et australiens. D'abord, une étape de diagnostic des modèles hydrologiques a montré que les plus grandes difficultés étaient rencontrées par les modèles lors des variations de la relation précipitations-débit des bassins versants. Ensuite, une comparaison de différentes méthodes de calage a abouti au constat que le choix de la fonction objectif optimisée lors du calage avait une influence non négligeable sur la robustesse des modèles. L'emploi du critère de KGE calculé sur la racine carrée des débits apparaît à cet égard le choix le plus pertinent. Le choix de la formulation de l'évaporation potentielle est également un facteur important d'amélioration de la robustesse des modèles hydrologiques, bien que les résultats de la comparaison de plusieurs formulations soient assez hétérogènes sur l'échantillon d'étude. Malgré ces résultats encourageants, des choix adaptés de la méthode de calage et du calcul de l'évaporation potentielle ne permettent d'améliorer la robustesse des modèles que de manière marginale.

La deuxième question a suscité la mise en place d'une méthodologie de diagnostic des lacunes structurelles des modèles hydrologiques nuisant à leur robustesse. La méthode, fondée sur une analyse des compétitions entre les performances d'un modèle et de l'effet de ces compétitions sur ses états internes, a été appliquée au modèle GR4J. Deux défauts majeurs ont ainsi pu être identifiés dans la structure du modèle, limitant sa capacité à se montrer robuste pour la simulation simultanée de plusieurs gammes d'écoulement. Une tentative de modification via l'ajout d'un paramètre au modèle s'est révélée encourageante, bien que de faible influence sur les performances du modèle.

Malgré la modestie des progrès auxquels ils ont abouti, ces travaux ouvrent la voie à des améliorations plus satisfaisantes de la robustesse des modèles hydrologiques. Les résultats de la thèse concernant les méthodes de calage et le choix des formules d'évaporation potentielle pourraient permettre l'élaboration de diagnostics plus ciblés des modèles hydrologiques, en réduisant partiellement ces sources d'incertitudes. Nous espérons que ces travaux pourront susciter de futurs développements de la structure des modèles hydrologiques, notamment pour la simulation des dynamiques interannuelles du comportement des bassins versants.

Abstract

The reliability and the robustness of modelling tools used to assess hydrological risks and water resources in catchments facing climatic changes is critical to efficiently foster management and prevention policies. The current anthropogenic climate change should particularly unsettle the water cycle and its dynamics. Among the many challenges that climate change poses, the ability of hydrological models to adequately represent hydrological processes over a large range of climatic conditions is key to simulate climate change impacts on the regime of rivers. However, modern hydrological models still fail to perform well in such various situations. Many studies actually show a general difficulty to simulate river streamflow outside of the conditions experienced during model parameter calibration. The causes of this lack of robustness are yet uncertain and may be manifold: parameter calibration methods, measurement errors, inadequate model structure... This thesis aims at studying these causes in order to identify solutions for model improvement and strives to answer the following questions:

- Is it possible to improve hydrological model robustness without modifying model structure and equations?
- Can we identify structural weaknesses in hydrological models impacting its robustness and resolve them?

Answers to the first question were seeked by analyzing the performances of a few hydrological models within the framework of a systematic split-sample calibration-evaluation experiment on a set of 558 catchments in France and Australia. Our diagnosis revealed that the hydrological models face major difficulties to simulate changes in catchments response to precipitation. The comparison of different calibration methods showed that the choice of the optimized objective function had a significant impact on model robustness. When computed on square-rooted streamflow, the KGE appeared as the most relevant choice. The way potential evaporation is computed also influences model robustness, although our comparison of a few potential evaporation models show rather heterogeneous results across the catchment set. In spite of encouraging results, only marginal improvements on model robustness were reached by a better selection of calibration techniques and potential evaporation models.

A method specifically designed to diagnose structural weaknesses impacting model robustness was proposed to answer the second question. The approach is based on an analyzis of performance trade-offs in a multi-objective framework and of the effects of these trade-offs on model internal behaviour. It was applied to the GR4J model. A couple of major structural deficiencies was identified. These deficiencies likely prevent the model from providing robust simulations in different streamflow ranges simultaneously. An attempt to modify the structure of GR4J yielded to an encouraging yet modest improvement of its performance.

Despite the light enhancement of hydrological model robustness achieved in this work, it may pave the way to further advances. The results obtained on calibration techniques and potential evaporation computation could foster the elaboration of a more in-depth analysis of model weaknesses by partially removing these sources of uncertainty. We hope that our work may stimulate model structural development, particularly towards better representation of catchment interannual dynamics.

Remerciements

Avant d'exprimer mes sentiments à tous ceux qui m'ont aidé à *écrire* cette thèse, de près ou de loin, j'aimerais commencer par remercier celui ou celle qui a entrepris de la *lire*. Quels que soient les motifs qui l'ont conduit à le faire – intérêt pour le contenu scientifique, goût pour la stricte lecture des remerciements, ou clic raté –, je le/la remercie et j'espère qu'il/elle trouvera ce qu'il/elle y cherche.

Les premières personnes dont je souhaiterais chanter les louanges sont mes directeurs de thèse. Le premier, Vazken, m'a guidé pour donner une direction pertinente à mon travail, accompagné pour communiquer convenablement mes résultats, bombardé d'idées nouvelles et de pistes à explorer, rappelé que la simplicité est toujours préférable à la complexité. Guillaume, quant à lui, a relu minutieusement chaque phrase de mes articles et manuscrit pour chasser imprécisions et incohérences, m'a aidé à mettre mes idées en ordre quand tout semblait flou, et plus généralement a cherché à m'enseigner la prudence et la rigueur dans l'interprétation des démarches et des résultats. Je leur dois d'avoir veillé sur mon travail et de m'avoir permis de le mener à son terme.

J'aimerais ensuite remercier sincèrement les cinq membres de mon jury : Gilles Drogue, Denis Ruelland, Agnès Ducharne, Anne-Catherine Favre et Joël Gailhard, d'avoir accepté d'évaluer la thèse. Lire un mémoire de plusieurs centaines de pages est une tâche incroyablement fastidieuse, et il faut un certain courage pour accepter un tel labeur. J'adresse à ce titre une mention spéciale pour M. Ruelland et M. Drogue pour avoir été les rapporteurs du manuscrit, dont les synthèses m'ont aidé à bien préparer la soutenance orale. J'aimerais enfin souligner à quel point j'ai apprécié les discussions que j'ai eu lors de la soutenance avec l'ensemble des membres du jury, dont les questions et la bienveillance m'ont très vite permis de me sentir à l'aise. Enfin, mon travail a été aiguillé au long de ces trois ans par un comité de suivi que je souhaite aussi chaleureusement remercier : Ludovic Oudin, Laurent Coron, Nathalie de Noblet-Ducoudré et Éric Martin. Je regrette qu'il n'ait pas été possible de se réunir tous ensemble pour célébrer la fin de ce travail.

Comment ne pas mentionner l'équipe Hydro : difficile d'imaginer cadre d'épanouissement plus propice pour un doctorant que cette équipe. Une grande part de ma gratitude revient à ses membres, pour l'accueil chaleureux qui est réservé à tout nouvel arrivant, pour l'enthousiasme scientifique qui y est partagé, pour le supplément d'âme insufflé à l'équipe par ses cadres, pour les séances de frisbee, de volley et de football à la pause déjeuner... J'ai une pensée particulière pour Daniela et Anthony, chers collègues de bureau et désormais également docteurs.

Je terminerais comme il convient en saluant mes géniteurs et le reste de leur progéniture, à qui je dédie mon premier chapitre et pour qui j'ai cherché à le rendre aussi didactique que possible. Qu'aucun ne se sente obligé de le lire après ça, bien sûr. Enfin, je souhaite exprimer toute ma reconnaissance à la personne qui m'accompagne depuis quelques années, dont la valeur ne se résume vraiment pas à ses possessions en bitcoins, ainsi qu'à sa famille, que j'aime beaucoup.

PS : il est toujours agréable de trouver son nom dans des remerciements, alors j'adresse aussi un immense merci à toutes les personnes que j'ai fréquenté à INRAE pendant ces années et que je n'ai pas cité nommément plus haut : Charles, Maria-Helena, Olivier, Alban, François, François, Gaëlle, Arnaud, Léonard, Paul, Antoine, Thibault, Myriam, Cyril, Romane, Lila, Pierre, Julie, Laure, José, Manon, Morgane, Cédric, Gaïa Valentin, Benoît, Bruno, Anne-Lise, Jean-Baptiste, Matthieu, Sylvain, Yannick, Julien, Cédric, Hocine, Maria, Alexis, Timo, Samy, Florent, Aya, Hoang, Subrata, Alienor, Marie-Line, Olivier, Christophe, Guillaume, Salif, Valérie, Ronan, Nathalie, Frédérique, Laurence, Stéphanie, Véronique, Aurélien, Oumar, Hicham, Benjamin, Peng, Sacha, Laura, Saran, Vandoir, Fernando, Ivan, Laurène, Chiara et Alberto.

Table des matières

Introduction générale **1**

I Enjeux, problématique et méthodes *7*

1 Généralités sur les modèles hydrologiques et leur évaluation **9**

1.1 La modélisation hydrologique : définitions et concepts 11
1.1.1 Objet d'étude 11
1.1.2 Introduction à la modélisation hydrologique 16
1.1.3 Approches de modélisation en hydrologie 19

1.2 Modéliser les conséquences du changement climatique en hydrologie 22
1.2.1 Enjeux et attentes vis-à-vis de la modélisation 22
1.2.2 Robustesse d'un modèle hydrologique en climat non-stationnaire 24

1.3 Synthèse 28

2 La robustesse des modèles hydrologiques : état de l'art et discussion **31**

2.1 Inventaire des faiblesses constatées des modèles 33
2.1.1 Constats généraux 33
2.1.2 Discussions suscitées par le constat d'un manque généralisé de robustesse . 36

2.2 Causes du manque de robustesse : rôle du calage des paramètres 39
2.2.1 Principes usuels du calage des modèles hydrologiques 39
2.2.2 Recherche d'un optimum hydrologique 42
2.2.3 Perspectives pour un calage plus robuste des modèles hydrologiques 45

2.3 Causes du manque de robustesse : structures inadaptées 47
2.3.1 Principes de l'évaluation de la structure des modèles 47
2.3.2 Lacunes identifiées en conditions climatiques variables 50

2.4 Synthèse 53

3 Matériel et modèles **57**

3.1 Base de données 59
3.1.1 Motifs de sélection des bassins versants 59
3.1.2 Origine des données 60
3.1.3 Caractéristiques des bassins versants de l'échantillon 61

3.2 Modèles hydrologiques 67
3.2.1 GR4J 67
3.2.2 TOPMO 68
3.2.3 SimHyd 69
3.2.4 CemaNeige 70

3.3 Formules d'évaporation potentielle 71
3.3.1 Oudin 72
3.3.2 Penman-Monteith 72

3.3.3 Morton . . . 74
3.3.4 Comparaison des chroniques d'évaporation potentielle . . . 76
3.4 Calage des modèles hydrologiques . . . 80
3.4.1 Fonctions objectif utilisées . . . 80
3.4.2 Algorithme de calage . . . 82
3.5 Critères d'évaluation des performances des modèles . . . 83
3.6 Synthèse . . . 85

II Diagnostic de la robustesse des modèles hydrologiques *87*

4 Procédures d'évaluation de la robustesse des modèles hydrologiques 89
4.1 Quelle stratégie de test pour quel objectif ? . . . 91
4.1.1 Le *Generalized Split-Sample Test* : stratégie explorative en extrapolation . . . 91
4.1.2 Intérêt d'une stratégie de test en interpolation . . . 92
4.2 Note technique : *A proxy metric to assess hydrological model robustness in a changing climate* . . . 93
4.2.1 Introduction . . . 94
4.2.2 Description of the Proxy for Model Robustness . . . 95
4.2.3 Material and methods . . . 97
4.2.4 Results: Reliability of the Proxy for Model Robustness to indicate model robustness . . . 99
4.2.5 Discussion . . . 101
4.2.6 Conclusions . . . 103
4.3 Compléments sur le *Proxy for Model Robustness* . . . 104
4.3.1 Vérification des résultats sur l'échantillon de bassins versants australien . . . 104
4.3.2 Comparaison de la robustesse des modèles hydrologiques utilisés dans la thèse . . . 105
4.3.3 Estimation de la robustesse d'un modèle à base physique . . . 106
4.3.4 Remarques pour la suite de la thèse . . . 110
4.4 Synthèse . . . 110

5 Étude de l'influence du calage sur la robustesse des modèles hydrologiques 111
5.1 Recherche d'une fonction objectif pour un calage robuste . . . 113
5.1.1 Comportement des fonctions objectif . . . 113
5.1.2 Comparaison des fonctions objectif à travers la robustesse des modèles hydrologiques . . . 116
5.1.3 Sensibilité de la robustesse aux choix de calage . . . 123
5.2 Liens entre robustesse des modèles et variables hydro-climatiques . . . 129
5.2.1 Liens identifiés en interpolation . . . 129
5.2.2 Liens identifiés en extrapolation . . . 131
5.3 Synthèse . . . 136

6 Étude de l'influence de l'évaporation potentielle sur la robustesse des modèles hydrologiques 139

6.1 Comparaison de formules d'évaporation potentielle à travers la modélisation hydrologique . . . 141
6.1.1 Comparaison générale des formules d'évaporation . . . 141
6.1.2 Interprétation des résultats obtenus avec la formule d'Oudin sur les bassins australiens . . . 150
6.1.3 Interprétation des résultats obtenus avec les formules de Penman-Monteith et d'Oudin sur les bassins français . . . 154
6.1.4 Synthèse . . . 157

6.2 Comparaison de formules d'évaporation potentielle au travers des variations interannuelles du débit . . . 158
6.2.1 Description de la méthode . . . 158
6.2.2 Élasticité du débit vis-à-vis des variables climatiques annuelles . . . 160
6.2.3 Élasticité du débit vis-à-vis des variables climatiques saisonnières . . . 162
6.2.4 Comparaison des performances des modèles linéaires saisonniers . . . 168
6.2.5 Discussion des résultats concernant les formules d'évaporation potentielle . . . 172
6.2.6 Tentative d'interprétation physique des effets de mémoire . . . 175

6.3 Synthèse . . . 177

III Évaluation des limites structurelles du modèle GR4J 181

7 Méthode d'évaluation des limites structurelles d'un modèle hydrologique 183

7.1 Principes de la méthode . . . 185
7.1.1 Intérêt de la méthode . . . 185
7.1.2 Description de la méthode . . . 186

7.2 Évaluation de la polyvalence du modèle GR4J . . . 192
7.2.1 Échantillonnage des paramètres du modèle . . . 192
7.2.2 Sensibilité des performances au choix des paramètres . . . 192
7.2.3 Polyvalence du modèle au regard du biais et du PMR . . . 195
7.2.4 Polyvalence du modèle au regard des gammes de débit . . . 199

7.3 Analyse du comportement du modèle GR4J . . . 202
7.3.1 Compatibilité entre les critères de performances . . . 202
7.3.2 Comparaison des paramètres associés aux gammes de débit . . . 205
7.3.3 Étude des états internes du modèle GR4J associés aux gammes de débit . . . 208
7.3.4 Synthèse et discussion de l'analyse de la structure de GR4J . . . 215

7.4 Discussion de la méthode mise en œuvre . . . 216
7.4.1 Possibilités d'application de la méthode à d'autres modèles . . . 216
7.4.2 Discussion sur le calcul de la polyvalence du modèle . . . 217

7.5 Synthèse . . . 219

8 Modifications de la puissance de la production du modèle GR4J 221

8.1 Principes et méthodologie . . . 223
8.1.1 Motivations et présentation de la modification . . . 223
8.1.2 Méthodologie d'évaluation . . . 225

8.2 Évaluation du paramétrage symétrique du réservoir de production . . . 227

8.2.1 Évaluation de GR4J_sym en calage . . . 227
8.2.2 Définition de l'intervalle de variations du paramètre α . . . 233
8.2.3 Performances de GR4J_sym en contrôle . . . 236
8.2.4 Discussion des résultats . . . 239
8.3 Évaluation du paramétrage asymétrique du réservoir de production . . . 243
8.3.1 Performances de GR4J_asym . . . 244
8.3.2 Analyse du comportement du modèle . . . 248
8.3.3 Discussion des résultats . . . 251
8.4 Synthèse . . . 253

Conclusion générale **256**

Annexes ***285***

A Suppléments à la note technique **287**
A.1 Characterisation of model bias across DSST setups . . . 289
A.2 The choice of an adequate mathematical expression . . . 289
A.3 Reliability of the metric for different DSST sub-period lengths . . . 291

B Discussions sur la définition des gammes de débit **293**
B.1 Introduction . . . 295
B.2 Problématique : définition d'une gamme de débit . . . 295
B.2.1 Gammes définies à partir de parts fixes des pas de temps . . . 295
B.2.2 Interprétation des problèmes soulevés par les parts fixes des pas de temps . . . 296
B.2.3 Gammes définies à partir de part fixes du débit cumulé . . . 298
B.2.4 Objectifs . . . 300
B.3 Méthodologie . . . 301
B.3.1 Définitions des gammes de débit et des fonctions objectif associées . . . 301
B.3.2 Méthode d'évaluation des fonctions objectif fondées sur les différentes définitions des gammes de débit . . . 301
B.4 Comparaison des fonctions objectif fondées sur les gammes de débit . . . 302
B.4.1 Performances en calage . . . 302
B.4.2 Performances en contrôle . . . 303
B.4.3 Comparaison des valeurs de KGE en contrôle . . . 303
B.4.4 Comparaison des valeurs de NSE calculé sur les débits transformés en contrôle 303
B.5 Synthèse . . . 306

C Implémentation d'une fonction objectif incluant explicitement la robustesse **307**
C.1 Introduction . . . 309
C.2 Méthodologie . . . 309
C.2.1 Description des fonctions objectif évaluées . . . 309
C.2.2 Méthode d'évaluation des fonctions objectif . . . 310
C.3 Résultats . . . 311

C.4 Discussion 314
C.5 Synthèse 315

D Vérification des données de précipitations de cinq bassins versants français 317
D.1 Méthodologie 319
D.2 Bassin de la Laines à Soulaines-Dhuys (H1333010) 319
D.3 Bassin de la Teyssonne à Changy (K1084010) 321
D.4 Bassin de la Corrèze à Saint-Yrieix-le-Déjalat (P3322510) 322
D.5 Bassin du Saint-Nicolas à Rougemont-le-Château (U2305210) 324
D.6 Bassin de la Siagne à Callian (Y5514040) 325

E Courbes des biais glissants des modèles pour les bassins français de l'échantillon 329
E.1 Introduction 331
E.2 Mise en œuvre des courbes 331
E.3 Courbes de biais glissants pour le sous-échantillon de bassins français 331

F Courbes des biais glissants des modèles pour les bassins australiens de l'échantillon371
F.1 Introduction 373
F.2 Mise en œuvre des courbes 373
F.3 Courbes de biais glissants pour le sous-échantillon de bassins australien 373

Introduction générale

Contexte

Le changement climatique engendré par les émissions anthropiques de gaz à effet de serre est l'un des bouleversements majeurs qui menacent nos sociétés dans le siècle à venir. Son lien étroit avec l'utilisation des énergies fossiles à l'origine des révolutions industrielles, dont il est encore difficile de se passer pour la production d'énergie, en fait un danger qui peut sembler inéluctable. De plus, la mondialisation des conséquences des émissions locales de gaz à effet de serre dilue le sentiment de responsabilité qui pourrait encourager les efforts nécessaires à la baisse de ces émissions. Alors que le sixième rapport d'évaluation du Groupe d'Experts Intergouvernemental sur l'Évolution du Climat (GIEC) est en préparation pour l'année 2022, force est de constater aujourd'hui que les intentions nationales de contribution à l'effort de réduction des émissions proposées lors de l'Accord de Paris de 2015 ne sauraient suffire pour limiter le réchauffement global à un niveau inférieur à 2 ℃ (UNFCCC, 2016). Une attention particulière doit donc être portée sur la caractérisation des modifications de l'environnement qui pourraient survenir dans les futures décennies, afin d'anticiper les risques et d'aiguiller les politiques d'adaptation aux conséquences du changement climatique.

Parce que l'eau est à la fois une ressource indispensable et un vecteur de catastrophes naturelles dévastatrices, les effets du changement climatique sur le cycle de l'eau se classent peut-être parmi les phénomènes les plus menaçants qu'il convient de circonscrire. La pérennité des infrastructures et des politiques de gestion liées à la maitrise de l'eau repose à la fois sur leur capacité à déployer une protection efficace lorsqu'elle est surabondante et sur les possibilités d'économiser la ressource lorsqu'elle vient à manquer. Pour vérifier la compatibilité de ces infrastructures et de ces politiques avec les modifications du cycle de l'eau et, le cas échéant, envisager leur redimensionnement, il est absolument nécessaire de disposer d'évaluations fiables de ces dynamiques. Ce besoin est d'autant plus pressant dans les régions où la gestion de l'eau est déjà sous tension et où les vulnérabilités aux risques hydro-météorologiques sont les plus préoccupantes.
L'appréciation des changements du régime hydrologique repose sur une cascade de simulations numériques réalisées par des modèles, dont les calculs permettent de disposer de projections climatiques et hydrologiques. Au début de la chaîne de modélisation se trouvent les modèles de circulation générale, qui estiment les conséquences des émissions de gaz à effet de serre sur le climat global. Ces émissions futures sont elles-mêmes regroupées en scénarios, reflétant une variété de politiques à l'échelle mondiale. En fin de chaîne, après plusieurs autres maillons, les modèles hydrologiques intègrent les effets des conditions climatiques futures sur le fonctionnement des bassins versants. Il n'existe pas aujourd'hui de modèle-monde pertinent aux échelles locales qui permettrait de remplacer une telle chaîne par une unique composante. De ce fait, chaque maillon propage vers le maillon suivant, à la fois les incertitudes attenantes aux maillons précédents, et les incertitudes de ses propres simulations. Bien que certains effets du changement climatique sur les dynamiques de l'eau continentale soient estimés de manière tangible, des incertitudes importantes subsistent encore vis-à-vis de nombreux phénomènes, limitant les marges de manœuvre en matière de décisions politiques et économiques.

Bien que, comme le suggèrent Blöschl *et al.* (2019), moins d'emphase ait peut-être été portée sur la communication des incertitudes associées à la modélisation hydrologique, celle-ci n'en pas exempte. Ces incertitudes proviennent en particulier de doutes vis-à-vis de la capacité des modèles hydrologiques à représenter la non-stationnarité du régime hydrologique des bassins versants, c'est-à-dire les modifications de ses caractéristiques statistiques. La non-stationnarité du régime hydrologique peut être induite par deux facteurs majeurs de changement : une évolution des conditions climatiques, c'est-à-dire une modification des conditions aux limites des bassins versants, d'une part, une transformation du comportement hydrologique du bassin, matérialisée par un fléchissement durable de la relation précipitations-débit, d'autre part. En général, les rôles respectifs de ces facteurs

sont très difficiles à distinguer, notamment parce que l'évolution du climat influence le comportement du bassin versant. Alors que la modélisation des changements du régime hydrologique liés à une modification directe du comportement d'un bassin versant, par exemple liée à un changement de l'usage des sols, est déjà une tâche particulièrement complexe, la modélisation hydrologique dans un climat en évolution semble être également un défi de taille.

Problématique

Les travaux d'évaluation des modèles hydrologiques ont démontré une capacité générale limitée pour simuler les évolutions passées des écoulements liées aux variations des conditions climatiques. Une application systématique du *Differential Split-Sample Test* (DSST), une procédure de test de modèles en calage-contrôle dans des conditions climatiques changeantes proposée par Klemeš (1986), a conduit à formuler le constat que les modèles hydrologiques montrent des performances fâcheusement dégradées lorsque placés dans des conditions de modélisation différentes de celles rencontrées en calage. Cette dégradation culmine dans la représentation des variations temporelles du bilan hydrique des bassins versants, aboutissant à des simulations largement biaisées. S'il est parfois possible de détecter des liens entre les variations du climat, des performances et des paramètres des modèles, ceux-ci sont difficiles à exploiter en pratique pour corriger les erreurs *a posteriori*. Malheureusement, le manque de robustesse des modèles hydrologiques face aux changements des conditions climatiques est aujourd'hui quasiment un consensus scientifique.

Ce constat soulève la question épineuse de la confiance qu'il est possible d'accorder aux modèles hydrologiques pour la simulation des effets du changement climatique sur le régime des cours d'eau. Pour améliorer la qualité des projections hydrologiques, il apparaît donc essentiel d'identifier les causes du manque général de robustesse des modèles hydrologiques. Une fois ces défaillances connues et diagnostiquées, des solutions pourraient être envisagées pour y remédier. Cette thèse prétend s'inscrire dans cette démarche et se propose de répondre à la question suivante :

« Comment améliorer la robustesse des modèles hydrologiques face aux changements des conditions climatiques ? »

La thèse est financée dans le cadre du projet *AquaClew*, inclus dans le programme *JPI Climate* de la European Research Area for Climate Services (ERA4CS). L'objectif du projet est de définir des voies d'amélioration des services climatiques destinés aux acteurs du secteur de l'eau. Au sein du projet, la thèse est rattachée plus spécifiquement au *Work Package 2*, constitué autour de réflexions sur la fiabilité des données mises à disposition dans les services climatiques. Bien que l'ensemble du programme soit étroitement lié à l'estimation des effets du changement climatique en Europe, les travaux réalisés dans le cadre de la thèse se sont focalisés exclusivement sur l'évaluation des modèles hydrologiques à partir de données historiques.

Un faisceau de pistes est généralement évoqué pour expliquer le manque de robustesse des modèles hydrologiques constaté dans les études publiées sur le sujet.
Une première piste concerne l'étape de calage des paramètres de modèles hydrologiques. Celle-ci est susceptible de jouer un rôle important sur la robustesse, dans la mesure où elle permet de dimensionner les équations des modèles pour qu'elles représentent les processus à l'œuvre dans les bassins étudiés. Le calage des paramètres d'un modèle sur des données est réalisé à partir de l'optimisation d'une fonction objectif matérialisant l'adéquation entre les simulations et les observations. Malgré l'utilisation très commune de quelques fonctions objectif populaires, cette pratique constitue un sujet récurrent de questionnement en modélisation. Il est de fait généralement admis que

le choix d'une fonction objectif adéquate n'est pas chose aisée. Une stratégie de calage pertinente doit par exemple tenir compte des interactions entre erreurs de mesures et erreurs structurelles des modèles (hétéroscédasticité), ou de la multiplicité des jeux de paramètres optimaux au sens mathématique (équifinalité). Ces facteurs éloignent le jeu de paramètres optimum mathématique, i.e. celui de la fonction objectif, du jeu de paramètres optimum hydrologique, i.e. représentant au mieux les processus hydrologiques, et contribuent vraisemblablement à détériorer la robustesse des modèles hydrologiques. Pourtant, en matière de robustesse face aux variations des conditions climatiques, nous recensons peu de références et d'études comparatives pour guider le choix de fonctions objectif pertinentes. Il nous semble donc primordial d'établir une hiérarchie des pratiques en matière de calage.
Une deuxième piste de travail est centrée sur la structure et les équations des modèles. Les modèles hydrologiques n'étant guère plus que des représentations imparfaites de la réalité, ils comportent nécessairement des approximations et des défauts. L'adéquation entre les processus représentés et les processus déterminant réellement les écoulements dépendent en ce sens du contexte géographique et temporel de leur utilisation, selon les caractéristiques des bassins versants étudiés. Tout l'enjeu de la modélisation consiste à sélectionner les processus les plus pertinents pour caractériser le comportement de ces bassins. Ce choix est toutefois rendu complexe en raison de la compensation des erreurs structurelles par les paramètres lors des phases de calage : des modèles de structures très différentes peuvent ainsi obtenir des performances similaires. Aussi la question de la sélection des modèles est-elle parfois réduite à des questions de principes plutôt qu'à une comparaison objective. Les performances des modèles hydrologiques sont également peu sensibles aux entrées d'évaporation potentielle, estimées elles-mêmes par des modèles, ou à la représentation interne du processus d'évaporation réelle. Les modèles hydrologiques semblent peiner à représenter des variations de la réponse des bassins versants aux précipitations, notamment lors de sécheresses persistantes. Dans la mesure où l'évaporation constitue un terme prépondérant du bilan hydrique, il semble important de questionner la pertinence de ses représentations, de même que la capacité des modèles à simuler des dynamiques interannuelles des réserves d'eau des bassins versants.

Organisation du mémoire

Pour tenter de répondre méthodiquement à la problématique posée, ce mémoire est divisé en huit chapitres, répartis en trois parties.

La première partie pose les bases méthodologiques.

Le premier chapitre présente les notions générales afférentes à la modélisation hydrologique dans un contexte de changement climatique. Une définition des objets et des concepts de la modélisation hydrologique introduit le chapitre. Une courte synthèse des enjeux liés à la mesure des variables hydrologiques y est notamment proposée, ainsi qu'une description des différentes philosophies de représentation des bassins versants sur lesquelles les modèles hydrologiques actuels sont fondés. Le chapitre aborde ensuite plus spécifiquement les enjeux de la modélisation des effets du changement climatique sur les régimes hydrologiques. Le chapitre est conclu par une définition de la robustesse des modèles hydrologiques et par une présentation des moyens de l'évaluer.
Le deuxième chapitre expose l'état de l'art en matière d'étude de la robustesse des modèles hydrologiques. Après un inventaire des principales faiblesses constatées vis-à-vis des modèles hydrologiques, le chapitre présente une analyse détaillée des causes principales relevées par la littérature scientifique pour expliquer ce manque général de robustesse. L'analyse porte d'abord sur le rôle du calage des paramètres des modèles et se concentre ensuite sur les défauts liés à leur structure.

Ces commentaires débouchent sur l'énoncé de la problématique et des objectifs poursuivis dans la thèse.
Enfin, le troisième chapitre introduit les données, les modèles et les méthodes employés dans la thèse. La base de données constituée d'un important échantillon de bassins versants français et australiens y est d'abord décrite, puis les modèles hydrologiques ainsi que les formules d'évaporation potentielle évalués dans la suite des travaux. Le chapitre se termine par une présentation des principales méthodes de calage et d'évaluation des performances des modèles.

La deuxième partie expose les principaux résultats concernant l'évaluation des choix de modélisation en dehors de la structure des modèles hydrologiques.
Le quatrième chapitre explique les procédures d'évaluation de la robustesse des modèles utilisées dans toute la deuxième partie. Le chapitre commence par une description des expériences de calage-contrôle mises en œuvre, qui correspondent à une évaluation des modèles en conditions d'extrapolation. Le reste du chapitre est consacré à l'introduction d'une nouvelle métrique pour l'estimation de la robustesse des modèles hydrologiques en interpolation. Cette métrique est censée procurer un aperçu de la capacité des modèles hydrologiques à représenter les effets des variations interannuelles du climat sur les écoulements sans nécessiter l'utilisation de procédures de calage-contrôle. Les paragraphes liés à cette métrique se présentent sous la forme d'une note technique, soumise dans une revue internationale à comité de lecture.
Le cinquième chapitre montre les résultats concernant l'influence des conditions de calage sur la robustesse des modèles hydrologiques. Le chapitre présente d'abord une comparaison de plusieurs fonctions objectif dans le but de déterminer celles sélectionnant les paramètres les plus robustes pour les modèles hydrologiques. La suite du chapitre est consacrée à l'analyse des performances des modèles en fonction des changements hydro-climatiques rencontrés.
Le sixième chapitre est constitué autour de la comparaison de plusieurs formulations pour le calcul de l'évaporation potentielle. Cette comparaison est dans un premier temps menée au regard de la robustesse des modèles hydrologiques. Elle débouche sur le diagnostic précis de quelques formulations pour une utilisation en modélisation hydrologique. L'analyse est dans un second temps portée sur les pouvoirs explicatifs respectifs de ces formulations vis-à-vis des variations interannuelles du débit des bassins versants, en s'affranchissant des modèles hydrologiques.

La troisième partie s'articule autour du diagnostic de la structure des modèles hydrologiques.
Une méthode de diagnostic des lacunes structurelles des modèles hydrologiques est proposée au septième chapitre. La méthodologie a pour but de déterminer les composantes des modèles limitant leur robustesse et leur capacité à simuler correctement différentes parties des hydrogrammes en même temps. Après une présentation de ses principes, elle est appliquée au modèle GR4J et aboutit à l'identification de plusieurs pistes à suivre pour améliorer la structure du modèle.
Le huitième et dernier chapitre expose les résultats obtenus à partir d'une modification du modèle GR4J, à partir des conclusions du chapitre précédent. Au-delà de l'influence sur les performances et la robustesse, le chapitre rapporte les effets de la modification sur le comportement du modèle. Le chapitre s'achève sur la proposition d'autres remaniements qui pourraient bénéficier à la robustesse de GR4J.

Partie I

Enjeux, problématique et méthodes

Chapitre 1.

Généralités sur les modèles hydrologiques et leur évaluation

Introduction

Cette thèse porte sur la robustesse des modèles pluie-débit en contexte de changement climatique. Afin d'inscrire ce sujet dans le paysage scientifique, ce chapitre s'attache à préciser les termes utilisés pour définir ce sujet et ses enjeux. Qu'est-ce que l'hydrologie ? Quel est son objet d'étude ? Comment le représente-t-on ? À quoi sert un modèle ? Comment modélise-t-on le changement climatique en hydrologie ? Que doit-on attendre d'un modèle hydrologique destiné à cette utilisation ? Autant de questions que les paragraphes suivants tentent d'aborder progressivement.

1.1. La modélisation hydrologique : définitions et concepts

1.1.1. Objet d'étude

1.1.1.a. Qu'est-ce que l'hydrologie ?

Traduit littéralement du grec ancien, l'hydrologie désigne la science de l'eau. Les scientifiques qui étudient l'hydrologie sont les hydrologues. Cette définition est un peu vague tant l'eau est omniprésente sur Terre et sous des formes variées. On en trouve dans l'atmosphère, dans les mers et dans les océans, sur les continents et même bien souvent dedans. Dans la plupart de ces lieux, qui composent l'hydrosphère, elle existe sous plusieurs états : gazeux, liquide et solide. Ces espaces sont interconnectés, et forment les grandes étapes du voyage des molécules d'eau, appelé communément cycle de l'eau. Schématiquement, l'eau à la surface des océans s'évapore vers l'atmosphère sous l'action du soleil, avant de se condenser en nuages, qui eux-mêmes alimentent les précipitations sur les continents. L'eau continentale s'écoule en formant torrents, rivières, fleuves et se jette dans la mer, et ainsi de suite depuis des millions d'années. Selon Dooge (1988), l'enjeu ultime de l'hydrologie est la compréhension et l'évaluation des quantités d'eau qui passent au fil du temps d'un milieu à un autre, autrement dit la résolution de l'équation de conservation de la masse (Figure 1.1).

Toutefois, il est très difficile pour une seule personne de connaître parfaitement l'ensemble des mécanismes qui font tourner un tel cycle, tant ils sont divers et complexes. Son étude est donc compartimentée en trois grandes sous-disciplines, ayant chacune pour objet d'étude l'une des trois grandes étapes : l'océan est aux océanographes, l'atmosphère aux météorologues, les continents sont aux hydrologues. Ce qui est généralement appelé hydrologie devrait donc s'appeler hydrologie continentale, et c'est dans cette discipline que cette thèse s'inscrit.

Penman (1961) suggère que l'hydrologie répond à une question d'apparence simple : « qu'arrive-t-il à la pluie ? ». Répondre à cette question revient à constater que la pluie tombée peut prendre des chemins très différents : écoulement direct dans un cours d'eau, infiltration dans le sous-sol, évaporation dans l'atmosphère, consommation par la biosphère... Qu'une goutte emprunte l'un ou l'autre de ces chemins dépend d'une grande diversité de facteurs. L'hydrologie continentale est donc à la confluence de nombreuses disciplines, qui constituent des domaines de recherche à part entière. Rodier en dresse une liste peut-être non-exhaustive dans l'ouvrage de Roche (1963) : « hydraulique, hydrodynamique fluviale, calcul des probabilités, climatologie, géologie, pédologie, géomorphologie ». Comme nous le verrons dans la suite, un des enjeux de la modélisation hydrologique est d'intégrer les principales connaissances issues de ces disciplines. Oublier l'une d'entre elles serait

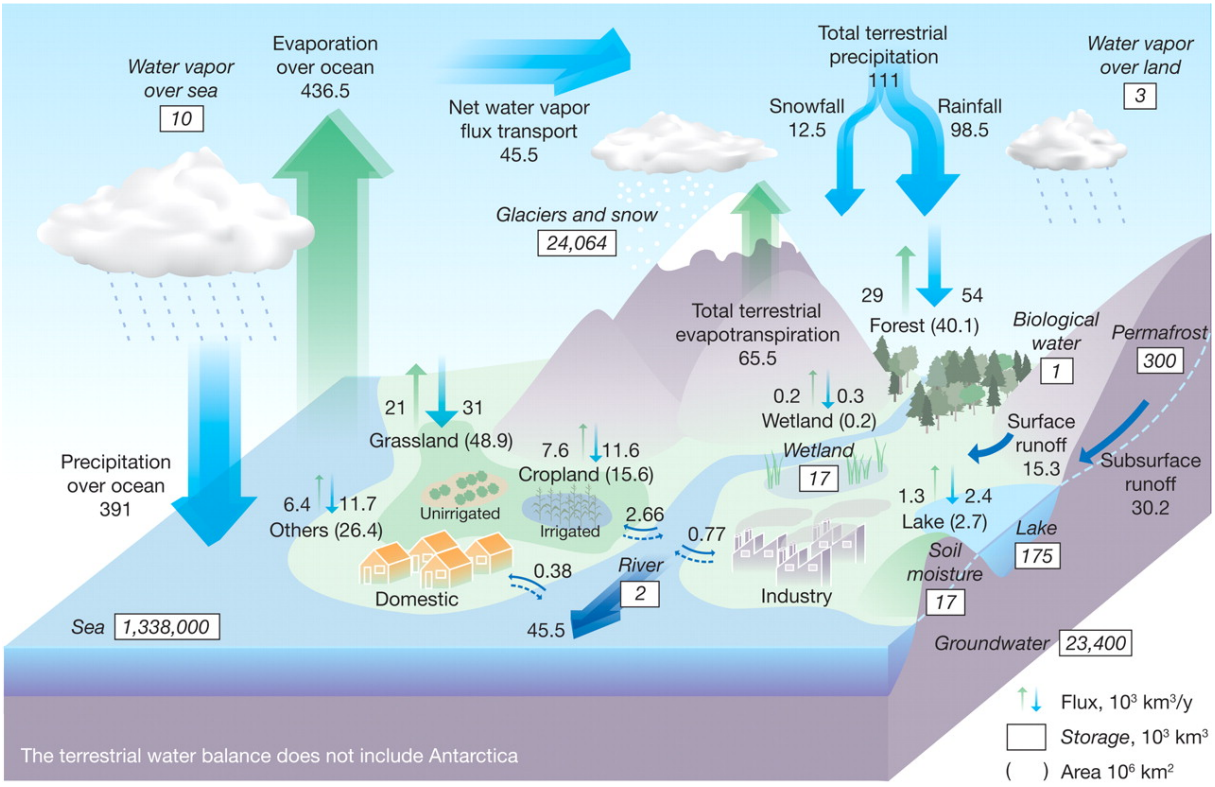


FIGURE 1.1 : Le cycle naturel et anthropique de l'eau, avec les flux (en $1{\times}10^3\,km^3\,an^{-1}$) indiqués par les flèches verticales et les volumes stockés (en $1{\times}10^3\,km^3$) par des rectangles. Les nombres entre parenthèses représentent les surfaces concernées en $1{\times}10^6\,km^2$. D'après Oki et Kanae (2006).

risquer de manquer un terme dans l'équation de conservation de la masse. Au contraire, représenter dans le détail chacune d'entre elles rendrait l'équation si complexe qu'elle en deviendrait insoluble.

1.1.1.b. Bassin versant

Comme toutes les sciences naturelles, l'hydrologie se propose de répondre aux questions qu'elle se pose par l'étude d'un système : le bassin versant. Le bassin versant est une unité géographique fermée hydrologiquement, c'est-à-dire définie de telle sorte qu'aucun écoulement n'y pénètre de l'extérieur, et que tout écoulement sortant passe par son exutoire. Un exutoire est une section droite d'un cours d'eau, parfois abusivement désigné comme un point si l'on considère un cours d'eau comme une ligne.

Un bassin versant est donc déterminé par la position de son exutoire. Déplacer un exutoire vers l'aval d'un cours d'eau revient à augmenter la taille du bassin versant considéré, le déplacer vers l'amont revient à la diminuer. En hydrologie de surface en principe, la frontière d'un bassin versant est délimitée par sa ligne de crête topographique, en supposant que l'eau qui alimente un cours d'eau provient du ruissellement de l'eau de pluie à la surface du sol. Néanmoins, un hydrogéologue considèrerait une telle définition comme une hérésie : le sous-sol d'un bassin versant topographique donné communique généralement avec celui du bassin versant topographique voisin (Figure 1.2). Les « lignes de crêtes piézométriques » n'ont pas de raison de se confondre avec les lignes de crêtes topographiques, et fluctuent d'ailleurs avec le temps (Le Moine, 2008). De fait, il est très difficile voire impossible de déterminer les frontières réelles d'un bassin versant. Par défaut, les hydrologues de surface considèrent cette complexité souterraine comme une source de « pertes » ou de « gains » du bassin versant topographique. Bien qu'elle soit imparfaite, c'est cette approche topographique qui est retenue dans cette thèse.

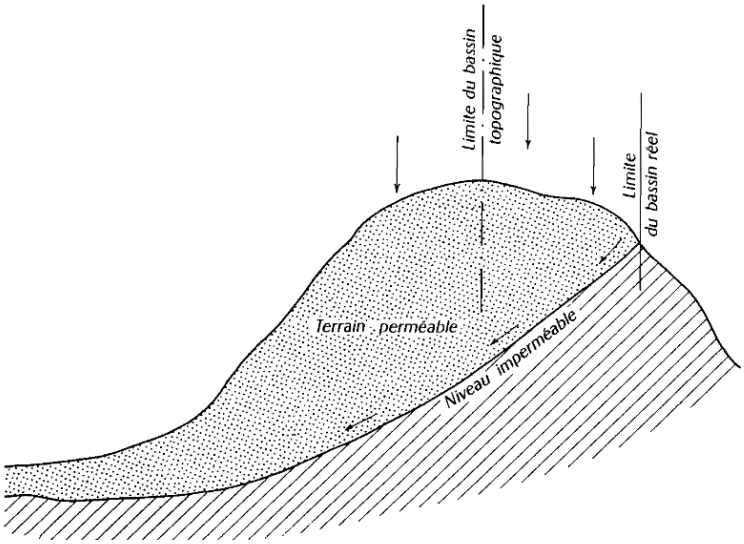


FIGURE 1.2 : Différenciation des limites du bassin versant topographique et de ses limites hydrogéologiques (bassin réel). Tiré de Roche (1963).

Les hydrologues écrivent donc l'équation du bilan de masse du bassin versant ainsi :

$$Q = P - E + \Delta S + L \tag{1.1}$$

Où Q représente le débit à l'exutoire, P le volume des précipitations, E la quantité évaporée dans l'atmosphère, et ΔS les variations de son « stock », et L le volume d'eau gagné ou perdu dans les échanges souterrains. L'unité de ces variables, généralement en mm mois^{-1} ou en mm an^{-1} dans l'équation, représente une lame d'eau répartie sur la surface du bassin versant accumulée sur une période de temps donnée. Les enjeux de la mesure de ces variables sont présentés brièvement dans la suite, dans un ordre croissant de difficulté d'observation.

1.1.1.c. Débit et précipitations

La manière la plus simple de connaître les précipitations sur un bassin versant consiste à recueillir la quantité d'eau tombée en un point dans un pluviomètre. Historiquement de simples seaux gradués totalisant le cumul des précipitations, les pluviomètres modernes permettent d'acquérir les mesures en temps réel et à des intervalles de temps de quelques minutes. Un unique pluviomètre suffit rarement à mesurer la pluie moyenne sur la surface d'un bassin versant, d'autant plus que celui-ci est grand et possède un fort relief. Pour disposer d'informations sur la distribution spatiale des précipitations, il est donc nécessaire de multiplier le nombre de ces points de mesure pour avoir un réseau aussi dense que possible. Si des méthodes sophistiquées de suivi du champ spatial des précipitations existent (radar, produits satellitaires), celles-ci nécessitent d'être étalonnées à partir de mesures au sol. Outre les limites liées à la densité du réseau pluviométrique, d'autres facteurs sont à l'origine d'incertitudes sur la mesure des précipitations à l'échelle d'un bassin, dont l'inventaire détaillé pourrait constituer un chapitre entier (Cosandey et Robinson, 2012). On peut citer pêle-mêle : positionnement et taille des pluviomètres, vent, évaporation, intensité des pluies, occurence de précipitations neigeuses, méthode d'intégration spatiale des données ponctuelles. Des mesures incertaines de la pluie peuvent avoir des conséquences néfastes pour la qualité des modélisations hydrologiques, notamment en ce qui concerne la sélection de paramètres « représentatifs » du bassin versant (Andréassian, 2002). Nous nous efforcerons d'intégrer les enjeux liés à la qualité des mesures de précipitations (et des autres variables) dans les développements à venir

dans la suite du manuscrit, bien que ce ne soit pas le sujet central du travail de thèse.

Le débit, généralement indiqué en $m^3\,s^{-1}$ ou en $mm\,j^{-1}$, s'obtient grâce à des jaugeages. Il existe plusieurs façons d'estimer un écoulement, choisies selon leur coût ou selon la taille du cours d'eau. Une méthode répandue consiste à mesurer les vitesses d'écoulement en différents points d'une section transversale du cours d'eau et d'en déduire le débit par intégration sur la section. Néanmoins, la majorité des mesures de débit utilisées par les hydrologues ne sont pas issues de mesures directes mais obtenues au moyen d'une relation hauteur-débit. Cette relation, appelée *courbe de tarage*, permet d'exploiter les mesures de hauteur d'un cours d'eau, qui sont simples à automatiser et peu coûteuses. Une courbe de tarage se calcule à partir de la régression d'un modèle sur une série de mesures ponctuelles de débit réalisées dans des conditions variées de hauteur d'eau. Sa simplicité en fait un outil très pratique pour les hydrologues, et a permis de produire un grand volume de séries de débit des cours d'eau français (e.g. Leleu *et al.*, 2014). Cette méthode est une source d'incertitudes supplémentaires aux erreurs de mesures de débit ou de hauteur d'eau. Ces incertitudes additionnelles sont liées à l'extrapolation de la courbe à des hauteurs non jaugées lors de crues ou d'étiages exceptionnels, ou encore, aux modifications de la morphologie du lit des cours d'eau (McMillan *et al.*, 2012).

1.1.1.d. Évaporation

Le terme d'évaporation de l'équation du bilan de masse correspond à la quantité d'eau vaporisée dans l'atmosphère sous l'action des contraintes météorologiques. L'évaporation totale dans un bassin versant correspond à la somme de la vaporisation de l'eau libre (gouttes à la surface des feuilles, flaques, lacs...) et de la transpiration des plantes à travers l'ouverture de leurs stomates. Il est donc fréquent de trouver l'évaporation continentale désignée par le terme « d'évapotranspiration ». Les questions terminologiques faisant l'objet de très nombreux débats au sein de la communauté scientifique dans lesquels ce travail n'a pas vocation à s'intégrer, nous nous tiendrons dans ce manuscrit au terme d'évaporation (Lhomme, 1997).
Bien que dans le bilan hydrique elle soit comptabilisée sous forme de volume ou de lame d'eau, l'évaporation peut aussi être considérée sous la forme d'un bilan d'énergie : c'est la partie du rayonnement net qui n'est pas utilisée pour réchauffer la première couche d'air (au-dessus) et la première couche de sol en-dessous. Il existe de nombreuses méthodes de mesure de l'évaporation (bilan d'énergie, covariance des turbulences, bac d'évaporation, lysimètre, scintillomètre...) ayant chacune leurs avantages et leurs inconvénients (Wang et Dickinson, 2012). Malgré ces nombreuses techniques, les données observées sont largement indisponibles à l'échelle des bassins versants en comparaison des données de précipitations et de débit : les chroniques de mesure disponibles sont soit concentrées sur certains sites spécifiques, soit trop imprécises aux échelles temporelles et spatiales d'étude du bassin versant, soit trop courtes pour être utilisables en entrée de l'équation du bilan de masse au même titre que les précipitations.

En modélisation hydrologique, il est très courant d'estimer l'évaporation effectivement produite (on parlera d'évaporation « réelle ») à partir d'une grandeur conceptuelle, l'évaporation potentielle (EP). Une description complète de cette notion est proposée par Oudin (2004), dont les principaux éléments sont brièvement résumés dans la suite. Thornthwaite (1948) introduit le premier l'évaporation potentielle comme « la quantité d'eau qui serait transpirée et évaporée si elle était disponible ». Les travaux agronomiques de Penman (1948) le conduisent la même année à proposer une formule pour calculer « l'évaporation d'une pelouse rase suffisamment étendue, en bon état et convenablement alimentée en eau ». Cette définition présente implicitement cette quantité comme un maximum à l'évaporation réelle observable sur une parcelle, avec l'hypothèse que les conditions atmosphériques et l'apport d'énergie restent constants au passage des conditions réelles de déficit

hydrique aux conditions potentielles de saturation. Depuis ces travaux fondateurs, la notion d'évaporation potentielle a fait l'objet de nombreuses discussions : critiques de l'hypothèse de conditions atmosphériques inchangées et rétro-actions entre évaporations réelle et potentielle (Bouchet, 1963; Morton, 1994), nature (eau libre ou sol végétalisé) et taille (suffisamment grande ou suffisamment petite) de la surface définissant l'évaporation potentielle (Shuttleworth, 1993; WMO, 1996), rôle de la température de surface du sol (Van Bavel, 1966; Priestley et Taylor, 1972). Ces débats ont donné naissance à de multiples définitions de l'évaporation potentielle, et des formulations mathématiques plus nombreuses encore, et suscité des travaux de synthèse tentant de rationaliser les différentes conceptions (e.g. Granger, 1989; McMahon *et al.*, 2013).

Quelle que soit la définition envisagée pour l'évaporation potentielle, elle n'est pas vraiment mesurable à l'échelle du bassin versant dans la mesure où les bilans d'énergie et les équations aérodynamiques ne sont pas valables à des échelles supérieures à quelques kilomètres (Brutsaert, 1986). Cette difficulté, associée au fait que d'un point de vue purement théorique, l'évaporation potentielle n'est pas mesurable dans des conditions non saturées (Lhomme, 1997), font de l'évaporation potentielle une grandeur purement fictive à l'échelle du bassin versant. Dans la mesure où ils ne sont pas falsifiables directement, les modèles d'évaporation potentielle utilisés en modélisation hydrologique ont pour but exclusif de fixer une limite supérieure à l'évaporation réelle et sont jugés à l'aune de la qualité des simulations de ces modèles hydrologiques (Oudin, 2004).

1.1.1.e. Variations de stock et échanges souterrains entre bassins

Au cours du transit de l'eau précipitée vers l'exutoire d'un bassin versant ou avant son évaporation, son temps de résidence au sein du bassin versant dépend fortement du chemin parcouru : de quelques minutes ou heures si elle est interceptée par la végétation ou si elle ruisselle le long des pentes, à plusieurs dizaines d'années si elle s'infiltre dans une nappe captive. De ce fait, les variations de la quantité d'eau stockée dans le bassin versant (surface, rivière en amont de l'exutoire, sol, et sous-sol) font intervenir nombre de processus, de zones d'influence et de temps caractéristiques très variés. Mesurer et tracer ces variations ne se fait donc pas de la même façon selon le pas de temps d'étude. À des échelles saisonnières ou inférieures, l'humidité du sol porte l'essentiel de ces variations, tandis qu'à des échelles pluri-annuelles celles-ci sont plutôt déterminées par l'évolution du niveau des nappes phréatiques. Dans tous les cas, des mesures des variations du stock d'eau sont rarement disponibles à l'échelle du bassin versant. Les mesures d'humidité du sol peuvent se faire par prélèvement direct d'un échantillon de sol (mesure contraignante, échelle spatiale caractéristique trop courte) ou par satellite (échelle spatiale caractéristique trop large, faible recul temporel). Les mesures des niveaux des nappes se fait par piezométrie, au niveau d'un puits. La topographie du sous-sol pouvant être très différente de celle rencontrée à la surface, l'interprétation de ces données nécessite de connaître l'extension spatiale de la nappe sondée. Face à ces difficultés, l'équation du bilan de masse est souvent abordée à l'échelle annuelle, en négligeant le terme des variations du stock (e.g. Zhang *et al.*, 2008). Pour certains bassins dominés par des dynamiques interannuelles, les variations de stock doivent néanmoins être prises en considération (Han *et al.*, 2020).

Une partie des eaux stockées dans les aquifères d'un bassin versant est également susceptible de circuler d'un bassin versant à l'autre sans passer par le réseau hydrographique. Ce phénomène d'échange souterrain est fréquent entre les sous-bassins amonts d'un bassin versant et les sous-bassins avals, vers lesquels l'eau s'écoule le long des lignes isopièzes. Certaines configurations géologiques rares peuvent également susciter des échanges importants entre bassins voisins, tels que les karsts où l'eau peut circuler en profondeur à travers des failles et des rivières souterraines. L'évaluation des échanges souterrains entre bassins versants est réalisable par l'étude des gradients piezométriques, qui nécessitent un réseau de mesure, ou dans certains rares cas par des

bilans de masse à l'exutoire de bassins voisins (e.g. Le Moine *et al.*, 2008). Le Moine (2008) note cependant l'impossibilité de connaître exactement le volume des échanges souterrains par l'étude des gradients piezométriques sans mesure de la perméabilité des sous-sols.

1.1.2. Introduction à la modélisation hydrologique

1.1.2.a. Qu'est-ce qu'un modèle ?

Comme mentionné précédemment, l'écoulement de l'eau à la surface des continents fait intervenir pléthore de processus, dont certains sont encore peu connus. L'hydrologie s'appuie donc en tant que science sur des représentations simplifiées de ces processus, c'est-à-dire de *modèles*, dont le caractère schématique témoigne de l'incapacité de l'entendement humain à saisir toute la complexité du réel, à toutes les échelles de temps et d'espace simultanément. Un modèle peut revêtir des formes variées : énoncé, schéma, carte, maquette, ou équation mathématique, et peut être employé dans différents contextes. La confrontation systématique des modèles et de leurs hypothèses associées aux mesures en laboratoire ou sur le terrain est à la base de la méthode scientifique utilisée en recherche fondamentale. Les modèles sont donc des outils servant à développer la compréhension que nous avons de notre environnement. Ils servent également à des fins prédictives ou comme outil de diagnostic de l'état d'un système dans des situations opérationnelles. Ces deux utilisations sont bien entendues fortement liées entre elles dans la mesure où une meilleure compréhension des systèmes réels diminue *a priori* le risque de formuler des prévisions très erronées, et qu'à l'inverse la comparaison des capacités prédictives de modèles pour calculer des grandeurs déjà connues permet d'éliminer les hypothèses inadéquates.

L'équation précédente de conservation de la masse (Équation 1.1) est un bon exemple de modèle hydrologique simple : il associe un ensemble de causes, en premier lieu desquelles les précipitations sur une surface délimitée, à une conséquence, l'écoulement en un point d'un cours d'eau. Le cadre descriptif qu'il offre du devenir des précipitations, entre écoulement et pertes par évaporation, stockage ou échange, en fait un outil très commode de classification des bassins versants, qu'il est par exemple possible de regrouper en catégories selon les rapports entre les flux. Toutefois, s'il donne un cadre pour appréhender simplement les principaux phénomènes à l'échelle du bassin versant, l'utilisation pratique du bilan de masse pour le calcul du débit est limitée par la difficulté de mesurer certaines des variables explicatives, comme l'évaporation réelle. Il est donc souvent nécessaire de formuler des hypothèses supplémentaires au sujet de certains processus pour être en mesure de se passer de données d'entrée indisponibles, ici par exemple en détaillant les interactions entre le sol et l'atmosphère pour quantifier l'évaporation réelle. À l'inverse, si des données de débit sont disponibles, le modèle de conservation de la masse peut être utilisé pour calculer une donnée d'entrée manquante.

L'exemple de l'équation de conservation de la masse illustre l'interaction forte qui existe entre la formulation d'un modèle et le contexte de son utilisation. Selon Perrin (2000), le développement d'un modèle hydrologique est conditionné par plusieurs éléments : la délimitation du système modélisé dans un cadre spatial et temporel, la mesure de flux internes et externes du système, l'objectif de modélisation, et le choix d'une formulation de la réalité. La palette extrêmement variée de conditions d'utilisation de modèles hydrologiques motive des approches toutes aussi variées pour répondre de façon aussi pertinente que possible aux besoins spécifiques de chacune d'entre elles. Ainsi, de nombreux modèles co-existent en modélisation hydrologique, fruits d'arbitrages entre généralité, réalisme, et précision, imposés par le contexte (Kauark-Leite et Nascimento, 1993, cité par Perrin, 2000).

1.1.2.b. Enjeux de la modélisation hydrologique

Outre leur utilisation à des fins de pure recherche fondamentale, les modèles hydrologiques répondent à des problématiques liées au rôle central qu'occupe l'eau dans nos sociétés. L'eau continentale constitue en effet à la fois une ressource indispensable qu'il convient de réguler, et un aléa lors des crues et étiages contre lequel il est nécessaire de se protéger. Dans ce contexte, l'emploi de modèles hydrologiques peut satisfaire des objectifs variés, généralement regroupés comme suit :

- Reconstitution de débits passés, pour combler des lacunes dans des mesures de débit ou pour étendre des chroniques existantes à partir de mesures de précipitations (e.g. Devers, 2019) ;
- Prévision de débits futurs, par exemple pour anticiper le risque de crues ou d'étiages ou bien optimiser la gestion de retenues d'eau, à des horizons temporels allant de quelques heures (e.g. Pagano *et al.*, 2014; Ficchì *et al.*, 2016) à quelques mois (e.g. Boucher *et al.*, 2018) ;
- Prédétermination des écoulements, c'est-à-dire caractérisation de la fréquence à laquelle des événements peuvent se produire. La modélisation peut ainsi servir à estimer la période de retour d'étiages ou de crues extrêmes définis par rapport à un certain seuil, ou inversement à estimer l'ampleur d'un événement de période de retour donnée, afin de dimensionner des ouvrages hydrauliques (e.g. Garavaglia, 2011) ;
- Caractérisation de non-stationnarité du régime hydrologique, liés à des modifications anthropiques ou naturelles. La modélisation peut par exemple permettre d'évaluer les effets de changement du couvert végétal par déforestation ou urbanisation (e.g. Andréassian, 2002; Sterling *et al.*, 2013), ou les conséquences régionales du changement global (e.g. Chiew *et al.*, 2009). Les recherches menées au cours du doctorat concernent les modèles utilisés dans cette dernière catégorie d'objectifs, c'est-à-dire en « projection » (Beven et Young, 2013).

1.1.2.c. Structure d'un modèle hydrologique

De nos jours, les modèles numériques constituent la très large majorité des modèles utilisés en hydrologie, dans la mesure où les capacités de calcul des ordinateurs permettent de résoudre des équations complexes dont la résolution « à la main » prendrait des années. Les modèles modernes sont capables de manipuler de très grands jeux de données dans des calculs beaucoup plus complexes que ceux du modèle de conservation de la masse. Malgré leur diversité, les modèles hydrologiques sont construits sur des bases communes, schématisées en Figure 1.3.
Les variables d'entrée d'un modèle hydrologique correspondent à l'ensemble des chroniques de données dont le modèle a besoin pour fonctionner et calculer les variables de sortie. Le modèle est « forcé » par les variables d'entrée (ex : les précipitations). Les variables de sortie sont les variables qui intéressent le modélisateur (ex : le débit). Au cours des calculs, le modèle peut faire intervenir des variables intermédiaires, nécessaires pour le calcul des variables de sortie. Ces variables intermédiaires sont désignées comme les variables d'état (ex : l'humidité du sol). Deux états différents du modèle peuvent aboutir à des variables de sortie différentes à partir des mêmes variables d'entrée.
L'utilisation d'un même modèle dans des contextes différents peut motiver un ajustement de ses équations pour prendre en compte les différences entre les caractéristiques physiques des milieux d'application. Les paramètres du modèle sont des constantes intervenant dans les équations dont le modélisateur choisit la valeur pour s'adapter aux bassins versants étudiés (ex : la porosité du sol). Il convient de noter que la signification physique d'un paramètre n'est pas nécessairement sans équivoque, et varie selon les modèles. Certains paramètres nécessitent d'être calés, c'est-à-dire estimés par une optimisation des performances du modèle, parce qu'il n'est pas possible de déterminer une valeur *a priori* à partir des caractéristiques physiques.
La performance d'un modèle pour représenter un système peut être jugée au regard de critères mathématiques (parfois géométriques ou visuels) qui comparent les variables de sortie simulées à des

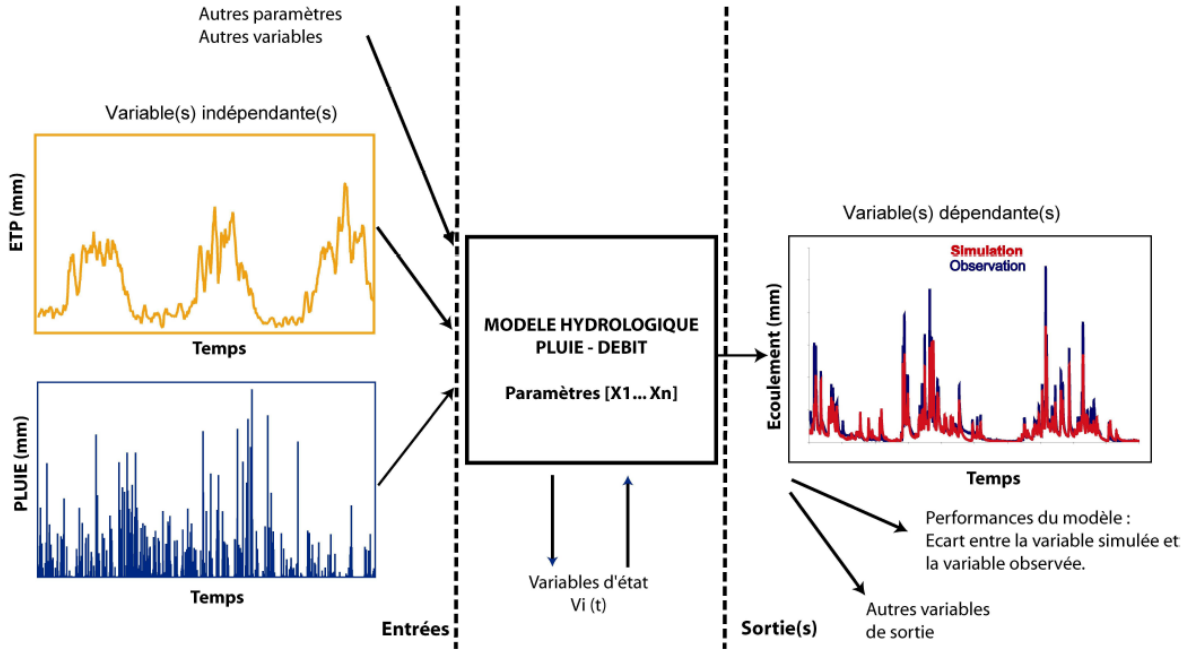


FIGURE 1.3 : Représentation du fonctionnement d'un modèle hydrologique. Tiré de Mathevet (2005)

observations. Les mesures de l'erreur permettent de diagnostiquer les modèles, de les comparer entre eux, ou bien de quantifier l'incertitude autour d'une simulation. Les manières de calculer les performances d'un modèle sont infinies, et le choix d'une méthode dépend à la fois des observations disponibles, des objectifs de modélisation, du confort d'utilisation, et de l'imagination du scientifique.

Il convient de noter que les paramètres d'un modèle peuvent être calés de façon manuelle ou automatique. Un calage manuel repose sur l'expertise du modélisateur, qui par une succession d'essais cherche à identifier les paramètres idéaux en inspectant visuellement les variables simulées ou en calculant divers indices de performance. Si elle peut permettre d'obtenir de bons résultats, cette méthode est longue et requiert une expertise, ce qui la rend difficilement applicable (Wagener et al., 2003). Le calage automatique consiste en l'optimisation d'une fonction de coût, appelée fonction objectif, qui reflète les performances du modèle. Au contraire du calage manuel, il a l'avantage d'être rapide et reproductible. Cependant, les résultats dépendent du choix de la fonction objectif, qui favorise et néglige nécessairement certains aspects du régime hydrologique simulé.
Dans tous les cas, le choix des paramètres d'un modèle dépend également de l'efficacité de l'algorithme servant à leur optimisation (un cerveau humain dans le cas d'un calage manuel, un algorithme informatique dans le cas d'un calage automatique), et du temps accordé à la recherche de l'optimum. Si l'optimisation n'est pas bien réalisée, le jeu de paramètres obtenu peut être sous-optimal, dans la mesure où il existe un autre jeu de paramètres, plus performant mais non identifié au calage. Le risque de sous-optimalité augmente exponentiellement avec le nombre de paramètres libres d'un modèle. Les enjeux liés à cette problématique seront abordés plus en détail au Chapitre 2.

1.1.3. Approches de modélisation en hydrologie

1.1.3.a. Principes généraux

Deux conceptions opposées orientent généralement le développement de modèles en science, et l'hydrologie n'y fait pas exception. La première suggère de modéliser un système complexe comme la somme de ses parties, par l'agrégation d'organes de petite échelle dont les caractéristiques physiques et le comportement sont théoriquement connus. Les sous-systèmes de cette catégorie de modèles pourraient d'ailleurs être considérés comme des systèmes à part entière, en interaction les uns avec les autres. C'est la raison pour laquelle cette approche est appelée ascendante, agrégative, mécaniste, ou réductionniste (*bottom-up* en anglais). En hydrologie, elle a naturellement vocation à valoriser les connaissances acquises dans plusieurs disciplines comme la mécanique, la géologie, l'agronomie, et la biologie, en fragmentant un bassin versant en sous-unités géographiques homogènes. Par construction, elle cherche à représenter la complexité d'un système de façon exhaustive et plausible.
Au contraire de cette approche agrégative, la deuxième conception considère un système réel comme un tout, irréductible à la somme de ses parties. En opposition à l'approche précédente, elle est appelée descendante, désagrégative, empirique, conceptuelle ou systémique (*top-down* en anglais). Selon Klemeš (1983), les modèles ainsi construits « commencent en tentant de faire émerger une vision conceptuelle du système à l'échelle d'intérêt, puis l'affinent en identifiant les procédés des échelles inférieures qui pourraient la justifier ». Cette approche est en général adoptée dans des situations opérationnelles, où la performance du modèle à simuler les variables d'intérêt prime sur son réalisme physique.

Ces deux approches ont chacune leurs avantages et leurs inconvénients, et la décision d'adopter l'une ou l'autre dépend évidemment du contexte de modélisation, des objectifs, des savoirs préalables et des données disponibles (Hrachowitz et Clark, 2017). Elles appellent généralement des discrétisations spatiales différentes, dans la mesure où la recherche de la frugalité guide la démarche de l'approche descendante.
On appelle modèles globaux les modèles qui ne tiennent pas compte explicitement de la variabilité spatiale des données d'entrée météorologiques et des caractéristiques du bassin versant, et modèles distribués ceux où les processus hydrologiques sont explicitement résolus en des points du bassin versant en fonction des forçages météorologiques et des caractéristiques physiques locales (Lobligeois, 2014). Le découpage d'un bassin versant peut être réalisé selon des schémas variés. En général, les modèles distribués désignent les découpages géométriques fondés sur des mailles rectangulaires ou triangulaires. Les modèles qui divisent un bassin versant en sous-bassins versants, par exemple ceux des affluents du cours d'eau étudié, sont appelés semi-distribués. D'autres techniques de discrétisation reposent également sur l'identification de zones au comportement hydrologique homogène à travers des considérations sur la végétation, le type de sol et le sous-sol, afin de réduire la prise en compte d'hétérogénéités n'ayant pas d'incidence sur la génération des écoulements. Les zones ainsi délimitées sont qualifiées d'unités hydrologiques représentatives (*HRU* en anglais).

Il convient de noter que tous les modèles hydrologiques conceptuels ne sont pas nécessairement globaux, et que les grandes familles de modèles hydrologiques sont issues d'approches qui dépassent le simple cadre général mentionné précédemment. Les paragraphes suivants dressent une liste succinctement détaillée (et certainement non-exhaustive) de ces familles, à partir de la description qu'en fait Perrin (2000). Dans l'ordre croissant de réalisme physique, on distingue les modèles dits « boîte noire », les modèles conceptuels « à réservoir » et les modèles « à base physique ». Les modèles « boîte noire » relient directement les entrées météorologiques au débit par des relations mathématiques sans fondement physique. Ce sont des modèles globaux, développés dans le seul

but de répondre efficacement à une problématique opérationnelle, généralement de prévision des débits à des échéances courtes. De ce fait, ils sont rarement adaptés pour simuler le régime hydrologique d'un bassin versant et ne s'intègrent donc pas dans les objectifs du travail de thèse. Ils ne sont pas détaillés dans la suite.

1.1.3.b. Modèles conceptuels « à réservoir »

Les modèles à réservoir, souvent simplement appelés modèles conceptuels, découlent d'une vision simplifiée des processus qui transforment les précipitations en débit au sein du bassin versant. Ces processus sont généralement globaux ou semi-distribués. L'utilisation de réservoirs pour représenter le cheminement et le stockage de l'eau dans le système leur permet en théorie de modéliser l'intégralité du régime hydrologique en gardant la mémoire des conditions antérieures. Les processus representés ne correspondent toutefois pas clairement à une réalité physique, ce qui peut rendre leur interprétation ambiguë.

On identifie deux composantes principales dans la plupart des modèles conceptuels :

- Le module de production est constitué des processus associés au bilan en eau. Il divise les précipitations brutes en précipitations nettes (qui participent à l'écoulement) en calculant les flux externes du bassin versant modélisé et les variations des variables d'état associées (évaporation, échanges souterrains, stockage). Son rôle est d'autant plus prépondérant que l'échelle temporelle d'étude est longue.
- Le module de routage est constitué des processus assurant la répartition temporelle de la quantité d'eau transitant dans le cours d'eau. À l'inverse du module de production, son rôle est d'autant plus prépondérant que l'échelle temporelle d'étude est courte.

Il est rare que les modules de production et de routage soient complètement dissociés dans la structure d'un modèle, et un même réservoir peut intervenir à la fois dans des processus de production et de routage.
Les premiers modèles à réservoirs étaient très complexes et possédaient de nombreux paramètres libres. Les modèles actuels sont en règle générale plus simples pour diminuer le nombre de dimensions en calage des paramètres. Parmi les nombreux modèles conceptuels existants, on peut citer notamment Topmodel (Beven et Kirkby, 1979), HBV (Bergström, 1976, 1995) ou GR4J (Perrin *et al.*, 2003).
Les modèles à réservoir ne représentent pas l'ensemble des processus connus impliqués dans la transformation des précipitations en écoulement, et se concentrent sur des processus a priori considérés comme dominants. Néanmoins, Wheater *et al.* (1993, cité par Perrin, 2000) note que la connaissance de ces mécanismes prépondérants dans un bassin versant suppose une analyse très détaillée du terrain, qu'il est très difficile de mener en pratique. En outre, au-délà de la difficulté technique et économique d'observer l'ensemble de ces mécanismes, il n'est pas garanti qu'un modèle construit à partir de processus spécifiques adaptés à une certaine catégorie de bassins versants soit utilisable sur un bassin versant choisi au hasard. Perrin (2000) suggère donc qu'il est préférable que le développement de la modélisation conceptuelle soit orienté par une recherche de généralité, afin d'être capable de simuler une large gamme des conditions sur des bassins très variés (voir aussi Gupta *et al.*, 2014).

1.1.3.c. Modèles « à base physique »

Au contraire des modèles conceptuels qui se concentrent sur la représentation du débit à l'exutoire, les modèles fondés sur la physique font intervenir des principes physiques théoriques pour décrire l'écoulement dans le bassin. Ces modèles sont très souvent distribués, afin de résoudre les équations aux dérivées partielles de propagation de l'eau, telles que celles de Saint-Venant dans le lit du

cours d'eau, ou celles de Boussinesq ou Richards dans les sols saturés ou non-saturés.

L'intérêt premier de cette représentation est qu'elle permet de représenter de nombreux processus liés de façon directe ou indirecte à l'écoulement à l'exutoire du bassin versant. Les liens de causalité entre ces processus et les variations des variables d'états représentées sont scientifiquement établis et garantissent a priori la capacité du modèle à simuler des situations de non-stationarité des forçages (Blöschl *et al.*, 2013). De même, le paramétrage des équations repose sur la connaissance des caractéristiques du milieu, et peut être modifié si celles-ci viennent à changer. L'exhaustivité et le réalisme des processus autorise enfin l'emploi de ces modèles pour des objectifs autres que le simple calcul du débit (e.g. Sterling *et al.*, 2013). Les modèles de surface continentale sont par exemple utilisés en couplage avec les modèles atmosphériques dans les modèles climatiques, ce qui leur permet d'intégrer des boucles de rétroaction entre les deux systèmes. Bien que la simulation du débit à l'exutoire des bassins versants soit parfois un objectif secondaire, ces modèles sont toutefois qualifiés d'hydrologiques dans la mesure où ils peuvent être employés à ce dessein. Le modèle ORCHIDEE (Krinner *et al.*, 2005) est un exemple de modèle de surface continentale.
D'autres approches sont issues des modèles de surface continentale mais témoignent d'un effort plus ou moins prononcé de simplification. C'est pourquoi on appelle sobrement cette catégorie de modèles « à base physique ». Ils intègrent parfois des réservoirs pour simplifier la résolution des équations, ce qui peut les rapprocher des modèles conceptuels. La distinction avec ces derniers se joue toutefois sur l'interprétation physique des modules de ces modèles, qui ont vocation à représenter des processus et des éléments clairement identifiables du bassin versant. Parmi les modèles à base physique, on peut citer par exemple ISBA-MODCOU (Habets *et al.*, 2008), J2K (Krause, 2001, 2002) ou encore mHM (Samaniego *et al.*, 2010; Kumar *et al.*, 2013).

Bien que la philosophie de modélisation « à base physique » semble la plus satisfaisante en théorie, leur utilisation en pratique n'est généralement pas aussi simple que celle des modèles conceptuels. En effet, la fragmentation du bassin versant en sous-unités requiert une grande quantité de mesures des caractéristiques physiques du terrain qui ne sont pas toujours disponibles. Face à ces situations, elle suppose des valeurs de paramètres parfois éloignées des caractéristiques réelles des processus, voire un recours à un calage qui rend caduque son interprétabilité physique. Par ailleurs, les restrictions liées au temps de calcul conduisent souvent à adopter une discrétisation spatiale trop large pour que les unités spatiales soient véritablement considérées comme homogènes. Les processus d'échelles plus fines que ces unités de calcul doivent dans ce cas être paramétrés. Enfin, certains processus importants demeurent encore mal connus ou difficiles à généraliser en dehors du cadre maîtrisé d'un laboratoire (Clark, 2018)[1].

Quelle que soit l'approche adoptée dans le développement et la formalisation d'un modèle hydrologique, on peut estimer qu'il doit posséder un certain nombre de qualités indispensables à une utilisation scientifique rigoureuse. À ce titre, Lindström *et al.* (1997, cité par Perrin, 2000) proposent une liste de règles que n'importe quel modèle devrait respecter :

- les fondements scientifiques du modèle doivent être sains ;
- les besoins en données doivent être en accord avec leur disponibilité sur la majorité des bassins ;
- la complexité du modèle doit être justifiée en termes de performance ;
- le modèle doit être convenablement validé ;
- le modèle doit être compréhensible.

Les études des conséquences du changement climatique sur le régime hydrologique des cours d'eau s'appuient sur des modèles hydrologiques issus des deux approches présentées. La section

[1] Clark, M. 2018, « Advances in continental-domain hydrologic modeling and prediction », Global Institute for Water Security (GIWS), University of Saskatchewan, `https://youtu.be/tt-xCJD42CA`

suivante présente les principaux résultats de ces études, ainsi que les enjeux et les attentes vis-à-vis des modèles hydrologiques utilisés.

1.2. Modéliser les conséquences du changement climatique en hydrologie

Il est aujourd'hui avéré que les activités humaines influencent significativement le système climatique. Le rejet de gaz à effet de serre tels que le CO_2 ou le CH_4, modifie l'équilibre radiatif terrestre et conduit à une augmentation de la température moyenne à l'échelle du globe. Cette augmentation est à l'origine de l'emploi du terme « réchauffement global » pour désigner le changement climatique initié au début de l'ère industrielle (Pachauri *et al.*, 2014).
Cette section présente dans un premier temps les principaux effets attendus de ce réchauffement sur le cycle de l'eau à l'échelle globale et les risques associés pour les sociétés humaines. La suite décrit succinctement les méthodes permettant d'estimer ces effets et les incertitudes associées notamment en ce qui concerne les modèles hydrologiques. Le propos se concentre ensuite sur la notion de robustesse climatique des modèles hydrologiques et sur les enjeux liés aux choix de modélisation sur lesquels le travail de thèse se fonde.

1.2.1. Enjeux et attentes vis-à-vis de la modélisation

1.2.1.a. Conséquences hydrologiques du changement climatique

Le changement climatique devrait avoir de nombreuses conséquences sur l'hydrosphère et sur le cycle de l'eau. Celles présentées dans la suite sont toutes issues du rapport n°5 du Groupe d'Experts Intergouvernemental sur l'Évolution du Climat (GIEC, Pachauri *et al.*, 2014), et du rapport spécial sur la cryosphère (Pörtner *et al.*, 2019).

La hausse des températures conduit notamment à l'intensification des processus de fonte de neige et de glace. Il est ainsi prévu que la fonte des glaciers et des calottes polaires participe fortement à l'élévation du niveau des océans. Le régime hydrologique des cours d'eau nivaux ou alimentés par des glaciers va également en subir les conséquences, avec une augmentation du débit hivernal, et des pics de fonte avancés dans l'année. Leur débit moyen devrait passer par un pic au cours du XXIe siècle dans toutes les régions du monde. Ces changements sont susceptibles d'affecter la disponibilité de la ressource en eau dans les zones de montagnes, utilisée pour l'hydroélectricité ou l'irrigation, ainsi que les risques naturels comme les crues et les glissements de terrain.

Les modifications induites par le changement climatique sur les précipitations ne seront pas uniformes. L'augmentation des températures devrait intensifier le cycle de l'eau à l'Équateur et conduire à une augmentation des pluies. En ce qui concerne les régions de latitude moyenne, les hétérogénéités spatiales actuelles sont suspectibles d'être accentuées. La variabilité des précipitations dans les régions sous l'influence de la mousson ou du phénomène El-Niño devrait également augmenter. Enfin, il prévu que les événements extrêmes de précipitations deviennent à la fois plus intenses et plus fréquents sur une grande partie du globe.

Ces changements feront courir à la population des risques majeurs en augmentant l'insécurité d'approvisionnement en eau de nombreuses communautés, en détruisant des infrastructures essen-

tielles lors des événements météorologiques extrêmes ou de crues dans les zones urbaines. Il est néanmoins possible de diminuer ces changements attendus, en réduisant les émissions anthropiques de gaz à effet de serre, ou bien de limiter leurs conséquences en adaptant nos sociétés face à l'émergence de ces risques. Ces deux leviers d'action requièrent une estimation précise et fiable des changements aux échelles caractéristiques des décisions politiques ou économiques. Une évaluation convenable des changements du régime hydrologique aidera notamment à dimensionner les infrastructures nécessaires aux politiques d'adaptation. Ces estimations s'appuient pourtant sur une chaîne de modélisation qui comporte de nombreuses sources d'incertitude.

1.2.1.b. Chaînes de modélisation et incertitudes

Les projections hydrologiques sont réalisées à partir de simulations successives de modèles, de l'échelle spatiale du globe à l'échelle du bassin versant. Ces modèles forment donc une chaîne, dont chaque maillon est indispensable pour aboutir à la simulation du débit futur (Figure 1.4).
La première étape consiste en l'élaboration de scénarios d'émission de gaz à effet de serre, qui illustrent chacun des trajectoires différentes d'augmentation des émissions au niveau mondial, de la plus volontariste à la plus laxiste. Les scénarios adoptés dans les travaux du cinquième rapport du GIEC sont appelés *Representative Concentration Pathways* (RCP). Ces trajectoires d'émission servent à contraindre les modèles de circulation générale (GCM) qui constituent le deuxième maillon de la chaîne de modélisation. Les GCM ont pour but principal d'estimer les effets des modifications du forçage radiatif sur les phénomènes climatiques de large échelle. Ces modèles extrêmement complexes ne peuvent pas modéliser les variables météorologiques aux échelles caractéristiques des bassins versants, essentiellement pour des raisons de temps de calcul. Le troisième maillon de la chaîne a pour rôle d'opérer cette descente d'échelle, soit par des méthodes statistiques (SDM), soit par la mise en place de modèles climatiques régionaux (RCM) associés parfois à des méthodes de corrections de biais. Les RCM peuvent être utilisés comme modèles hydrologiques dans la mesure où ils intègrent généralement une modélisation couplée des surfaces continentales et de l'atmosphère. Néanmoins, le calcul du débit est généralement réalisé par un modèle hydrologique utilisé comme quatrième et dernier maillon de la chaîne, forcé par les variables météorologiques estimées à l'étape de descente d'échelle.

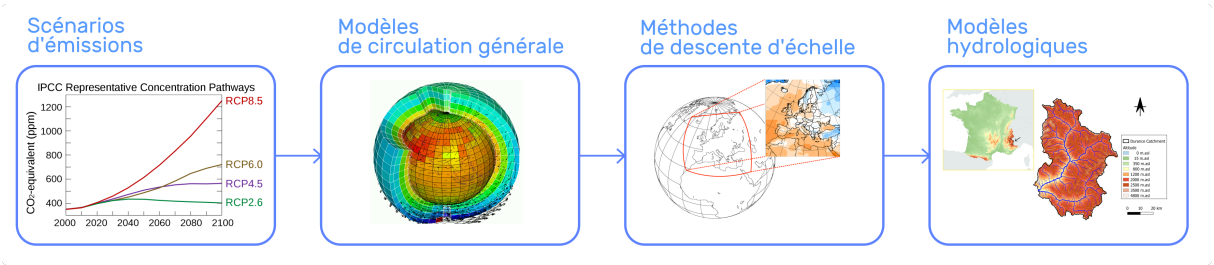


FIGURE 1.4 : Chaîne de modélisation utilisée en projection hydrologique.

Cette chaîne de modélisation fournit des projections hydrologiques marquées généralement par de larges incertitudes, émanant de plusieurs sources.
La première source d'incertitude est inhérente à l'exercice de projection climatique, dans la mesure où elle relève de l'impossibilité de prédire les émissions de gaz à effet de serre. Contrairement aux sources suivantes d'incertitude, celle-ci ne provient pas d'une erreur de modélisation à proprement parler. Les scénarios RCP matérialisent des trajectoires d'émissions considérées comme plausible, pour lesquelles il convient de déterminer *a posteriori* les décisions politiques, économiques et sociétales futures qui les conditionnent.
La deuxième source d'incertitude concerne la variabilité interne du climat. Pour comprendre cette notion, il convient de rappeler que le climat est une description statistique des variables météoro-

logiques, décrites comme des variables aléatoires dont on étudie les caractéristiques moyennées en général sur des périodes de 30 ans (Pachauri *et al.*, 2014). Entre l'échelle temporelle caractéristique typique de la météorologie (quelques jours) et du climat (quelques décennies), ces variables peuvent décrire des variations liées au fait que certains événements météorologiques laissent une trace durable à cause de l'inertie du système climatique (Liu, 2012). Ces variations peuvent être aléatoires, ou bien s'inscrire dans des schémas temporels et spatiaux réguliers, tel que le phénomène El-Niño (Trenberth, 1997). Le volume des précipitations qui sera effectivement mesuré en un lieu pendant plusieurs décennies futures pourra être supérieur à celui prévu par les modèles, qui est simplement le volume des précipitations ayant le plus de chance d'être observée. Le rôle de l'incertitude liée à la variabilité interne du climat est prépondérant dans les premières décennies du XXIe siècle et diminue à mesure que les écarts entre les climats résultant des différentes trajectoires d'émission s'accentuent (Hingray et Saïd, 2014).
La dernière source d'incertitude est enfin liée directement aux erreurs et aux différences entre les prédictions des modèles numériques. Celles-ci peuvent par exemple être liées à la paramétrisation des processus physiques ou à la discrétisation temporelle et spatiale choisie. La question de la contribution de chaque maillon de la chaîne de modélisation à cette incertitude est un champ de recherche actif. Certaines études soutiennent par exemple que les modèles climatiques ou les méthodes de descentes d'échelle produisent la majeure partie de l'incertitude (e.g. Dobler *et al.*, 2012; Lafaysse *et al.*, 2014), à l'exception des bassins influencés par la neige pour lequels les modèles hydrologiques sont également très incertains (e.g. Addor *et al.*, 2014; Giuntoli *et al.*, 2015). D'autres travaux font valoir le rôle majeur joué par les modèles hydrologiques dans l'incertitude totale des projections dans de nombreuses régions (e.g. Hagemann *et al.*, 2013; Schewe *et al.*, 2014), pour la simulation des étiages (Vidal *et al.*, 2016) ou pour celle du manteau neigeux (Melsen *et al.*, 2018). L'attribution des sources d'incertitudes varie donc fortement d'un contexte à l'autre. Il convient de noter que la non-linéarité des interactions entre les maillons de la chaîne de modélisation participe à la multiplication des sources d'incertitudes entre elles (Bosshard *et al.*, 2013).

Il convient de noter que malgré ces limites, les principales conclusions exposées précédemment relèvent de consensus scientifiques forts (Pachauri *et al.*, 2014; Pörtner *et al.*, 2019). L'incertitude quant aux changements à venir est principalement liée aux échelles locales ou à certaines zones du globe où la variabilité naturelle est très forte (Knutti et Sedláček, 2013). Dans cette perspective, il apparaît essentiel de chercher à réduire l'incertitude et à garantir la fiabilité des modèles hydrologiques utilisés en projection. Cette recherche nécessite l'évaluation de la capacité des modèles à représenter les écoulements dans les bassins versants, tant pour l'évaluation de la ressource que pour la simulation des événements extrêmes.

1.2.2. Robustesse d'un modèle hydrologique en climat non-stationnaire

1.2.2.a. Définition de la robustesse

Malgré les incertitudes fortes qui planent encore sur certaines conséquences du changement global, les conditions climatiques risquent de changer drastiquement dans une grande partie du globe au cours du XXIe siècle. La question de savoir si les modèles utilisés en projection sont capables de représenter les systèmes hydro-climatiques est dans ce contexte absolument cruciale. Cette capacité fait appel à la notion de « robustesse » du modèle. Avant de la détailler en contexte de changement climatique, il est nécessaire d'expliquer à quoi cette notion correspond en modélisation.

En français, la notion de robustesse renvoie à celles de vigueur, de solidité, ou de résistance, face aux altérations que pourraient causer des contraintes extérieures. En algorithmique, la robustesse d'un programme informatique est la capacité d'un programme à garder un comportement « accep-

table » dans des conditions « exceptionnelles ou imprévues » (Fernandez *et al.*, 2005). En modélisation, la robustesse est liée à l'évaluation de la perte de performances du modèle entre des conditions nominales et des conditions non-nominales, pour lesquelles il n'a pas été spécifiquement conçu ou même prévu d'être utilisé. Il apparaît donc essentiel de déterminer avant tout quelles sont les conditions nominales d'utilisation d'un modèle hydrologique.
Un modèle étant développé à partir des connaissances et des constats formulés sur le fonctionnement d'un bassin versant, ses équations n'ont de pertinence a priori que dans le cadre spatial et temporel dans lequel ces données ont été produites. La simulation des débits à une période donnée sur l'ensemble des bassins versants sur lequel le modèle a été construit constitue le cadre strictement nominal de son utilisation. Dans le cas d'un modèle disposant de paramètres libres, le calage de ces paramètres sur les données d'un bassin ou d'une période n'ayant pas été utilisées pour son développement permet malgré tout de maximiser ses performances en dimensionnant ses équations. Cette situation place le modèle dont les paramètres sont calés dans des conditions plus favorables pour fonctionner qu'un modèle dont les paramètres ne seraient pas calés. On peut donc considérer que caler un modèle sur des données n'ayant pas été utilisées pour son développement le place dans des conditions de simulation presque nominales.
Les conditions non-nominales de simulation correspondent aux situations restantes, c'est-à-dire dans des bassins et/ou à des périodes qui ne font pas partie de l'échantillon de développement du modèle, et/ou pour lesquels ses paramètres n'ont pas fait l'objet d'un calage. Simuler le débit d'un bassin sur une période B différente de celle du calage des paramètres (période A) est un exemple de situation non-nominale. Les performances ainsi obtenues seront nécessairement moins bonnes que celles obtenues directement en calage sur la période B. S'il est suspecté qu'entre la période A et la période B, des perturbations ont modifié le régime hydrologique du bassin versant, il est probable que la chute de performance sera plus forte que si le régime était resté inchangé. Cet exemple montre que plusieurs conditions non-nominales peuvent perturber la qualité des simulations d'un modèle à des degrés divers. Plus un modèle hydrologique est robuste, plus les conditions doivent être défavorables pour observer une dégradation de ses performances par rapport à la situation nominale.

Dans le cadre des projections hydrologiques, les perturbations susceptibles d'altérer les performances d'un modèle peuvent être causées par des modifications du climat (régime des précipitations, température), de l'usage des sols (urbanisation, déforestation, etc.), ou de régulation des cours d'eau (barrages, prélèvements, etc.). Nous désignerons comme robustesse climatique la robustesse des modèles face aux changements climatiques. La robustesse climatique des modèles peut aussi être envisagée comme leur capacité à extrapoler le régime hydrologique en dehors des conditions climatiques connues. Les questions de l'usage des sols et de la régulation des cours d'eau ne sont pas traitées dans le cadre de cette thèse, bien que ces phénomènes aient une influence importante sur les débits passés et futurs (e.g. Sterling *et al.*, 2013; Lauri *et al.*, 2012).

1.2.2.b. Évaluation de la robustesse climatique

Il n'est pas possible d'évaluer directement la robustesse climatique des modèles hydrologiques face aux changements climatiques à venir dans la mesure où les conditions futures ne sont pas observables (Refsgaard *et al.*, 2014). Les archives paléohydrologiques disponibles sont utilisées en reconstruction des conditions passées mais pas pour l'évaluation de modèles hydrologiques. De ce fait, l'estimation de la robustesse des modèles ne peut se faire qu'avec les mesures de l'ère moderne, c'est-à-dire dans les mêmes conditions climatiques qui ont servi à leur développement. L'exigence des tests qu'il est possible de déployer pour analyser la capacité d'extrapolation climatique des modèles se heurte donc nécessairement à certaines limites, qui seront abordées après la présentation des méthodes couramment utilisées pour évaluer la robustesse des modèles hydrolo-

giques.

Dans ses travaux visionnaires sur l'évaluation opérationnelle des modèles hydrologiques, Klemeš (1986) propose quatre méthodes de difficultés croissantes pour tester la robustesse des modèles hydrologiques :

- Le *Split-sample test* (SST) consiste en la division de la chronique des données observées d'un bassin en deux parties indépendantes. La première partie est utilisée dans un premier temps pour le calage des paramètres du modèle et la deuxième pour l'évaluation (le contrôle) de ses performances. Les rôles s'inversent ensuite : la deuxième sert au calage et la première à l'évaluation. Ce test est une simple procédure de calage-contrôle. Le SST est une procédure qui vise à estimer la possibilité de transférer un modèle et ses paramètres dans le temps sans changer de lieu. Cette qualité est généralement désignée comme la transférabilité temporelle du modèle et de ses paramètres.
- Le *Proxy-basin test* s'effectue en transférant un modèle et ses paramètres calés sur un bassin à un autre bassin de la même région, et vice-versa. La qualité ainsi évaluée est la transférabilité spatiale d'un modèle et de ses paramètres, puisque son principe est le même que le SST à l'exception du fait que le transfert se fait dans l'espace et non dans le temps.
- Le *Differential split-sample test* (DSST) est une version plus exigeante du SST, qui repose sur une sélection des périodes de calage-contrôle qui témoignent de conditions aussi contrastées que possible, par exemple une période très pluvieuse et une autre très sèche. Les conditions de simulation sont de la sorte aussi défavorables que possible en transfert dans le même bassin. Ce schéma est plus adapté à l'évaluation de la robustesse climatique que les deux précédents dans la mesure où il peut être adapté aux changements projetés par les modèles climatiques pour les futures décennies, en identifiant les variables climatiques les plus susceptibles d'évoluer. L'auteur recommande par exemple de procéder à un calage-contrôle d'une période sèche vers une période plus humide s'il est prévu que les conditions futures seront plus humides.
- Le *Proxy-basin differential split-sample test* est construit comme la combinaison du proxy-basin test et du DSST, ce qui en fait le test le plus exigeant des quatre. C'est un pré-requis essentiel dans l'hypothèse où un modèle doit servir à évaluer les conséquences du changement climatique dans des bassins non jaugés où l'application d'un DSST n'est pas possible.

Refsgaard *et al.* (2014) notent que le proxy-basin test est en général appliqué pour les modèles non calés, tels que les modèles de surface continentale inclus dans les modèles climatiques. Le DSST est naturellement le schéma de référence pour l'évaluation de la robustesse des modèles hydrologiques conceptuels, qui pour la grande majorité possèdent des paramètres libres. Plusieurs variantes en ont été proposées, qui seront évoquées au chapitre suivant. Refsgaard *et al.* (2014) constatent cependant que si le DSST est censé reproduire des conditions non-stationnaires de simulation proches de celles rencontrées en projection, les changements effectivement observés dans le passé sont vraisemblablement de moindre ampleur que ceux risquant d'advenir dans le futur. De ce fait, le DSST repose en pratique sur la sélection de périodes témoignant en « apparence » de différents climats.

1.2.2.c. Choix de modélisation en climat non-stationnaire

À ce stade de l'état de l'art, la notion de stationnarité peut sembler obscure après les discussions précédentes au sujet de la variabilité interne du climat et des différences climatiques recherchées par le DSST, et il convient d'en donner une définition.
En physique, la non-stationnarité s'oppose à la stationnarité. D'après le dictionnaire Larousse, un phénomène physique est dit stationnaire lorsqu'il se reproduit identiquement à lui-même au cours

du temps. Carbonnel et Hubert (1995) donnent une définition plus précise de la stationnarité sur le site du Comité National Français des Sciences Hydrologiques (CNFSH) : « une série chronologique de réalisations d'une grandeur aléatoire, à un pas de temps donné, est dite stationnaire si ses réalisations sont issues d'un même processus aléatoire dont les paramètres (moyenne, variance, asymétrie, autocorrélation...) restent constants au cours du temps ». Les réalisations d'un processus stationnaire peuvent donc exhiber des déviations ou des tendances persistant pendant plusieurs décennies ou centenaires (Cohn et Lins, 2005). À l'inverse, un processus non-stationnaire se reconnaît lorsque ses propriétés statistiques évoluent dans le temps de façon déterministe (Koutsoyiannis et Montanari, 2015).
Dans ce cadre, le réchauffement global apparaît comme un processus non-stationnaire, dans la mesure où le système climatique est perturbé par l'adjonction de gaz à effet de serre émis par la combustion de matières fossiles issues du sous-sol. L'addition de gaz à effet de serre anthropique n'est pas un processus stochastique et peut se décrire en termes déterministes. La variabilité interne du climat, en revanche, est par définition un processus stationnaire. De ce fait, un changement climatique passé ou à venir dans un bassin versant n'est pas nécessairement le résultat de la non-stationnarité du climat. L'enjeu de l'évaluation de la robustesse climatique d'un modèle hydrologique étant de vérifier qu'il est capable de simuler un même bassin versant dans des conditions climatiques variables, celles-ci peuvent être à l'origine de processus stationnaires (variabilité naturelle) ou non-stationnaires (changement climatique).

Dans cette perspective, le DSST présente un certain nombre de limites pour l'évaluation stricte de la robustesse climatique des modèles hydrologiques utilisés en projection. Stephens *et al.* (2020) en répertorient notamment deux :

- Les chroniques historiques sont susceptibles de ne pas contenir de périodes caractérisant des anomalies climatiques suffisamment longues pour observer une influence notable dans la réponse hydrologique du bassin versant. L'inertie des réserves d'eau souterraines peut par exemple masquer les effets qu'aurait un changement du régime des précipitations équivalent si celui-ci persistait sur le long terme.
- Il n'est pas possible d'observer les effets d'une augmentation de la concentration en gaz carbonique sur les comportements hydrologiques observables dans le passé. Une telle augmentation pourrait pourtant affecter les cycles de croissance de la végétation et leur utilisation des précipitations.

Pour ces raisons, la robustesse climatique des modèles hydrologiques telle qu'évaluée par des méthodes basées sur le DSST pourrait être sur-estimée. À cela s'ajoute le fait que ces études concernent en particulier les modèles hydrologiques conceptuels qui, au contraire des modèles de surface continentale, n'ont pas vocation à être couplés aux modèles de circulation atmosphérique. Ils ne modélisent donc pas certains mécanismes de rétroaction, comme l'intensification des précipitations extrêmes liée à la hausse de l'évaporation continentale (Trenberth, 1999).

On pourrait considérer que les modèles conceptuels ne sont pas les candidats idéaux pour une utilisation en projection hydrologique, tant pour des raisons de philosophie de modélisation que pour des raisons méthodologiques d'évaluation. Il faut cependant noter qu'à notre connaissance, les études d'impact qui impliquent des modèles à base physique non calés à l'échelle locale sont en pratique toujours fondées sur l'analyse de simulations en « offline » de ces modèles, c'est-à-dire forcés par des projections climatiques pré-calculées par des simulations couplées préalables. En cela, ces études d'impact ne diffèrent pas de celles impliquant des modèles conceptuels. Il convient également de rappeler que la robustesse des modèles de surface continentale n'est, à notre connaissance, jamais testée dans un cadre strict d'évaluation de la transférabilité temporelle de leurs paramètres et/ou de leurs équations. Ces examens ne sont en général menés que pour les méthodes de correction de biais des sorties des modèles climatiques régionaux (e.g. Teutschbein et

Seibert, 2012). L'évaluation des modèles de surface continentale est réalisée en général à partir de multiples variables à plusieurs échelles spatiales (e.g. Dutra *et al.*, 2010; Guimberteau *et al.*, 2018), dans des conditions où la validation n'est pas clairement indépendante du choix des équations et des paramètres du modèle. La méthodologie d'évaluation des modèles hydrologiques conceptuels les exposent donc naturellement davantage à la critique, au moins en ce qui concerne la simulation des débits.
Par ailleurs, les modèles hydrologiques conceptuels sont généralement utilisés en bout de chaîne de modélisation pour les projections climatiques, en tant que modèle d'impact pour la simulation spécifique du débit à l'exutoire de bassins versants. Les données météorologiques utilisées pour les forcer proviennent donc de simulations couplées de modèles climatiques, qui modélisent les effets des boucles de rétroactions entre atmosphère et surfaces continentales. Cantonnés à une utilisation en aval de ces modèles couplés, les modèles hydrologiques conceptuels peuvent se passer en première approche de l'intégration de ces boucles. L'intégration des effets de la hausse des concentrations en gaz carbonique reste toutefois un chantier inévitable pour le développement des modèles conceptuels en contexte de changement climatique, bien qu'il ne soit pas abordé dans ce travail de thèse.

1.3. Synthèse

Notions générales

L'objectif de ce chapitre était de présenter les principales notions manipulées dans la thèse. Ces notions concernent en premier lieu le champ de la modélisation hydrologique : le bassin versant et les enjeux de la mesure des flux qui le caractérisent, l'utilisation des modèles hydrologiques et les éléments de leur structure, les différentes philosophies de représentation des processus hydrologiques. Les enjeux liés à la simulation des effets du changement climatique en hydrologie ont été abordés en second lieu : les conséquences estimées du changement climatique sur le régime hydrologique, et les incertitudes entourant ces estimations. Le sujet de la thèse se situe à la confluence entre la question des incertitudes du changement climatique et la question de la modélisation hydrologique : les modèles hydrologiques sont-ils capables d'être utilisés en projection climatique ? Si non, comment les améliorer pour qu'ils le deviennent ?

Robustesse climatique des modèles hydrologiques

Pour répondre à cette question, nous avons introduit la notion de robustesse climatique des modèles hydrologiques, qui détermine leur capacité à simuler le régime hydrologique dans des conditions climatiques différentes de leurs conditions de développement et de calage. Dans la perspective de la simulation des effets du changement climatique sur le régime des cours d'eau, cette question est centrale. En effet, le changement climatique risque de perturber fortement les caractéristiques des forçages atmosphériques en entrée des modèles hydrologiques, voire la réponse des bassins versants à ces forçages.
L'impossibilité d'évaluer les modèles hydrologiques directement dans les conditions climatiques futures exige l'utilisation d'observations passées. Pour reproduire des conditions de changement climatique, la méthode d'évaluation la plus utilisée en modélisation hydrologique conceptuelle est la technique du *Differential Split-Sample Test* (DSST). La méthode consiste en une procédure de calage-contrôle d'un modèle hydrologique sur deux sous-ensembles des observations disponibles caractérisant des conditions hydro-climatiques contrastées.
Le chapitre suivant présente les principaux résultats concernant la robustesse climatique des mo-

dèles conceptuels. La grande majorité est obtenue par l'application du DSST ou de variantes. Ces conclusions montrent assez généralement que les modèles hydrologiques actuels ne sont pas assez robustes pour représenter des situations de changements climatiques passées.

Chapitre 2.

La robustesse des modèles hydrologiques : état de l'art et discussion

Introduction

Ce chapitre s'attache à dresser un état des lieux de la littérature sur la robustesse climatique des modèles hydrologiques conceptuels.
Pour rappel, la notion de robustesse climatique d'un modèle est intimement liée à celle de transférabilité temporelle de ses paramètres. Ceux-ci font en effet nécessairement l'objet d'un calage pour adapter le fonctionnement du modèle aux spécificités d'un bassin versant. Les données utilisées pour le calage n'étant pas nécessairement représentatives de l'ensemble des conditions climatiques subies par le bassin, tester la validité de ces paramètres dans d'autres conditions est déterminant pour l'utilisation du modèle en extrapolation.

2.1. Inventaire des faiblesses constatées des modèles

2.1.1. Constats généraux

2.1.1.a. Mise en exergue du manque de robustesse par le DSST

Pour évaluer un modèle en vue d'une utilisation opérationnelle, Klemeš (1986) proposait le Split Sample Test (SST), qui consiste à caler et évaluer ses paramètres sur deux périodes indépendantes. Il proposait également une variante plus exigeante, le Differential Split Sample Test (ou DSST), pour lequel les deux sous-périodes de calage-contrôle sont choisies de façon à ce qu'elles soient climatiquement contrastées. Malgré la précocité de ces travaux méthodologiques, le DSST est resté peu employé jusqu'à la dernière décennie, exception faite de quelques études pionnières (Refsgaard et Knudsen, 1996; Donnelly-Makowecki et Moore, 1999; Seibert, 2003). La question de l'évaluation et de l'amélioration du régime hydrologique s'est toutefois peu à peu imposée comme une question importante auprès de la communauté scientifique. Une série d'articles d'opinion (Milly *et al.*, 2008; Wagener *et al.*, 2010) a peut-être contribué à cette prise de conscience. Ce questionnement, en suscitant un besoin d'identifier les qualités et défauts des modèles, a débouché sur une utilisation de plus en plus fréquente du DSST, qui s'inscrit parfaitement dans la logique des *crash-tests* (Andréassian *et al.*, 2009).
Alors que le DSST était présenté par Klemeš (1986) comme une condition « nécessaire mais non suffisante » à l'utilisation de modèles en extrapolation, la plupart de ceux ayant été testés jusqu'à présent montrent des résultats peu satisfaisants.

Dans une étude menée sur quatre bassins suédois, Seibert (2003) a constaté que l'application du DSST au modèle HBV entre un calage sur une période de crues modérées et un contrôle sur une période de fortes crues pouvait donner des performances très médiocres. Son analyse démontre qu'en contrôle le modèle prédit correctement les crues modérées, de même amplitude que celles rencontrées en calage, mais sous-estime systématiquement les crues de grande ampleur. Le modèle calé directement sur la période de fortes crues ne sous-estime pourtant pas ces crues. L'auteur observe également que la dégradation des performances est plus forte dans les bassins où les différences entre la période de calage et celle de contrôle sont plus marquées.
Les conclusions de cette étude ont été généralisées par celles d'autres travaux menés avec le même modèle dans d'autres cadres et avec d'autres critères de distinction des périodes de calage et de contrôle. Hartmann et Bárdossy (2005) testent par exemple le même modèle sur un bassin versant allemand dans différentes configurations de calage-contrôle (période chaude vers période froide,

froide vers chaude, pluvieuse vers sèche, sèche vers pluvieuse). Leurs résultats montrent que le modèle témoigne de défaillances d'autant plus importantes entre calage et contrôle que les performances sont calculées à des échelles temporelles agrégées. Les critères calculés sur des pas de temps aggrégés à l'échelle saisonière montrent ainsi de nettes diminutions de performance selon la saison et la nature du transfert des paramètres. A l'échelle annuelle, les écarts de performance entre calage et contrôle sont forts quelle que soit la nature du transfert ou la fonction objectif de calage. Les travaux de Merz *et al.* (2011) poursuivent le diagnostic du modèle HBV sur un large échantillon de bassins autrichiens par la mise en place d'une procédure de calage-contrôle systématique fondée sur la division des chroniques en plusieurs sous-périodes de 5 ans. Cette méthode permet aux auteurs de constater l'augmentation du biais du modèle avec l'augmentation de l'écart temporel entre les périodes de calage et de contrôle. Les biais observés concernent autant les écoulement médians que les étiages et les crues.

Des conclusions similaires à celles tirées pour les études concernant le modèle HBV sont également valables pour d'autres modèles. Vaze *et al.* (2010) formulent un constat similaire en appliquant le DSST à quatre modèles de complexités variées sur un échantillon de bassins australiens en montrant que des changements dans les précipitations moyennes provoquent des simulations biaisées en contrôle. Brigode *et al.* (2013) retrouvent des motifs similaires pour deux autres modèles sur des bassins français. Broderick *et al.* (2016) ajoutent dans une étude sur un échantillon de bassins irlandais que la robustesse des modèles diffère selon les modèles utilisés, les scénarios de test, et les bassins. Ainsi, certains modèles testés produisent des simulations plus robustes dans certains types de bassins, bien qu'aucun ne soit robuste dans toutes les conditions de transfert évaluées. Fowler *et al.* (2016) notent que les pires performances obtenues sur une période de contrôle peuvent correspondre à de très bonnes performances sur la période de calage.
Enfin, un atelier organisé par l'Association Internationale des Sciences Hydrologiques (AISH) renforce le consensus sur le manque général de robustesse des modèles hydrologiques à travers une large expérience associant les modèles de plusieurs équipes de recherche sur un échantillon de bassins marqués par des changements climatiques ou d'usage des sols (Thirel *et al.*, 2015a,b).

2.1.1.b. Caractérisation du manque de robustesse des modèles hydrologiques

Le manque général de robustesse des modèles relevé dans les travaux cités précédemment invite à interpréter les résultats des projections hydrologiques avec précaution. Pour évaluer la confiance qu'il faut accorder aux modèles hydrologiques en extrapolation, il est toutefois nécessaire de caractériser plus en détail les problèmes de transférabilité en fonction des situations. Dans cette optique, certaines études ont été réalisées à partir de variantes du SST et du DSST afin de raffiner l'analyse des défauts des modèles. Dans leurs formes originales, le SST et le DSST ont l'avantage majeur d'être simples, faciles d'application et d'interprétation. Cependant, appliquées telles quelles, ces procédures présentent quelques défauts qui limitent la possibilité de dresser un diagnostic exhaustif des modèles testés. Un inventaire assez exhaustif des différents formats de ces expériences est établi par Dakhlaoui *et al.* (2019).

L'application du DSST nécessite la détermination d'une ou plusieurs variables climatiques pour distinguer des périodes hétérogènes. Il n'est cependant pas possible de connaître à l'avance les changements climatiques les moins propices au transfert des paramètres d'un modèle pour définir un crash-test aussi rude que possible. Pour étudier *a posteriori* l'influence de changements des conditions climatiques sur la robustesse des modèles hydrologiques, Coron *et al.* (2012) ont développé une procédure augmentant le nombre et la variété des expériences de calage-contrôle, nommée *Generalized Split-Sample Test* (GSST). La méthodologie repose sur la division des séries chronologiques d'un échantillon de bassins australiens en sous-périodes de 10 ans définies par le passage itératif d'une fenêtre temporelle glissant d'un an à chaque itération. Deux sous-périodes consécu-

tives contiennent donc 9 années en commun. En calant trois modèles hydrologiques sur chaque sous-période « glissante » et en évaluant les performances sur les sous-périodes indépendantes, les auteurs ont donc produit un grand nombre d'expériences de calage-contrôle et comparé les baisses de performance aux écarts climatiques entre les périodes donneuses et receveuses. Leurs résultats montrent dans de nombreux cas une baisse significative des performance des modèles lors des changements dans la pluviométrie moyenne, tandis qu'aucun lien significatif n'est observable avec les variations de la température moyenne ou de l'aridité des bassins. En général, les calages sur des périodes plus humides provoquent une surestimation des débits en contrôle sur des périodes sèches, et inversement les calages sur des périodes plus sèches provoquent une sous-estimation des débits en contrôle sur des périodes humides.

Un certain nombre d'études impliquant un DSST construit à partir des précipitations moyennes prolonge le constat en observant que la transférabilité des modèles et de leurs paramètres est généralement meilleure en calage sur une période sèche et en contrôle sur une période pluvieuse que lorsque les rôles s'inversent (Vaze *et al.*, 2010; Li *et al.*, 2012; Bisselink *et al.*, 2016). Dakhlaoui *et al.* (2017) rajoutent que les transferts de périodes sèches et chaudes vers des périodes humides et froides sont plus favorables que l'opposé, en implémentant un DSST *bootstrap* considérant à la fois les températures et les précipitations annuelles moyennes. Le terme bootstrap indique que les sous-périodes du DSST ne sont pas nécessairement constituées d'années consécutives, ce qui autorise ainsi à maximiser l'hétérogénéité des sous-périodes entre elles et à limiter d'éventuelles erreurs évolutives des données d'entrée (Coron, 2013).
Li *et al.* (2012) avancent plusieurs hypothèses pour expliquer l'asymétrie de transférabilité entre les différents types de conditions climatiques. Les auteurs relèvent notamment que les périodes sèches sont dans la plupart des cas marquées par des débits faibles et un rendement des précipitations (part des précipitations effectivement transformées en écoulement à l'exutoire) faible, ce qui rend l'erreur du modèle très sensible aux variations des paramètres et aux erreurs dans les données de précipitations. La prise en compte de la variabilité spatiale des précipitations et de l'humidité du sol est d'ailleurs considérée comme un facteur potentiel d'amélioration des simulations en conditions sèches. Les hétérogénéités spatiales d'humidité du sol peuvent en particulier être plus prononcées en conditions sèches qu'en conditions humides. Les processus liés à la génération de l'écoulement sur les sols étant caractérisés par des fortes non-linéarités et la présence de paliers, les erreurs en modélisation en conditions sèches en sont d'autant plus importantes.

2.1.1.c. Irréductibilité des erreurs de volume

Les modèles évalués par le DSST et quelques variantes montrent des défaillances en contrôle prononcées dans l'estimation des volumes moyens écoulés. Il est communément reconnu dans ces études que les biais constatés limitent considérablement la confiance qu'il est possible d'accorder aux simulations en extrapolation réalisées avec les modèles conceptuels, en particulier pour la modélisation de la ressource en eau des futures décennies. Pourtant, les auteurs de ces diagnostics relèvent que les modèles sont capables de simuler correctement le bilan en eau pour chacune de ces périodes.
L'analyse de leurs paramètres peut expliquer les baisses importantes de performance observées entre calage et contrôle des modèles hydrologiques. Kelleher et Shaw (2018) observent par exemple que quelques paramètres du modèle HBV contrôlent son bilan hydrique et montrent que recaler un seul de ces paramètres en contrôle permet de réduire fortement les erreurs sur les volumes moyens simulés. De nombreuses études portant sur l'évaluation de modèles hydrologiques en conditions climatiques changeantes ont relevé une grande incertitude liée à l'estimation des paramètres des modèles en fonction de la période utilisée en calage (e.g. Wilby, 2005). La dépendance de certains paramètres à la période utilisée en calage peut être considérée comme un signe préoccupant de défaillances structurelles d'un modèle, dans la mesure où ses paramètres sont censés être repré-

sentatifs de propriétés physiques inhérentes à un bassin versant, et donc indépendants des forçages météorologiques (Bard, 1974, cité par Wagener *et al.*, 2003). Ces considérations soulèvent la question de la possibilité d'identifier un jeu de paramètres capable de résoudre simultanément le bilan en eau d'un bassin versant dans différentes conditions.

Pour répondre à cette question, Coron *et al.* (2014) proposent une procédure adaptée du GSST (Coron *et al.*, 2012) fondée sur i) le calage successif de trois modèles hydrologiques de complexités variées sur des sous-périodes glissantes de chroniques de 20 bassins français de montagne, puis ii) le calage sur l'ensemble de la chronique, et iii) le calcul du biais des modèles munis des jeux de paramètres obtenus sur chacune des sous-périodes. Les résultats montrent que les courbes des biais des modèles munis des différents jeux de paramètres sont toutes parallèles entre elles, y compris la courbe obtenue en calage sur l'ensemble des données disponibles (Figure 2.1). Les auteurs constatent également que chacune de ces courbes est positionnée de sorte à minimiser le biais des modèles sur la sous-période de calage. Le parallélisme des courbes des biais sur les sous-périodes glissantes concerne les trois modèles testés, de l'équation à 1 paramètre libre au modèle semi-distribué comportant 19 paramètres libres. Ce résultat illustre l'insensibilité des problèmes de bilan des modèles à la longueur des périodes de calage. Moins les courbes de biais sous les sous-périodes glissantes sont « plates » (i.e. horizontales), moins il semble possible d'identifier un unique jeu de paramètre permettant de simuler correctement le bilan hydrologique d'un bassin versant à toutes les périodes. Au contraire, l'existence d'un jeu de paramètres produisant une courbe plate serait un signe positif de la robustesse d'un modèle dans la mesure où l'erreur relative resterait constante et donc amendable par une correction *a posteriori*.

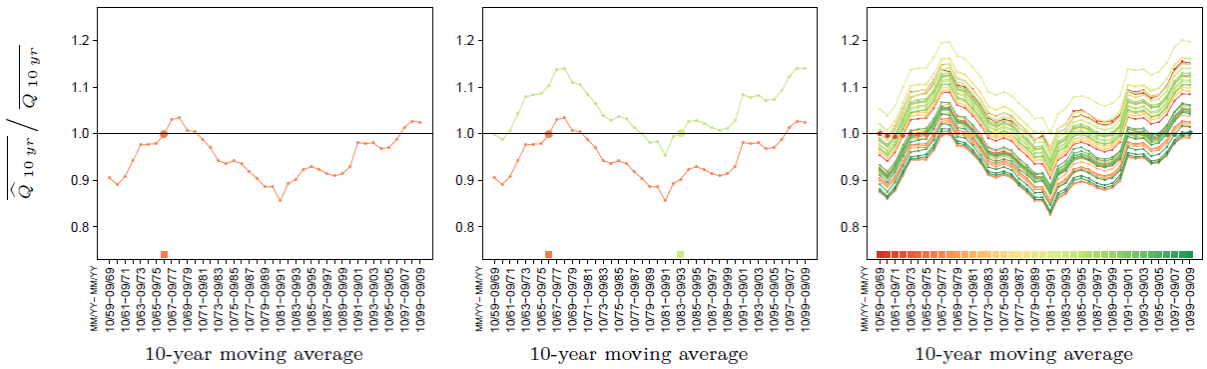


FIGURE 2.1 : Construction de la courbe des biais sur les sous-périodes glissantes. Tiré de Coron *et al.* (2014). Le modèle est d'abord calé sur une sous-période des données longue de 10 ans et son biais est calculé sur toutes les autres sous-périodes de 10 ans (à gauche). Le modèle est ensuite calé sur une autre sous-période (au milieu). Le modèle est calé itérativement sur toutes les sous-périodes disponibles (à droite).

Fowler *et al.* (2016) prolongent le constat en montrant que l'optimisation simultanée des performances de modèles hydrologiques variés sur deux sous-périodes très contrastées aboutit à un aplatissement très modéré des courbes de biais. Les 5 modèles utilisés et le cadre de l'étude, 86 bassins australiens ayant connu un assèchement considérable, suggèrent que les conclusions concernant le parallélisme des courbes de biais sur sous-périodes glissantes sont généralisables à d'autres contextes de modélisation.

2.1.2. Discussions suscitées par le constat d'un manque généralisé de robustesse

2.1.2.a. Liens observés entre conditions climatiques et paramètres des modèles

La recherche de liens entre les variations des paramètres et celles des conditions climatiques peut permettre de détecter les faiblesses structurelles des modèles hydrologiques. Merz *et al.* (2011) rapportent l'existence de tendances communes entre certains paramètres du modèle HBV, notamment ceux liés à la représentation de l'humidité du sol et de la neige, et la température d'un échantillon de bassins autrichiens. Avec le même modèle, Osuch *et al.* (2015) et Kelleher et Shaw (2018) constatent également des dépendances non linéaires entre les paramètres de sol et des variables climatiques, notamment la température. Dakhlaoui *et al.* (2017) étendent ces résultats aux paramètres de deux autres modèles sur quelques bassins tunisiens. Tout en reconnaissant que cette pratique n'est théoriquement pas satisfaisante, Merz *et al.* (2011) suggèrent de corriger *a posteriori* les paramètres sensibles du modèle HBV en fonction des variations des conditions climatiques.
Cette idée est mise en pratique par Coron (2013), qui expérimente le transfert de relations linéaires entre la température et certains paramètres des modèles GR4J (Perrin *et al.*, 2003) et Mordor (Garçon, 1996; Mathevet, 2005) entre deux échantillons de bassins. L'auteur constate des effets différenciés selon les modèles : la robustesse de Mordor bénéficie nettement de cette correction, tandis que les résultats sont plus mitigés pour GR4J. L'explication avancée souligne le plus grand nombre des processus de GR4J impliqués dans le bilan du modèle et leur imbrication, au contraire de celui de Mordor qui est principalement réglé par un unique facteur correctif des précipitations. Malgré ces résultats encourageants, Coron (2013) évoque en outre plusieurs limites générales à la méthode de correction climatique des paramètres des modèles hydrologiques :

- Nécessité de démontrer l'existence de corrélations à la fois entre des variations climatiques, les erreurs du modèle et certains de ses paramètres. Souvent, des relations entre paramètres et climat ne sont pas clairement identifiables ou sont différentes d'un bassin à l'autre (e.g. Li *et al.*, 2012; Motavita *et al.*, 2019; Fowler *et al.*, 2020). La méthode mise en place par Coron (2013) est d'ailleurs appliquée sur une sous-partie de l'échantillon d'étude, constituée uniquement des bassins où des liens significatifs entre température et biais des modèles ont été observés.
- Les variations simultanées de plusieurs variables climatiques peuvent nuire à la transférabilité des facteurs correctifs des paramètres, ou exiger des relations multi-paramètres et multi-variables climatiques impliquant de nombreux paramètres *ad hoc* à caler. La robustesse de ces relations correctives est donc difficile à garantir.

L'idée de permettre aux paramètres de modèles hydrologiques de varier dans le temps pour adapter les processus modélisés est donc difficile à implémenter en pratique, bien qu'il ne soit pas exclu qu'elle puisse être utile dans certaines situations (e.g. Westra *et al.*, 2014). Certaines études cherchent à développer cette technique dans les cas où des modifications d'usage des sols sont connues (e.g. Pathiraja *et al.*, 2018). Cette problématique ne rentre toutefois pas dans le cadre de celle de la robustesse climatique.

2.1.2.b. Limites d'un calage « mimétique »

Le constat de la faible transférabilité des modèles et de leurs paramètres entre différentes conditions climatiques est fréquemment accompagné d'invitations à choisir une période de calage proche des conditions attendues dans le futur pour diminuer les erreurs potentielles des modèles hydrologiques (e.g. Li *et al.*, 2012; Bisselink *et al.*, 2016; Broderick *et al.*, 2016; Dakhlaoui *et al.*, 2017). L'idée serait donc de caler un modèle hydrologique sur une sous-période des observation « mimant » le climat futur. Faute de mieux, cette précaution semble logique. Il convient néanmoins d'évoquer succinctement les raisons pour lesquelles elle risque de n'être pas suffisante pour garantir des projections hydrologiques de bonne qualité.

L'existence d'incertitudes liées à chaque maillon de la chaîne évoquées au Chapitre 1 suggère qu'il n'est parfois pas possible d'obtenir une information claire et robuste sur les tendances climatiques attendues dans une région donnée. Le signal modélisé pour les précipitations moyennes est parfois peu cohérent entre les différents couples GCM–SDM ou GCM–RCM (e.g. Wilby et Harris, 2006). La variabilité interne du climat peut être une des principales causes d'incohérence des projections (Deser *et al.*, 2012). Dans ces situations, il n'est pas possible de choisir une unique période reflétant les conditions futures.
Il a par ailleurs déjà été mentionné au Chapitre 1 qu'il n'était pas garanti que les conditions futures puissent être contenues dans les données historiques disponibles. Dakhlaoui *et al.* (2019) démontrent ainsi non seulement que les changements climatiques projetés dans un ensemble de bassins tunisiens vont au-delà de la variabilité historique observé, mais aussi que ces changements dépassent les variations climatiques « acceptables » pour la transférabilité de trois modèles hydrologiques, même comptés en terme relatif par rapport à la sous-période historique la plus proche des conditions futures.

2.1.2.c. Diagnostic des modèles hydrologiques

Parmi toutes les causes possibles des difficultés constatées des modèles hydrologiques conceptuels à simuler les débits hors période d'apprentissage, certaines ressortent de la littérature existante avec plus d'emphase. La structure des modèles conceptuels est une des explications évoquées le plus fréquemment : les modèles semblent trop simples, trop dépendants des données de calage, et donc manquant de degrés de liberté pour représenter des comportements hydrologiques non-stationnaires. La diversité des modèles testés et jugés insuffisamment robustes renforce la plausibilité de cette interprétation. Certains travaux ont exploré des pistes d'amélioration de la structure de certains de ces modèles dans le but précis d'en améliorer la robustesse. Le nombre de ces tentatives est toutefois limité à ce jour dans la mesure où la phase de diagnostic ne semble pas achevée, comme en attestent des trouvailles récentes (e.g. Bouaziz *et al.*, 2020; Duethmann *et al.*, 2020; Fowler *et al.*, 2020).
Le besoin de diagnostiquer avec finesse la structure et les équations des modèles hydrologiques se heurte d'ailleurs au problème du calage des paramètres. S'il est reconnu qu'une structure inadéquate ne peut aboutir à des résultats robustes, la phase de calage joue également un rôle majeur tant le choix des paramètres est soumis à un grand nombre d'aléas, nuisant au potentiel prédictif d'un modèle (Beven, 2006). Le choix de la (ou des) fonction(s) objectif, de l'algorithme, des données d'apprentissage sont autant de causes d'incertitude sur la transférabilité des paramètres. Des techniques de calage inadaptées sont de ce fait peut-être à l'origine d'une sous-estimation de la robustesse des modèles. Cette question est d'autant plus importante que l'intégration de mécanismes manquants à sa structure peut en accroître la complexité et rendre cet exercice de calage plus délicat encore.

Les développements suivants présentent les principales conclusions de la littérature concernant ces deux pistes d'explication majeures du manque observé de robustesse des modèles pluie-débit, c'est-à-dire la structure et le calage des modèles. Dans la mesure où ce constat est étroitement lié à la procédure de calage-contrôle du DSST et de ses variantes, et pour les raisons présentées au paragraphe précédent, la piste du calage des modèles est exposée avant celle concernant leur structure. Le rôle des erreurs dans les données, également invoqué parmi les causes de l'inadaptation des modèles hydrologiques, est abordé de façon transversale.

2.2. Causes du manque de robustesse : rôle du calage des paramètres

2.2.1. Principes usuels du calage des modèles hydrologiques

2.2.1.a. Calages manuel et automatique

Il existe deux grandes façons de procéder au calage des paramètres d'un modèle hydrologique, évoquées brièvement au Chapitre 1. Le calage manuel repose sur l'expertise préalable d'un hydrologue, c'est-à-dire sur sa connaissance du fonctionnement du modèle et de celle du bassin versant d'étude, pour identifier des paramètres permettant aux processus modélisés de correspondre aux processus réels. Généralement, cette identification est faite par étapes successives de raffinement de la valeur des paramètres, « en tournant les boutons », aux moyens d'analyses graphiques et de différents critères d'erreur pour juger de la qualité des simulations. Cette technique s'oppose aux méthodes de calage automatique où l'estimation des paramètres est déléguée à un algorithme chargé de minimiser l'erreur du modèle matérialisée par une ou plusieurs fonctions objectif.

Bien qu'aujourd'hui la plupart des modèles hydrologiques soient calés de manière automatique, la méthode manuelle présente quelques avantages qu'il convient de rappeler, dans la mesure où ils peuvent concourir à éviter quelques écueils rencontrés par les méthodes automatiques dans l'estimation de paramètres robustes. Au contraire des machines binaires capables de ne traiter qu'une information à la fois, même à grande fréquence, le cerveau humain peut considérer un problème de grande dimension de manière synoptique et arbitrer des compromis entre différents aspects avec subtilité et flexibilité (Gupta *et al.*, 1998). Cette aptitude permet par exemple d'intégrer la présence d'erreurs potentielles dans les données. Par ailleurs, la connaissance des mécanismes du modèle prévient l'identification de valeurs « non physiques » des paramètres. Néanmoins, Boyle *et al.* (2000) soulignent que le calage manuel exige beaucoup de temps et de travail préalable. Dans la mesure où il est recommandé de développer et évaluer les modèles hydrologiques sur de larges échantillons de bassins versants pour identifier des tendances robustes (Andréassian *et al.*, 2006; Gupta *et al.*, 2014; Addor *et al.*, 2020), cette exigence réduit considérablement son potentiel d'utilisation. Par ailleurs, la subjectivité intervenant nécessairement dans le choix des paramètres limite la reproductibilité du processus de calage par d'autres hydrologues (Boyle *et al.*, 2000) et donc à sa vérifiabilité et sa crédibilité (Hutton *et al.*, 2016).

En comparaison du calage manuel, le calage automatiquel a l'avantage d'être plus rapide, plus objectif, et facile à appliquer à des modèles hydrologiques conceptuels. Appuyé par l'augmentation des capacités de calcul, le calage automatique fait aujourd'hui office de norme. La quasi-totalité des études mentionnées précédemment et participant au consensus de manque de robustesse des modèles hydrologiques repose sur ce procédé, exceptés les travaux de Refsgaard et Knudsen (1996) et de Donnelly-Makowecki et Moore (1999). En général, une procédure de calage automatique consiste en l'optimisation d'une unique fonction objectif à partir d'une exploration de l'espace des paramètres conduite par un algorithme, bien qu'il existe également des méthodes d'optimisation simultanée de plusieurs fonctions objectif, qui seront abordées dans la suite.

Pour une fonction objectif donnée, l'enjeu des procédures de calage automatique se situe dans la capacité de l'algorithme de calage à explorer correctement l'espace des paramètres dans les limites fixées par le scientifique. Mathevet (2005) décrit deux principales attentes à l'égard des algorithmes de recherche de l'optimum :

- efficacité, c'est-à-dire atteinte du jeu du paramètres optimisant strictement la fonction objectif ;
- efficience, quantifiée par le nombre de tentatives effectuées avant l'identification du jeu de paramètres final.

Ces deux attentes sont relativement contradictoires, dans la mesure où tester un unique jeu de paramètres est très efficient mais très inefficace, et en tester une infinité est très efficace mais très inefficient. En pratique, la capacité d'un algorithme à satisfaire cette double exigence peut-être contrariée par des difficultés liées à la forme de la surface de réponse du modèle, qui dépend à la fois des données de calage, de la fonction objectif et de la structure du modèle. Une surface de réponse plus irrégulière, c'est-à-dire revêtant l'apparence d'une succession de pics et de vallées abruptes, augmente le risque que l'algorithme ne converge pas vers le jeu de paramètres optimal, mais plutôt vers un optimum secondaire (Duan *et al.*, 1992). De même, l'insensibilité de la performance du modèle à certains paramètres (que l'on pourrait aussi appeler « indétermination » de ces paramètres) ou au contraire son hypersensibilité à d'autres sont des facteurs supplémentaires de sous-optimalité. Un jeu de paramètres sous-optimal peut donner l'illusion fausse que la structure d'un modèle est inadaptée (Andréassian *et al.*, 2012), bien qu'on puisse considérer qu'un modèle dont le calage convergerait systématiquement vers des optima secondaires est effectivement inadéquat à la modélisation hydrologique (modèle sur-paramétré).

Face à ces enjeux, de nombreux algorithmes ont été développés dans la recherche d'un compromis entre efficacité et efficience (e.g. Wang, 1991; Duan *et al.*, 1992; Hansen *et al.*, 2003; Reed et Devireddy, 2004). Les problématiques liées au sur-paramétrage ont également été intégrées au développement de certains modèles cherchant explicitement une forme de « frugalité » (e.g. Perrin *et al.*, 2003; Mathevet, 2005). Malgré ces efforts, il convient généralement d'interpréter les résultats d'un calage automatique avec vigilance, tant les facteurs de risques de convergence vers un optimum secondaire sont nombreux et dépendants des conditions de simulation.

2.2.1.b. Fonctions objectif usuelles

En considérant que les modèles sont des représentations imparfaites de la réalité vouées à ne jamais reproduire un hydrogramme observé à l'identique, un jeu de paramètres ne saurait atteindre une erreur du modèle strictement nulle. Le choix de la fonction objectif optimisée affecte donc nécessairement la nature du jeu de paramètre optimal, si tant est qu'il soit identifiable par un algorithme de calage. Bien que le choix de la fonction objectif dépende des besoins du problème à traiter et qu'il existe des possibilités infinies, la plupart des recherches évoquées précédemment sont fondées sur un ensemble limité de métriques. La raison de la popularité de ces quelques métriques est probablement à la fois liée à des questions de formalisme mathématique, d'habitudes, de facilité d'interprétation et d'efficacité.
En statistiques, il est courant d'évaluer le pouvoir explicatif d'un modèle en considérant la dispersion des erreurs à travers leur variance. La régression linéaire d'une variable observée par rapport à une autre s'obtient par exemple en minimisant cette variance, c'est-à-dire la moyenne des carrés des écarts entre les débits simulés et observés (en anglais *Mean Square Error*, ou MSE). Moyennant certaines hypothèses sur les erreurs des variables observées en question, cette méthode d'optimisation dite « des moindres carrés » permet d'obtenir les paramètres les plus probables. En hydrologie, les modèles ont été traditionnellement calés en minimisant des fonctions objectif adaptées de la racine carré de la MSE (ou RMSE), qui permet d'obtenir une estimation de l'erreur de la même dimension que les débits observés. Le critère de Nash et Sutcliffe (1970), généralement appelé *Nash-Sutcliffe Efficiency* (NSE) ou sobrement critère de Nash, est peut-être l'une des fonctions objectif les plus populaires. Originellement conçu pour être équivalent au coefficient de détermination (ou « variance expliquée », R^2) utilisé en statistiques, ce critère présente l'avantage de normaliser l'erreur par la comparaison avec un modèle naïf, valant la moyenne du débit observé, et donc dans une certaine mesure de rendre comparables les performances d'un modèle hydrologique sur plu-

sieurs bassins versants.

La pertinence du formalisme mathématique au fondement de ces fonctions objectif est néanmoins contrariée par l'invalidité des hypothèses de base sur les erreurs du modèle, à savoir qu'elles sont en moyenne nulles, indépendantes (non corrélées) et de variance constante (homoscédasticité). Par exemple, les erreurs du modèle sont généralement corrélées dans la mesure où les débits d'un cours d'eau et donc les erreurs de mesure sont fortement autocorrélés (Sorooshian et Dracup, 1980), remettant ainsi en question l'hypothèse sur l'indépendance des erreurs. La structure à « mémoire » des modèles hydrologiques participe par ailleurs à propager les erreurs dans les données d'entrée sur plusieurs pas de temps consécutifs (Kavetski *et al.*, 2003). La crédibilité de l'hypothèse d'indépendance des erreurs est donc questionnable. En outre, les erreurs ne respectent généralement pas l'hypothèse d'homoscédasticité (Sorooshian et Dracup, 1980). Les erreurs de mesure sur les débits de crue sont typiquement plus grandes qu'en étiage, notamment du fait de la concavité habituelle des courbes de tarage. D'autres arguments moins théoriques ont été invoqués dans les nombreuses discussions du critère de Nash (listées par Gupta *et al.*, 2009). La comparaison d'un modèle hydrologique avec un modèle aussi naïf que le modèle « moyenne » conduit par exemple à surestimer ses performances sur les bassins très saisonniers, où ce modèle de référence est particulièrement inadapté (Schaefli et Gupta, 2007). Ces critiques ont donné lieu à des propositions de modifications du NSE (e.g. Mathevet *et al.*, 2006; Schaefli et Gupta, 2007) ou à des critères alternatifs.

La fonction objectif la plus couramment utilisée est certainement le KGE (i.e. *Kling-Gupta Efficiency*), proposé par Gupta *et al.* (2009), qui décompose l'erreur du modèle sous la forme d'un critère agrégeant la corrélation, le biais et le rapport des écarts-types des débits simulés et observés. Cette décomposition permet l'optimisation de ces sous-composantes en limitant leurs interactions comme dans le cas de la RMSE et du NSE. Ces interactions peuvent notamment produire des biais volumiques importants sur les bassins au régime très variable. L'intérêt du KGE se situe également dans la grande facilité d'interprétation physique de ses sous-composantes.

2.2.1.c. Enjeux

Les principes du calage des paramètres des modèles hydrologiques exposés dans les paragraphes précédents abordent les principaux enjeux liés à l'estimation de jeux de paramètres satisfaisants. La difficulté majeure en calage tient dans les fortes interactions entre la structure du modèle, les données de calage et la formulation de l'erreur du modèle. Ces interactions influencent directement la complexité du problème d'optimisation à résoudre pour les algorithmes de calage, éprouvant la capacité des algorithmes de calage à atteindre l'optimum de la surface de réponse. Néanmoins, les difficultés liées à l'application des théories statistiques de base à la réalité du contexte de la modélisation hydrologique pour l'élaboration de fonctions objectif appropriées conduisent à douter de la pertinence de cet optimum, même dans l'hypothèse où l'on disposerait de l'algorithme de recherche idéal.
Andréassian *et al.* (2012) invitent à considérer l'existence de deux optima distincts en calage des modèles hydrologiques : l'optimum mathématique est le maximum atteignable de la fonction objectif par un jeu de paramètres existant dans l'espace des paramètres accessibles, l'optimum « hydrologique » est le jeu de paramètres permettant d'obtenir les simulations les plus robustes. On peut ajouter que cet optimum ne dépend pas des données d'apprentissages, pourvu qu'elles soient en quantité suffisante. Il caractériserait au contraire les propriétés inhérentes du bassin versant, son identification permettant donc de se passer de paramètres variant avec le temps. Les paragraphes suivants font état des raisons de la distinction entre l'optimum mathématique et l'optimum hydrologique et des solutions proposées dans la littérature pour y remédier.

2.2.2. Recherche d'un optimum hydrologique

2.2.2.a. Alternatives classiques aux moindres carrés

Les défauts évoqués des méthodes fondées sur l'approche des moindres carrés sont connus pour orienter les paramètres des modèles hydrologiques vers la simulation des hautes eaux. Krause *et al.* (2005) relèvent ainsi que cette tendance se fait au détriment des capacités des modèles à reproduire certains processus, tels que la génération du débit de base. En considérant que certains contextes de modélisation requièrent des modèles fiables dans la représensation des crues, ces méthodes peuvent apparaître comme adaptées. Néanmoins, il est nécessaire de corriger cette tendance lorsqu'il est attendu le modèle soit performant dans de nombreuses conditions.
Certains travaux s'appuient sur la transformation préalable des débits observés et simulés pour rétablir la validité de l'hypothèse d'homoscédasticité (Box et Cox, 1964). L'utilisation du logarithme des débits a par exemple été expérimentée pour donner plus d'emphase aux débits d'étiage (e.g. Dawdy et Lichty, 1968; Oudin *et al.*, 2006a; De Vos *et al.*, 2010), ou encore celle d'exposant comme la racine carrée dans l'objectif d'atteindre une répartition plus équilibrée (e.g. Chiew *et al.*, 1993; Oudin *et al.*, 2006a). Une alternative pour éviter l'emploi d'une transformation arbitraire inadaptée à la distribution des erreurs peut consister à accorder un poids à l'erreur carrée sur chaque pas de temps (e.g. Sorooshian et Dracup, 1980; Thyer *et al.*, 2009). L'utilisation de critères de performance construits à partir des écarts absolus plutôt que quadratiques peut également permettre de diminuer l'influence des crues dans le processus de calage (e.g. Criss et Winston, 2008).
D'autres auteurs contournent quant à eux les limites de l'hypothèse d'indépendance des erreurs du modèle dans les approches fondées sur les moindres carrés en testant la capacité des modèles hydrologiques à reproduire l'autocorrélation des débits observés. Ces études impliquent généralement des méthodes de traitement du signal et cherchent à valoriser l'information sur les dynamiques du régime hydrologique contenue dans les données de calage (e.g. Montanari et Toth, 2007; Schaefli et Zehe, 2009).

2.2.2.b. Calage multi-objectif

On peut considérer qu'il existe un dénominateur commun aux différentes techniques de sélection des paramètres d'un modèle : la nécessité d'exploiter au mieux l'information disponible pour contraindre les paramètres à dimensionner les processus modélisés. Certains processus dominent la génération de l'écoulement dans des conditions spécifiques, par exemple en étiage prolongé ou en crues soudaines, et nécessitent parfois des ajustements des fonctions objectif traditionnelles pour peser sur le calage des paramètres. Plutôt que d'agréger l'erreur sur chaque pas de temps en une seule valeur synthétique au prix d'une perte d'information (Wagener *et al.*, 2003), certaines méthodes consistent en l'optimisation simultanée de plusieurs critères d'erreur. Gupta *et al.* (1998) suggèrent que ces méthodes de calage, qualifiées de « multi-objectif », devraient impliquer des aspects complémentaires et indépendants du régime hydrologique.

A cette fin, l'emploi de « signatures hydrologiques » peut aider à diagnostiquer la capacité du modèle à reproduire les caractéristiques des processus dominant le régime hydrologique d'un bassin versant (Gupta *et al.*, 2008). McMillan *et al.* (2017) définissent les signatures hydrologiques comme des « indices calculés à partir de séries de données observées et modélisées telles que les précipitations, le débit ou l'humidité du sol ». Shamir *et al.* (2005b) remarquent par exemple la stabilité des temps caractéristiques d'ascension et de récession hydrologique dans le temps sur un échantillon de bassins versants et évaluent leur utilisation en calage d'un modèle hydrologique. Le calage multi-objectif de modèles hydrologiques peut impliquer des critères variés, par exemple des signatures caractérisant les caractéristiques du débit à différentes échelles temporelles (e.g. Shamir *et al.*,

2005a; Shafii et Tolson, 2015), relatives à la dispersion des débits à partir de la courbe des débits classés (e.g. Winsemius *et al.*, 2009; Vrugt et Sadegh, 2013) et mélanger signatures et critères aux moindres carrés (e.g. Shamir *et al.*, 2005b; Vrugt et Sadegh, 2013; Shafii et Tolson, 2015). La valorisation d'autres sources de données que le débit et les précipitations peut également participer à contraindre les paramètres dimensionnant des processus internes aux modèles. Si ces données sont disponibles, le calage peut inclure des critères liés à des données de qualité des eaux (e.g. Kuczera et Mroczkowski, 1998), de niveaux piezométriques (e.g. Seibert *et al.*, 1997; Madsen, 2003), d'humidité du sol (e.g. Milzow *et al.*, 2011), de gravimétrie (e.g. Milzow *et al.*, 2011; Bai *et al.*, 2018; Mostafaie *et al.*, 2018), d'évaporation réelle (e.g. Fowler *et al.*, 2018a), ou de couverture neigeuse (e.g. Riboust *et al.*, 2019). La disponibilité ou la qualité de ces sources de données ne suffisent pas toujours pour calculer une erreur entre une variable observée et une variable simulée. Certains travaux rapportent leur incorporation comme « données douces » (*soft data* en anglais) pour forcer les modèles à adopter certains comportements, comme reproduire certaines signatures hydrologiques dans des bassins non-jaugés (Winsemius *et al.*, 2009) ou simuler des tendances interannuelles dans les stocks d'eau souterraine modélisés (Fowler *et al.*, 2020).

L'inclusion d'objectifs multiples en calage peut se faire suivant plusieurs degrés de complexité. L'approche la plus simple consiste à intégrer ces objectifs dans un unique critère scalaire, par exemple sous la forme d'une distance dans un espace à plusieurs dimensions. C'est par exemple l'approche au fondement du KGE, qui est construit comme une distance dans un espace tri-dimensionnel. D'autres fonctions objectif sont constuites de façon plus triviale par l'addition de critères pondérés calculés sur les débits (e.g. Zhang *et al.*, 2008; Duethmann *et al.*, 2020). L'intégration de critères dans le cadre de « limites d'acceptabilité » (c.f. paragraphes sur les incertitudes) est également envisagée pour agréger les performances en comptabilisant le nombre d'objectifs remplis par chaque jeu de paramètres (e.g. Blazkova et Beven, 2009; Shafii et Tolson, 2015). Parce qu'elles réduisent le calage multi-objectif à l'optimisation d'un scalaire, ces fonctions objectif sont simples d'utilisation. Toutefois, elles négligent les phénomènes de compensation et de compétition entre les sous-critères, qui, combinés à des différences de sensibilité, peuvent compromettre l'atteinte de scores satisfaisants pour chacun d'entre eux (Gupta *et al.*, 2009).
Au contraire, envisager le problème d'optimisation dans l'espace des sous-critères permet d'explorer ces interactions. L'estimation du front de Pareto, constitué des jeux de paramètres non dominés, tels qu'il n'est pas possible d'améliorer le score dans un sous-critère sans diminuer celui d'un autre (Pareto, 1909), permet de sélectionner les jeux de paramètres qui établissent le meilleur compromis entre les différents objectifs de performance. Cette approche a été utilisée en calage de modèles hydrologiques pour caractériser les performances dans différentes gammes d'écoulement et à des échelles temporelles variées (e.g. Madsen, 2000; Wöhling *et al.*, 2013). Cherchant un jeu de paramètres climatiquement robuste, Gharari *et al.* (2013) considèrent par exemple la maximisation simultanée des performances d'un modèle hydrologique sur des sous-périodes de la période de calage, et obtiennent des résultats encourageants.
Le calage dans un espace de plusieurs dimensions requiert l'emploi d'algorithmes spécifiques de calcul de Pareto (e.g. Vrugt et Robinson, 2007; Monteil *et al.*, 2020) ou de méthodes de Monte Carlo.

2.2.2.c. Intégration des incertitudes

La recherche du jeu de paramètres optimal au sens hydrologique peut prendre des voies variées, à partir d'approches sur la minimisation des erreurs à chaque pas de temps, dans le domaine fréquentiel, à partir d'objectifs multiples visant différents moments statistiques du débit observé ou encore par la reproduction de signatures caractéristiques du bassin versant. La robustesse climatique d'un tel jeu de paramètres suppose qu'il soit déterminé par la physique intrinsèque du bassin versant et non par les conditions de son estimation. Cette ambition de représenter la réalité d'un bassin par un jeu unique de paramètres de modèle se heurte néanmoins à plusieurs limites.

La première de ces difficultés est liée à la présence inéluctable d'erreurs dans les données d'apprentissage. Dans la plupart des approches décrites précédemment, le débit simulé en sortie de modèle est confronté directement aux débits observés, qui sont sujets à des imprécisions et des erreurs de mesure abordées brièvement au Chapitre 1. Ces erreurs n'ont pas la même influence selon qu'elles concernent certaines catégories de débit ou d'autres. Berthet *et al.* (2010) montrent que les jeux de paramètres dépendent en réalité fortement des débits d'un nombre réduit d'événement de crues. Bien que cette étude soit réalisée avec des méthodes s'inscrivant dans le cadre de la prévision hydrologique, ce qui semble renforcer la dépendance constatée, elle montre qu'un jeu de paramètres optimal censé représenter la totalité du cycle hydrologique est conditionné par des erreurs de mesure de débits sur quelques pas de temps seulement. Les recherches menés par Le Moine (2008) sur les pas de temps d'influence de l'erreur des modèles en fonction de plusieurs transformations de débits appuient d'ailleurs ces conclusions.
Outre les erreurs dans les données de débits, les erreurs dans les données d'entrée influencent également le résultat du calage d'un modèle hydrologique. Duethmann *et al.* (2020) rapportent par exemple que des variations dans le réseau pluviométrique peuvent expliquer une part importante des erreurs d'un modèle en contrôle. Même pour les modèles capables de corriger les erreurs dans les données d'entrée grâce à certains paramètres, ces erreurs peuvent éloigner ces paramètres des caractéristiques physiques des bassins versants (Andréassian, 2002). Oudin *et al.* (2006b) ajoutent que l'inclusion d'erreurs aléatoires dans les précipitations en entrée de modèle affectent considérablement ses performances et ses paramètres. Les mêmes auteurs notent que des erreurs systématiques dans les données d'évaporation potentielle influencent également les paramètres de la fonction de production du modèle. Bien que les erreurs ou les différences entre formules d'évaporation potentielle n'engendrent pas forcément de différences significatives dans les performances des modèles hydrologiques (Oudin *et al.*, 2005b, 2006b), ces données sont issues de conceptions et de modèles difficiles à comparer directement, hors du cadre de la modélisation hydrologique. De ce fait, le choix du modèle d'évaporation potentielle et les données nécessaires à son calcul participent également à l'incertitude dans l'estimation de paramètres robustes.

Une limite supplémentaire à l'identification d'un unique optimum hydrologique réside dans la teneur variable des données d'apprentissage en informations relatives à certains processus, notamment ceux dont l'influence sur le régime hydrologique ne se manifeste que sous certaines conditions. La manière la plus triviale d'accroître le contenu en informations utilisables pour caler les paramètres d'un modèle consiste à allonger la longueur des données de calage (e.g. Mathevet, 2005; Vrugt *et al.*, 2006; Arsenault *et al.*, 2018). Certains travaux montrent toutefois que la robustesse des paramètres plafonne à partir d'une certaine quantité de données (e.g. Harlin, 1991; Seibert et Beven, 2009), et qu'un choix adapté de ces données aide à réduire la longueur des chroniques nécessaire à l'identification de paramètres robustes (e.g. Perrin *et al.*, 2007). Ainsi, Singh et Bárdossy (2012) montrent que la teneur en information de longues chroniques pouvait être réduite à une poignée de pas de temps épars, témoignant d'événements météorologiques « inhabituels », suffisant au calage de paramètres robustes. Il se peut néanmoins que certains paramètres ne soient pas identifiables de manière robuste selon les conditions de modélisation (e.g. Perrin *et al.*, 2007), ce qui soulève la question déjà succinctement évoquée de la complexité des modèles que de telles ambitions permettent (Wagener *et al.*, 2003). Il n'est par exemple pas possible de conditionner les paramètres d'un modèle à représenter des situations de sécheresses prolongées tant que des sécheresses pluri-annuelles n'ont pas été observées dans un bassin versant (Fowler *et al.*, 2020). Certains auteurs recommandent toutefois d'intégrer malgré tout des comportements complexes, par exemple pour la représentation de dynamiques basse-fréquence, dimensionnés par des paramètres faiblement identifiables, plutôt que de renoncer totalement à les représenter (e.g. Reichert et Omlin, 1997).

La décision de n'adopter qu'une seule représentation paramétrique du comportement d'un bassin versant se heurte enfin à l'existence de jeux de paramètres équifinaux, c'est-à-dire qui atteignent des performances égales. Beven (2006) souligne que l'impossibilité de rejeter strictement plusieurs configurations d'un modèle à partir des éléments disponibles devrait conduire à les adopter ensemble. Cette incertitude devrait ainsi témoigner des lacunes structurelles des modèles hydrologiques. L'absence d'un processus essentiel à la dynamique de l'écoulement dans un bassin versant suscite par exemple une compensation par les paramètres dimensionnant les autres processus, par de multiples moyens. A l'inverse, le calage d'un modèle sur-paramétré induit l'existence de paramètres dont la valeur n'influence pas les performances.

Il est courant d'intégrer la présence d'incertitudes en modélisation par la production d'un ensemble de simulations simultanées. Certains travaux sont ainsi fondés sur la détermination de limites « d'acceptabilité », afin de déterminer les jeux de paramètres membres de l'ensemble des jeux de paramètres susceptibles d'incarner l'optimum hydrologique (e.g. Beven, 1993; Seibert, 2003; Blazkova et Beven, 2009; Liu *et al.*, 2009). Les jeux de paramètres à l'intérieur des frontières fixées par ces limites sont considérés comme également probables. Les limites d'acceptabilité sont établies à partir de l'expertise subjective qu'a le modélisateur de la qualité des données à disposition, des limites d'un modèle et des marges d'erreur qu'il est prêt à tolérer. Madsen (2000) souligne la différence entre l'approche par limites d'acceptabilité et celle de Pareto : l'une est une interprétation statistique de l'incertitude, l'autre représente l'équivalence de différents jeux de paramètres dans un cadre multi-objectif. La méthode des limites d'acceptabilité présente l'avantage majeur d'être simple d'application, dans la mesure où elle ne nécessite pas de connaissances préalables en statistiques, mais elle est est parfois critiquée comme manquant de rigueur (Beven *et al.*, 2008) et ne permet pas de distinguer les sources d'incertitudes (Beven et Freer, 2001). D'autres méthodes sont à l'inverse fondées sur une modélisation explicite de la structure des différentes sources d'incertitude, comme par exemple des erreurs dans les données d'entrée (e.g. Kuczera *et al.*, 2006), ou afin d'intégrer la transgression des hypothèses statistiques dans les méthodes de moindres carrés (e.g. Sorooshian et Dracup, 1980; Evin *et al.*, 2013).
Admettant la dépendance du jeu de paramètres identifié en calage au choix de la fonction objectif, certaines recherches explorent le recours à des techniques de pondération de différentes réalisations du calage d'un modèle en fonction des conditions de simulation. Oudin *et al.* (2006a) montrent par exemple l'intérêt d'indexer le poids respectif de deux jeux de paramètres obtenus avec des fonctions objectif visant respectivement les crues et les étiages à l'humidité du sol estimée par un modèle. D'autres approches consistent à estimer des poids différents à chaque pas de temps à partir des performances obtenus avec plusieurs jeux de paramètres ou modèles (e.g. Najafi *et al.*, 2011).

2.2.3. Perspectives pour un calage plus robuste des modèles hydrologiques

2.2.3.a. Manque de comparaison de fonctions objectif simples

Le calage robuste des modèles hydrologiques est confronté à plusieurs difficultés, parmi lesquelles on peut lister l'inefficacité des algorithmes, les propriétés statistiques des débits et de l'erreur des modèles défavorables à l'application des théorèmes classiques de la statistique, la multiplicité des processus hydrologiques à considérer, les incertitudes dans les données d'apprentissage, l'équifinalité des paramètres... Ces difficultés ont suscité nombre de recherches méthodologiques pour réussir à les contourner, afin de se rapprocher de l'optimum hydrologique. Ces recherches ont débouché sur une pléthore de fonctions objectif utilisables seules ou en association dans un cadre

multi-objectif.
Les conclusions du début du Chapitre 2 sur le manque de robustesse des modèles sont toutefois issues de travaux fondés pour la grande majorité sur des fonctions objectif aux moindres carrés (RMSE, NSE), parfois complexifiées pour inclure une prise en compte du biais. Le KGE est également utilisé dans quelques études récentes (Coron *et al.*, 2014; Bisselink *et al.*, 2016; Fowler *et al.*, 2016; Dakhlaoui *et al.*, 2017, 2019), de pair avec une transformation des débits par la fonction racine carrée dans le cas du diagnostic de Coron *et al.* (2012). Un poignée d'études sur la robustesse climatique des modèles se situent enfin dans un cadre plus fin de prise en compte de l'incertitude paramétrique, et déterminent un ensemble de paramètres dans un cadre bayésien (Brigode *et al.*, 2013; Broderick *et al.*, 2016; Motavita *et al.*, 2019), ou plus simplement dans le cadre de limites d'acceptabilité (Seibert, 2003; Li *et al.*, 2012; Kelleher et Shaw, 2018).
De ce fait, le rôle des fonctions objectif aux moindres carrés dans le consensus de manque de robustesse des modèles hydrologiques n'est pas clair. Les défauts qui leur sont connus expliquent peut-être en partie le manque de transférabilité des paramètres identifiés en calage. Il convient de noter que les approches bayésiennes sont généralement évaluées au regard de la cohérence des jeux de paramètres sélectionnés en comparaison des méthodes aux moindres carrés, afin de témoigner de leur faible dépendance aux données utilisées pour le calage (e.g. Thyer *et al.*, 2009). Bien que cette caractéristique soit intimement liée à celle de transférabilité de l'optimum hydrologique, il n'est pas certain que ces deux objectifs se recoupent tout à fait. De plus, ces méthodes sont pour la plupart présentées dans des études dont la forme est proche de la *proof of concept*, à l'essai sur une poignée de bassins versants pour un seul modèle hydrologique. Ces fonctions objectif ayant rarement fait l'objet de comparaisons dans un cadre de simulation en conditions climatiques changeantes, aucune fonction objectif ne fait véritablement office de référence par défaut pour cette application.

Quelques pistes suggèrent pourtant que les fonctions objectif habituelles contribuent au manque de robustesse des modèles hydrologiques diagnostiqués. En effet, Fowler *et al.* (2016) ont établi dans une étude sur plusieurs dizaines de bassins australiens que le calage avec le KGE ne permettait pas de sélectionner des jeux de paramètres capables de simuler convenablement deux périodes très contrastées, même dans les cas où de tels jeux existaient. Ces situations se produisant pour un nombre remarquable de bassins versants et de modèles, les auteurs concluent que certaines structures de modèles hydrologiques sont injustement considérées comme inadaptées à des conditions climatiques changeantes. Les mêmes auteurs comparent sur le même échantillon plusieurs fonctions objectif identifiées parmi celles utilisées dans les travaux sur la robustesse des modèles hydrologiques, et montrent que les jeux de paramètres sélectionnés par le KGE et le NSE sont parmi les moins robustes (Fowler *et al.*, 2018b). Ils proposent en remplacement l'utilisation d'une fonction objectif adaptée du KGE, appelée *Split* KGE, calculée comme la moyenne des KGEs annuels sur la chronique de calage, ou bien celle du *Refined Index of Agreement* (Willmott *et al.*, 2012), qui met en perspective la magnitude des erreurs absolues du modèle et la magnitude de la variabilité des observations. Les performances obtenues avec le NSE associé à la racine des débits sont également parmi les meilleures.
L'évaluation de l'efficacité des fonctions objectif traditionnellement utilisées et de leurs alternatives dans un cadre de calage-contrôle contrasté apparaît donc comme une piste intéressante pour l'amélioration de la robustesse des modèles.

2.2.3.b. Vers une caractérisation plus détaillée des objectifs de calage

Au-delà des strictes performances en DSST, les découvertes de Coron *et al.* (2014) exigent que les modèles hydrologiques soient non seulement capables de modéliser le débit dans des conditions climatiques différentes de celles rencontrées en calage, mais également de représenter ses variations interannuelles au sein de la période de calage. En effet, ces travaux ont montré que le

calage sur une partie ou la totalité des données ne changeait pas la structure de la courbe des biais interannuels du modèle. Ce constat pose la question des rôles respectifs de la structure du modèle et du processus de calage. Il se peut en effet que les méthodes actuelles ne contraignent pas suffisamment la sélection des paramètres dans la représentation des phénomènes caractéristiques des échelles interannuelles, ou bien que les lacunes se situent principalement dans la structure des modèles, et que le seul degré de liberté réellement exploité en calage soit l'ajustement du biais du modèle.

De ce fait, la recherche de meilleures techniques de calage doit répondre à un double objectif :

- Diminuer les pertes de performances en contrôle dans des conditions d'extrapolation climatique, dans le cadre du DSST ou d'expériences adaptées. Cet objectif répond au besoin d'exploiter au mieux l'information hydrologique des données de calage même lorsque celles-ci ne témoignent pas de toute la variabilité climatique subie par le bassin versant.
- Aplatir la courbe des biais sur les sous-périodes glissantes, dans des conditions d'interpolation. Cet objectif est associé à la nécéssité de valoriser l'information contenue dans les données de calage lorsque celle-ci est complète.

L'aplatissement de la courbe de biais sur les sous-périodes glissantes passe nécessairement par la diminution des écarts entre les maxima et les minima, dont la caractérisation peut orienter la recherche de méthodes de calage plus performantes. Ces extrema renseignent d'ailleurs également sur les pires conditions de transfert de jeux de paramètres, qui ne sont pas bien définies et varient selon les contextes des études relevées précédemment. Caractériser ces situations demande donc de chercher au delà des simples variations climatiques en termes de température ou de précipitations, par l'analyse de caractéristiques hydro-climatiques plus fines.
Ces recherches pourraient permettre d'articuler un diagnostic plus abouti des défauts structurels des modèles hydrologiques, par l'identification des degrés de liberté manquants pour représenter l'évolution de la réponse des bassins versants avec les conditions climatiques.

2.3. Causes du manque de robustesse : structures inadaptées

Dans les constats formulés sur le manque de robustesse des modèles conceptuels, la structure et les équations de ces modèles sont susceptibles de jouer un rôle majeur. Néanmoins, les interactions des défaillances structurelles avec les données à travers l'estimation des paramètres rendent leur identification difficile. Dans cette section sont présentés les enjeux et les principes de l'évaluation de la structure des modèles, ainsi qu'un inventaire des lacunes connues ou perçues des modèles actuels limitant leur robustesse.

2.3.1. Principes de l'évaluation de la structure des modèles

2.3.1.a. Compensation des erreurs structurelles en calage

Le choix des paramètres d'un modèle hydrologique est déterminant dans sa capacité à simuler convenablement le régime d'un bassin versant dans une certaine variété de conditions. Un calage idéal permet de dimensionner les processus décrits dans les équations du modèle à la mesure des caractéristiques physiques intrinsèques du bassin. L'un des enjeux d'un tel calage est de correc-

tement établir la valeur des paramètres associés aux processus du modèle ayant peu d'influence sur les écoulements, au moins dans les données d'apprentissage. En revanche, la robustesse d'un modèle hydrologique est nécessairement limitée dans le cas inverse, c'est-à-dire lorsque certains processus prépondérants ne sont pas intégrés à sa structure, ou que leur formulation est inadéquate. On peut considérer que l'absence ou la description maladroite de tels processus relèvent d'une erreur de modélisation, qu'on désignera généralement comme une « erreur structurelle ». La compensation des erreurs structurelles par le calage des paramètres ajoute à l'incertitude sur les paramètres estimés en calage (Wagener *et al.*, 2003; Beven, 2006). Dans la mesure où les incertitudes causées par les erreurs structurelles ne sont pas facilement séparables des autres sources d'incertitudes, l'évaluation indépendante de la structure des modèles est un exercice complexe.

On peut estimer que dans le meilleur cas, le calage sélectionnera correctement les paramètres des processus du modèle qui correspondent à la réalité du bassin versant. Formulé plus précisément, les paramètres des processus décrits dans le modèle ne doivent pas servir à compenser l'absence d'autres processus nécessaires à la représentation de certains phénomènes réellement à l'œuvre dans le bassin. Hooper *et al.* (1988) parviennent par exemple à identifier des voies de simplification d'un modèle hydrogéochimique à partir d'un calage multi-objectif visant la reproduction des débits et de la concentration d'un traceur chimique. Les auteurs suggèrent que les paramètres trop peu contraints par l'intégration du traceur chimique dans le processus de calage sont superflus en l'absence de données supplémentaires et que ce sur-paramétrage conduit à des simulations irréalistes du traceur.
Dans le pire des cas, les paramètres des processus présents seront utilisés pour compenser les processus absents, nuisant à sa robustesse. Comme le note Kirchner (2006), l'évaluation de la structure du modèle nécessite alors des procédures spécifiques de test pour vérifier si le modèle donne les bonnes réponses pour les bonnes raisons. Bien que le DSST soit un outil pertinent pour diagnostiquer et caractériser les maux dont souffrent les modèles, leur attribution à des erreurs structurelles n'est pas triviale dans la mesure où les choix faits en calage peuvent fortement influencer les résultats (Fowler *et al.*, 2016, 2018a). Une méthode possible de repérage des erreurs structurelles consiste en l'analyse des variations des paramètres en fonction du temps, afin de reconnaître ceux jouant un rôle de compensation. Cette analyse peut être réalisée en décomposant les données en de multiples périodes de calage (e.g. Merz *et al.*, 2011; Kelleher et Shaw, 2018), ou par l'analyse de la sensibilité des erreurs sur chaque pas de temps par un tirage aléatoire de paramètres (e.g. Wagener *et al.*, 2003).

Ces méthodes font toutefois face à des difficultés liées à l'interprétation des paramètres instables. En effet, les variations d'un paramètre ne témoignent pas nécessairement d'un défaut de formulation du processus déterminé par ce paramètre, mais parfois d'un défaut présent ailleurs dans la structure du modèle. Par exemple, Kelleher et Shaw (2018) relèvent que les variations temporelles des paramètres liés aux caractéristiques du sol, qui peuvent bénéficier aux performances d'un modèle, sont généralement irréalistes dans la mesure où dans la plupart des bassins les propriétés réelles du sol évoluent lentement. Les auteurs suggèrent plutôt que ces variations signalent l'incapacité des formules d'évaporation à représenter fidèlement les changements interannuels d'évaporation.
Par ailleurs, les paramètres des modèles complexes sont inévitablement peu identifiables et sont amenés à prendre des valeurs instables (Kuczera et Mroczkowski, 1998). Ces variations simultanées peuvent masquer les éléments structurels défaillants. Kirchner (2006) souligne que la grande flexibilité des modèles complexes nuit de ce fait non seulement à l'identification de leurs paramètres mais également à celle des formulations incorrectes des processus décrits.

2.3.1.b. Plausibilité des processus internes

Une autre approche que celle inférant les erreurs structurelles aux paramètres consiste à questionner la plausibilité des processus et des flux simulés. D'après Fowler *et al.* (2020), cette approche fait défaut aux études d'impact du changement climatique, dans lesquelles la description n'est que très rarement accompagnée d'une critique des choix de modélisation. Même lorsqu'elle est proposée, la simple comparaison des performances ne suffit pas à évaluer la plausibilité de différents modèles. La confrontation à partir de critères généraux sur les débits simulés permet en effet seulement de juger de « l'agrégation des hypothèses » constituant un modèle, mais pas des « hypothèses prises individuellement » aux fondements de chacun de ses composants (Bouaziz *et al.*, 2020). Les difficultés de validation des composants de la structure de modèles sont dûes à un manque d'information permettant d'isoler et critiquer individuellement ces composants. L'implication de différentes sources d'information hydrologiques peut à cet égard guider les diagnostics et disqualifier les hypothèses les moins plausibles (Kuczera et Mroczkowski, 1998; Kirchner, 2006).

Dans le cas de modèles distribués ou semi-distribués, les conceptualisations régionales peuvent être testées par une validation interne sur les observations de débit aux exutoires de sous-bassins versants (e.g. Ambroise *et al.*, 1995; Moussa *et al.*, 2007). Néanmoins, cette approche n'est pas envisageable pour les modèles globaux tels que ceux typiquement utilisés dans les études sur la robustesse climatique.
Pour cibler certains processus décrits dans les modèles, l'analyse de leur plausibilité interne peut s'appuyer sur l'utilisation des signatures hydrologiques, déjà mentionnées dans le cadre du calage des paramètres des modèles (e.g. Euser *et al.*, 2013). En analysant les simulations de débits à l'échelle d'événements spécifiques témoignant de processus contrastés, de Boer-Euser *et al.* (2017) notent que des modèles calés selon un même protocole par un ensemble de laboratoires, montrant des performances similaires au regard de critères généraux, pouvaient produire des prédictions très hétérogènes des épisodes analysés. Les auteurs dégagent des structures plus performantes pour la représentation de certaines catégories d'événements, mais constatent qu'aucune structure unique ne se démarque pour simuler l'ensemble. Dans une étude plus récente, Bouaziz *et al.* (2020) poursuivent l'analyse par la confrontation des états internes des modèles à des produits satellitaires d'évaporation, de surfaces enneigées, d'humidité du sol et de masse d'eau totale (gravimétrie). Les auteurs questionnent la cohérence des états et flux internes entre les modèles, en suggérant que des débits simulés identiques qui résultent de descriptions différentes des états internes ne peuvent pas être simultanément plausibles (Bouaziz *et al.*, 2020).

2.3.1.c. Structures flexibles ou modèle unique ?

La recherche de structures adéquates s'envisage enfin au moyen d'approches flexibles, consistant à tester plusieurs configurations de modèles et ne retenir que la plus performante pour chaque bassin. Cette approche résulte du constat que la transposabilité d'une unique structure dans l'espace est difficile à garantir, dans la mesure où elle procède d'une hypothèse unique sur les processus à l'œuvre dans n'importe quel bassin versant. L'approche flexible se fonde sur l'essai d'hypothèses multiples (Clark *et al.*, 2011). On notera que plutôt que de clamer que certains types de bassins appellent certains types de modèle, l'approche par structure flexible entérine plus modestement l'idée que la modélisation de bassins très différents nécessite vraisemblablement des représentations différentes (Kavetski et Fenicia, 2011). Différentes méthodologies de structures flexibles ont été proposées. La méthode SUPERFLEX (Kavetski et Fenicia, 2011; Fenicia *et al.*, 2011) suggère la complexification incrémentale d'une structure de base, notamment afin d'éviter l'emploi de modèles sur-paramétrés pour la simulation de bassins simples. La structure flexible SUMMA (Clark *et al.*, 2015a,b) est une approche à base physique fondée sur un adaptation préalable des termes intervenant dans le bilan de masse, puis de l'échelle de discrétisation spatiale et du paramétrage des processus.

Les approches structurelles flexibles répondent au défi que pose, entre autres, l'identifiabilité et le calage des paramètres, notamment ceux liés aux processus simulés n'ayant que peu d'influence réelle sur le débit observé. Ces méthodes s'appuient sur le constat de bon sens que ces processus devraient pouvoir être inclus ou mis de côté selon les bassins considérés. Il paraît à ce titre envisageable que les capacités extrapolatives de modèles choisis en connaissance de cause, ou simplement par comparaison et élimination, seraient meilleures que celles de structures uniques appliquées dans des contextes très hétérogènes. Les études impliquant des approches flexibles font particulièrement état de la plus-value de ces approches pour la simulation de bassins particuliers, où les modèles empiriques « rigides » sont sous-performants, bien que ces derniers s'avèrent être de bons compromis pour la majorité des bassins étudiés (Kavetski et Fenicia, 2011; Van Esse *et al.*, 2013). Néanmoins, Seiller *et al.* (2017) relèvent que certaines difficultés de l'approche par structures flexibles résident dans le choix *a priori* des composantes pertinentes à tester, ainsi que dans l'analyse des raisons des succès ou échecs des structures qu'elles composent.
S'inscrivant également dans la perspective d'exploiter plusieurs structures de modèles pour chaque bassin, les approches ensemblistes multi-modèles permettent de contourner ce problème. Utiliser les réalisations d'un ensemble de modèles dispense de la délicate étape du choix d'une structure finale, supposée être la plus adaptée au contexte de modélisation. En conditions d'extrapolation climatique, cette approche montre de bons résultats du fait de la compensation des erreurs structurelles des modèles moins pertinents (e.g. Seiller *et al.*, 2012).

Bien que l'utilisation d'un ensemble de modèles ou de structures flexibles soit prometteuse, on peut considérer que le développement de modèles rigides, destinés a priori à un emploi « en solitaire », n'a pas atteint son stade terminal. En effet, les travaux et réflexions à ce sujet participent à l'identification des composantes essentielles que doivent posséder les modèles pour fournir des réponses efficaces aux problèmes qu'ils sont amenés à résoudre. Dans le cadre d'usages en extrapolation climatique, certains défauts connus ou suggérés des modèles actuels, comme ceux évoqués dans les paragraphes qui suivent, ne semblent pas complètement solubles dans leur mélange et leur combinaison (Fowler *et al.*, 2020). Le développement de structures uniques stimule la recherche de mises en équation robustes des processus décrits et proposés pour régler ces défauts, par exemple pour en diminuer la complexité matérialisée par le nombre de variables d'états ou de paramètres libres (Seiller *et al.*, 2017). Enfin, Kavetski et Fenicia (2011) notent que pour les cas de bassins complexes, des structures uniques peuvent être utilisées comme structure de base pour une modélisation plus adaptée au contexte local.
Pour ces raisons, les travaux de cette thèse portent sur la robustesse des structures rigides, non incorporées à des simulations d'ensemble de modèles, malgré les liens forts entre transposabilités temporelle (robustesse) et spatiale (régionalisation) des modèles. Les paragraphes suivants sont consacrés à l'inventaire des lacunes structurelles communes aux modèles hydrologiques.

2.3.2. Lacunes identifiées en conditions climatiques variables

2.3.2.a. Représentation de l'évaporation

À l'échelle globale, environ 60% de l'eau des précipitations continentales est restituée à l'atmosphère sous forme d'évaporation, qui constitue de ce fait une part essentielle du bilan en eau des bassins versants (Oki et Kanae, 2006). La capacité des modèles à représenter les flux d'eau évaporée est donc déterminante pour pouvoir espérer simuler correctement les effets du changement global sur les régimes hydrologiques régionaux. Toutefois, les perspectives d'évaluation de cette capacité sont fortement limitées par les difficultés majeures auxquelles fait face la modélisation hydrologique à l'échelle caractéristique du bassin versant, liées à la faible disponibilité d'observations d'évaporation réelle et à la conceptualisation théorique de l'évaporation potentielle, dont la mesure

n'est pas possible.
La structure même des modèles hydrologiques complexifie en outre l'évaluation de l'évaporation modélisée via les simulations de débit, du fait des spécificités du calcul de l'évaporation, traditionnellement décomposé en deux étapes. La première consiste à estimer l'évaporation potentielle à partir des variables météorologiques, la seconde à déterminer la part véritablement évaporée de cette quantité potentielle à partir de la disponibilité en eau du bassin versant. Cette décomposition permet aux paramètres libres des modèles de compenser jusqu'à un certain point les erreurs présentes dans l'évaporation potentielle estimée au préalable et réduit leur influence sur le débit simulé à l'exutoire (Section 2.2.2.c). De ce fait, l'amélioration de la reproduction du bilan hydrique du modèle exige à la fois des estimations plus justes de l'évaporation potentielle et un perfectionnement de la représentation des mécanismes de transformation de l'évaporation potentielle en évaporation réelle.

De nombreux modèles d'évaporation potentielle coexistent, sans qu'il soit évident de départager les formulations les plus robustes des moins robustes à partir des performances des modèles hydrologiques. Dans la perspective d'une utilisation en étude d'impact du changement climatique pourtant, le choix n'est pas anodin selon les contextes régionaux. Guo *et al.* (2017b) montrent par exemple que les évaporations potentielles estimées par deux modèles différents présentent des sensibilités hétérogènes aux variables climatiques, qui dépendent des caractéristiques hydro-climatiques des bassins considérés. Les auteurs notent néanmoins que la température est le facteur dominant de variation de l'évaporation potentielle quel que soit le modèle hydrologique considéré. Les différents facteurs déterminant les changements d'évaporation potentielle peuvent également influencer les projections hydrologiques réalisées avec des modèles de surface continentale, qui sont particulièrement sensibles à la formulation de l'évaporation potentielle utilisée dans les zones limitées en énergie, où la demande évaporatoire est en moyenne supérieure aux précipitations moyennes (Milly et Dunne, 2016).
Le choix d'une formule adéquate pour l'évaporation potentielle n'est toutefois que faiblement guidé par les expériences de validation par le biais de la modélisation hydrologique, pour les raisons de compensation évoquées précédemment. Jayathilake et Smith (2020b) notent que les formules d'évaporation sont généralement assorties au contexte de modélisation en fonction de la disponibilité des observations des variables climatiques et de la complexité des modèles hydrologiques employés. Les formules utilisées les plus simples ne sont calculées qu'à partir de la seule température moyenne journalière (e.g. Oudin *et al.*, 2005a), tandis que d'autres dérivant de conceptions moins empiriques utilisent jusqu'à quatre variables différentes : température, humidité relative, rayonnement et vitesse du vent (e.g. Penman, 1948; Monteith, 1965). Si les modèles simples semblent suffire pour la modélisation hydrologique conceptuelle (Oudin *et al.*, 2005a), certains travaux insistent sur l'importance de sélectionner des modèles plus réalistes pour des modèles plus complexes, comme les modèles éco-hydrologiques fondés sur une représentation plus explicite de la végétation, et dans les régions limitées en énergie (Jayathilake et Smith, 2020b).

L'estimation de l'évaporation potentielle n'est pas la seule source d'incertitude qui affecte la modélisation de l'évaporation. Le processus de conversion de l'évaporation potentielle en évaporation réelle suscite également de nombreuses représentations différentes dans les modèles hydrologiques, qui illustrent à la fois la diversité des spécificités régionales et des conceptions qui en découlent et le souci perpétuel d'arbitrer en simplicité et réalisme. Fowler (2017) relève par exemple la diversité des réponses de la végétation à l'assèchement des conditions climatiques en Australie et de leurs conséquences sur le régime hydrologique dans les travaux consacrés à ce sujet : mortalité des arbres (Allen *et al.*, 2010; Williams *et al.*, 2013), évolution de l'équilibre entre différentes espèces (e.g. Hughes, 2003), régénération post-incendies (Kuczera, 1987), fertilisation liée à l'augmentation de la teneur atmosphérique en carbone (Peel, 2009; Ukkola *et al.*, 2016; Ajami *et al.*, 2017), augmentation du rendement en eau de la production primaire (Troch *et al.*, 2009), évolution du réseau racinaire (Troch *et al.*, 2015)... De tels phénomènes peuvent avoir des conséquences

variées sur le rendement des précipitations et la résistance à la demande évaporatoire, qu'il n'est pas aisé de réduire à des formulations parcimonieuses.
Les représentations diverses du processus d'évaporation dans les modèles semblent n'avoir que peu d'influence sur les performances des modèles hydrologiques à simuler le débit (Bai *et al.*, 2015). Pourtant, Guo *et al.* (2017a) démontrent l'influence majeure qu'ont ces représentations sur la sensibilité du débit et de l'évaporation réelle simulés aux variations de variables climatiques déterminant l'évaporation potentielle, et soulignent ainsi l'importance du choix des modèles hydrologiques dans les études d'impact du changement climatique. Les différences constatées seraient explicables dans le contexte de ces travaux notamment par les représentations hétérogènes de l'évaporation pendant les jours pluvieux dans les modèles hydrologiques testés. Les auteurs notent enfin que la représentation du processus d'évaporation dans les modèles semble avoir d'autant plus d'effet dans les bassins secs et limités en eau, marqués par une forte saisonnalité de l'humidité du sol. Cette conclusion paraît compléter celles tirées des études précédemment citées sur l'évaporation potentielle, dont l'estimation correcte est primordiale dans les régions humides et limitées en énergie.

2.3.2.b. Représentations des dynamiques interannuelles

La question du rôle de la réponse de la végétation succinctement abordé dans les paragraphes précédents s'incorpore dans une problématique plus vaste que celle de l'évaporation, qui concerne les dynamiques de modification de la réponse du bassin versant aux forçages climatiques. La compréhension et la simulation de ces dynamiques, en particulier aux échelles temporelles saisonnières ou interannuelles, représentent un enjeu important de la compréhension des cycles hydrologiques. Les réserves d'eau d'un bassin versant répondent en effet aux anomalies climatiques (e.g. Luque-Espinar *et al.*, 2008) et induisent des variations de la relation pluie-débit (e.g. Hughes *et al.*, 2012). D'une certaine façon, ces réserves capitalisent la « mémoire » qu'aurait le bassin versant des événements climatiques passés. Dans le cadre du changement climatique, lorsque ces anomalies sont persistantes ou correspondent à des modifications durables des forçages climatiques, il est essentiel que les modèles soient en mesure de reproduire ces variations. Toutefois, certains indices montrent que cette capacité fait en général défaut aux modèles hydrologiques.

Le parallélisme quasi systématique des courbes de biais glissants relevé par Coron *et al.* (2014) constitue une première piste de réflexion. En effet, celui-ci traduit une grande stabilité des variations interannuelles du débit simulé par les modèles : l'écart entre le débit simulé d'une période A et celui d'une période B semble n'être pas dépendant des jeux de paramètres, du moins ceux obtenus en calage. Pourtant les paramètres libres de modèles pluie-débit leur procurent une grande flexibilité, notamment dans la reproduction de la saisonnalité et des volumes moyens écoulés. Cette flexibilité ne se détectant apparemment presque pas dans les dynamiques interannuelles modélisées, cela pourrait être le signe que celles-ci ne dépendent pas des paramètres, ou qu'elles ne sont pas représentées par les composantes des modèles. Il convient toutefois de noter que cette analyse est une interprétation, et que sa crédibilité nécessite notamment de généraliser les résultats d'une seule étude impliquant trois modèles et vingt bassins. La transposabilité des résultats à d'autres méthodes de calage, d'autres régions et d'autres modèles devrait donc être évaluée.

Les travaux de Saft *et al.* (2016a) éclairent néanmoins plus directement une difficulté des modèles à simuler correctement les dynamiques interannuelles des bassins versants, à partir de l'analyse de la modélisation de sécheresses décennales sur des bassins australiens. En effet, leurs résultats suggèrent que les modèles hydrologiques testés produisent des simulations biaisées de ces événements en contrôle uniquement lorsque la relation pluie-débit du bassin a changé entre les conditions de calage et les épisodes des sécheresses évalués en contrôle. Les auteurs notent que des représentations inadaptées des processus à l'œuvre dans les propriétés du sol, dans la végétation ou les réserves souterraines pourraient expliquer pourquoi les modèles peinent à simuler correctement

les variations de la réponse des bassins versants concernés. Les recherches de Saft *et al.* (2016b) confirment que les changements observés dans les relations pluie-débit des bassins étudiés sont liés à des mécanismes internes au bassin versant et écartent l'hypothèse selon laquelle les erreurs seraient provoquées par des données de débit ou de précipitations altérées.

Enfin, Fowler *et al.* (2020) ont montré récemment que de nombreux modèles n'étaient pas en mesure de reproduire le déclin significatif des réserves souterraines observé dans un ensemble de bassins australiens. L'impossibilité des modèles à simuler des dynamiques longues et lentes serait liée au fait que les cycles saisonniers de leurs réservoirs couvrent déjà presque toute l'amplitude de variation autorisée. Ceux-ci étant dès lors presque vides en période d'étiage, aucun déficit supplémentaire ne peut s'accumuler au fil des années. Fowler (2017) ajoute que les modèles dont le réservoir le plus lent varie sur les temps caractéristiques de la récession du débit de base ne peuvent probablement pas simuler correctement des bassins à mémoire hydrologique interannuelle. L'utilisation de modèles « à déficit », dont le remplissage des réservoirs n'est pas limité par des bornes supérieures et inférieures, est perçue par l'auteur comme une voie d'amélioration pour la modélisation de sécheresses persistantes. Néanmoins, bien qu'elles apparaissent prometteuses, ces formulations nécessitent des développements futurs pour être à même de reproduire de manière plausibles l'accumulation observée de déficits en eau (Fowler *et al.*, 2019).

2.4. Synthèse

Caractérisation de la robustesse climatique des modèles hydrologiques conceptuels

La robustesse des modèles hydrologiques correspond à la capacité de ces modèles à simuler le régime hydrologique de bassins versants dans des conditions climatiques différentes de celles rencontrées lors du calage de leurs paramètres. Pour évaluer la robustesse des modèles, on peut recourir à la méthode du *Differential Split-Sample Test* qui consiste à distinguer dans une chronique historique d'observations deux périodes témoignant de conditions hydro-météorologiques contrastées. Les deux périodes ainsi définies peuvent servir à tour de rôle de période de calage et de période de contrôle.
Il a été constaté dans la grande majorité des études sur le sujet que les modèles manquaient généralement de robustesse, c'est-à-dire que leurs performances se dégradent dans des conditions climatiques différentes de celles rencontrées en calage. La simulation des variations temporelles du bilan hydrique des bassins versants constitue le défaut majeur relevé dans ces situations. Des liens qui unissent les variations du climat, des performances et des paramètres des modèles sont parfois détectables mais difficiles à exploiter en pratique pour corriger les erreurs. Enfin, il est apparu que si les conditions de calage déterminent le biais des modèles, elles n'ont qu'une influence limitée sur les variations interannuelles de ce biais, dont la forme reste stable.
Les travaux ayant abouti à ces constats évoquent souvent un faisceau de pistes pour expliquer le manque de robustesse des modèles hydrologiques. Nous avons identifié deux pistes majeures dont la pertinence est corroborée par des études récentes.

Rôle du calage des paramètres des modèles hydrologiques

La première piste s'articule autour du rôle des méthodes de calage des modèles dans le manque de robustesse constaté. Bien qu'aucun algorithme de recherche ne soit parfait et ne garantisse l'identification du jeu de paramètres optimal, il semble que le facteur limitant de la robustesse des modèles soit avant tout le choix de la fonction objectif de calage.

Traditionnellement, les fonctions objectifs employées sont des estimateurs aux moindres carrés issus de la théorie statistique des régressions linéaires. Toutefois, de nombreux aspects amènent à questionner la validité des hypothèses formulées pour l'utilisation de tels estimateurs dans le contexte particulier de la modélisation hydrologique. On constate par exemple que les interactions entre les erreurs de mesure et les erreurs structurelles des modèles éloignent l'optimum mathématique des fonctions objectif de calage usuelles de l'optimum hydrologique, c'est-à-dire du jeu de paramètres le plus robuste.
Par ailleurs, de nombreux jeux de paramètres mènent à des performances similaires, dénotant une équifinalité. Enfin, la distribution de l'information utile pour l'estimation des paramètres est inégalement répartie dans les données utilisées en calage. Cette inégale répartition invite à utiliser des critères de performances mettant en valeur l'erreur du modèle aux pas de temps décisifs afin de valoriser au mieux l'information contenue dans les données de calage. De ce fait, il existe une pléthore de critères de performance utilisables comme fonctions objectif pour contraindre les modèles à reproduire différents aspects du régime hydrologique observé.
Malgré cette richesse, un petit ensemble de fonctions objectif a été utilisé dans les travaux relevant les problèmes de robustesse des modèles hydrologiques en extrapolation climatique. En outre, ces méthodes de calage usuelles ont été très rarement comparées entre elles dans le cadre spécifique de la transférabilité des paramètres.

Rôle de la structure des modèles hydrologiques

La deuxième piste a été identifiée à travers les questionnements relevés dans la littérature à propos du rôle de la structure des modèles hydrologiques dans les défaillances des modèles en extrapolation. Un frein majeur à l'identification de ces défaillances est la possibilité pour les paramètres de compenser ces erreurs structurelles lors du calage. Ainsi, des débits modélisés ou des performances similaires peuvent masquer des différences entre les états et flux internes modélisés par plusieurs modèles, voire plusieurs jeux de paramètres d'un même modèle.
Les performances des modèles sont notamment peu sensibles au modèle d'évaporation potentielle et à la représentation interne du processus d'évaporation réelle. Dans la mesure où le manque de robustesse des modèles est marqué par une difficulté importante à reproduire les variations des volumes écoulés, on peut soupçonner que la simulation de l'évaporation est un point critique de l'évaluation et du développement de modèles plus robustes puisque c'est un terme prépondérant du bilan hydrique dans la plupart des bassins. Le terme d'évaporation fait toutefois intervenir des processus complexes, liés aux dynamiques interannuelles de la végétation et des réserves souterraines des bassins versants. Or des études récentes indiquent justement que de nombreux modèles semblent peiner à représenter ces variations, notamment lors de sécheresses persistantes. Ce constat témoigne d'une difficulté des modèles à garder en mémoire les effets de conditions climatiques passées sur les réserves d'eau des bassins versants.

Problématique et objectifs de la thèse

Les défauts des modèles hydrologiques sont préoccupants dans le cadre d'une utilisation en changement climatique. Nous allons tenter de contribuer à leur diagnostic approfondi en vue de leur amélioration, en répondant aux deux questions principales suivantes, décomposées en six sous-questions :

- **Comment améliorer la robustesse des modèles hydrologiques sans modifier leurs équations et leur structure ?**
  Cette question centrale pose avant tout celle de la manière d'évaluer la robustesse des modèles, c'est-à-dire du choix des méthodes et des critères de performance (sous-question n°1). En complément des méthodes relevées dans la littérature, nous avons proposé un estimateur de la robustesse des modèles synthétique et utilisable également pour des modèles plus

complexes que ceux étudiés dans le cadre de la thèse. Une fois ce cadre d'analyse défini, nous avons cherché à déterminer les changements hydro-climatiques les plus défavorables aux performances afin de mieux caractériser les difficultés rencontrées par les modèles (sous-question n°2). La suite des développements s'est concentrée sur la recherche de stratégies de calage permettant d'identifier les jeux de paramètres les plus robustes, à travers la comparaison de plusieurs fonctions objectif (sous-question n°3), ainsi que sur le choix du modèle d'évaporation potentielle (sous-question n°4). En répondant à ces questions, nous avons tenté de définir un cadre de modélisation favorisant la robustesse des modèles existants dans des conditions de simulation difficiles. Outre la possibilité de dresser une liste des « bonnes pratiques » pour l'élaboration d'études d'impacts hydrologiques du changement climatique, ce cadre est censé permettre de révéler des défauts structurels des modèles en limitant le plus possible les doutes concernant le calage et les données utilisées.

- **Peut-on identifier au sein d'un modèle hydrologique simple des défauts structurels limitant sa robustesse et peut-on les corriger ?**
  Nous avons considéré qu'une étape essentielle à franchir avant de pouvoir amender le modèle de façon pertinente consistait en l'exploration des interactions entre les éléments de la structure et la robustesse du modèle. Nous avons pour y parvenir cherché à identifier les composantes jouant un rôle limitant dans la robustesse du modèle GR4J (sous-question n°5). Nous avons pour cela mis en œuvre un cadre d'analyse multi-objectif pour l'étude des interactions entre robustesse du modèle, paramètres et états internes du modèle. Les derniers travaux menés ont enfin été focalisés sur la portée de modifications structurelles du modèle, proposées à la lumière des réponses aux questions précédentes (sous-question n°6).

Pour répondre à la première question, plusieurs modèles hydrologiques ont été employés afin de tirer des conclusions dépassant les particularités d'un modèle unique. Dans cette optique, nous avons également choisi de raisonner sur un grand échantillon de bassins versants situés en France et en Australie. La description des données et des modèles utilisés est proposée au Chapitre 3. Le Chapitre 4 est dévolu à la détermination du cadre d'évaluation de la robustesse. Le Chapitre 5 est construit autour de la recherche des meilleures méthodes de calage et de l'identification des changements hydro-climatiques mettant les modèles hydrologiques le plus en difficulté. Le Chapitre 6 traite du choix de la formule d'évaporation potentielle pour favoriser la robustesse des modèles hydrologiques. Une méthode d'identification des limites structurelles des modèles hydrologiques est exposée au Chapitre 7. Enfin, le Chapitre 8 fait état d'une tentative de modification de la structure de GR4J.

Chapitre 3.

Matériel et modèles

Introduction

Dans ce chapitre, nous présentons les bases de données hydro-météorologiques utilisées dans l'ensemble des chapitres suivants. Une description des modèles hydrologiques, d'évaporation potentielle ainsi que des critères de calages et critères d'évaluation des modèles hydrologiques est ensuite donnée.

3.1. Base de données

3.1.1. Motifs de sélection des bassins versants

3.1.1.a. Intérêt d'un large échantillon de bassins

Nous avons fait le choix de fonder notre analyse de la robustesse des modèles hydrologiques sur la simulation d'un grand nombre de bassins versants. Gupta *et al.* (2014) rapportent que les exercices de modélisation sur de larges échantillons de bassins présentent plusieurs avantages en comparaison des analyses réalisées sur des bassins spécifiques, même très instrumentés. Un premier avantage est la possibilité de comparer rigoureusement plusieurs modèles ou hypothèses de modélisation en facilitant l'estimation de leurs limites et de leurs domaines d'application. L'utilisation de tests statistiques permet à ce titre de mesurer la robustesse et la généralité des résultats obtenus et d'identifier des bassins remarquables (*outliers*). Mathevet (2005) en fait la démonstration en estimant le nombre minimal requis de bassins dans un échantillon pour différencier les performances de plusieurs modèles entre elles selon les différences moyennes entre elles et l'écart-type de leurs performances sur l'échantillon. L'auteur montre que dans la plupart des cas, un échantillon de plusieurs centaines de bassins est largement préférable à à un échantillon de plusieurs dizaines de bassins pour observer des différences significatives. Gupta *et al.* (2014) relèvent enfin que cette approche peut servir à estimer les incertitudes qu'il faudrait attendre en appliquant un modèle dans un contexte arbitraire, par exemple pour des bassins non jaugés.

3.1.1.b. Sélection des bases de données et tri des bassins

Les bassins versants utilisés dans le cadre de cette thèse proviennent de l'échantillon utilisé par Coron (2013). Cet échantillon est constitué de bassins versants situés en France métropolitaine et au sud-est de l'Australie. Ces deux régions connaissent des climats assez différents l'un de l'autre, et sont marquées par des fortes disparités climatiques intra-régionales. Le sud de l'Australie a subi un épisode de sécheresse très sévère entre 1997 et 2009, connu sous le nom de *Millenium Drought*, et dont une caractérisation a été proposée par Chiew *et al.* (2014) et par le *Bureau of Meteorology*[1]. De ce fait, les bassins de l'échantillon australien sont d'un intérêt particulier vis-à-vis du sujet de cette thèse dans la mesure où les modèles pourront être évalués dans des conditions de transfert très exigeantes. L'échantillon de bassins français présente quant à lui l'avantage d'être très souvent utilisé dans les travaux de l'équipe et donc bien connu, et d'avoir été sélectionné par Coron (2013) selon des critères de longueur et de qualité des chroniques de données. Ces deux échantillons ont le point commun de rassembler des bassins peu influencés par les activités humaines. Ils ne comportent pas de retenue notable et subissent des prélèvements nuls ou faibles compte tenu de leur débit. Enfin, les bassins des deux échantillons ont fait l'objet d'une critique des données par

[1] http://www.bom.gov.au/climate/updates/articles/a010-southern-rainfall-decline.shtml

Coron (2013) basée sur l'analyse des courbes de doubles-cumuls entre chaque bassin versant et ses plus proches voisins. Cette méthode simple permet d'identifier des ruptures anormales de cumul du débit observé en postulant que les débits cumulés de deux bassins proches suivent une loi de proportionnalité. Cette critique a débouché sur échantillon de 580 bassins français et 204 bassins australiens. Ces deux ensembles ont constitué notre échantillon de départ, que nous avons réduit en tenant compte de plusieurs critères.
Nous avons dans un premier temps sélectionné des bassins peu influencés par la neige en retirant de l'échantillon les bassins dont la part annuelle moyenne de précipitations neigeuses est supérieure à 10%. Cet égard nous a permis de simplifier le paramétrage de la modélisation de la neige dans les modèles hydrologiques. Certains bassins ont également été retirés de l'échantillon au cours de la thèse en découvrant des données douteuses non déclarées comme telles dans les bases de données. La mise à l'écart de ces bassins a été choisie même lorsque les doutes ne concernaient que de courtes périodes, afin d'écarter le risque que d'autres parties des chroniques soient invalides sans que nous ne nous en soyons aperçus. Finalement, ce filtrage a permis de constituer un échantillon de 558 bassins, dont 377 français et 181 bassins australiens.

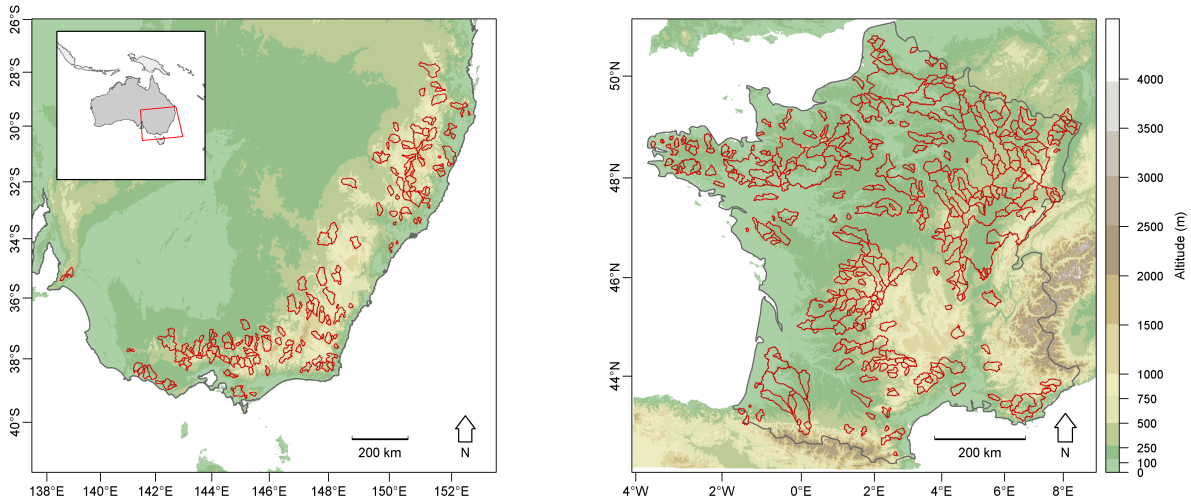


FIGURE 3.1 : Contours des 181 bassins versants australiens de l'échantillon et des 377 bassins versants français.

3.1.2. Origine des données

3.1.2.a. Données des bassins versants français

Les données des bassins versants français proviennent de la base de données construite par l'équipe Hydro d'*INRAE*, conçue pour la modélisation hydrologique journalière au moyen de modèles hydrologiques globaux (Delaigue *et al.*, 2020). Les séries chronologiques journalières à disposition s'étendent sur des périodes situées entre août 1958 et juillet 2018.
Les données de débit ont été fournies par la Banque HYDRO, qui opère un travail régulier de vérification des valeurs mesurées (accessibles à l'adresse `http://www.hydro.eaufrance.fr/`, dernier accès aux données le 21 Janvier 2019, décrite par Leleu *et al.*, 2014). Les séries de débit utilisées ont été choisies pour comprendre au moins 20 années non nécessairement consécutives présentant moins de 10% de données lacunaires.
Les données météorologiques sont issues de la réanalyse SAFRAN, fournie par Météo-France à un pas de temps journalier sur une grille régulière de 8 km de résolution (Vidal *et al.*, 2010). Dans

la base de données utilisée dans la thèse, ces valeurs ont été agrégées sur la surface des bassins versants. Les chroniques disponibles regroupent les données de précipitations totales, la fraction solide des précipitations, les températures moyenne, minimale et maximale de l'air, la vitesse du vent à 10 m, l'humidité spécifique de l'air et les rayonnements incidents hautes fréquences et basses fréquences. Seules les chroniques de précipitations totales et de température moyenne ont servi directement en forçage des modèles hydrologiques, tandis que les autres ont été utilisées pour calculer les chroniques d'évaporation potentielle (c.f. Section 3.3), elles aussi calculées maille par maille puis agrégées sur la surface des bassins.

On notera qu'en France, par convention, les années hydrologiques commencent le 1er octobre et terminent le 30 septembre de l'année calendaire suivante.

3.1.2.b. Données des bassins versants australiens

Les données des bassins versants australiens ont été réunies par le *CSIRO Land&Water* (Canberra, Australie). Ces chroniques sont également disponibles au pas de temps journalier et couvrent des périodes variables selon les bassins versants au sein de la période 1974-2006. L'échantillon dont elles proviennent a été décrit en détail par Vaze *et al.* (2011).
Les séries de débit ont été mesurées par les différentes agences gouvernementales australiennes, qui ont la charge du contrôle de leur qualité.
Les données météorologiques proviennent de la base *SILO Data Drill* (accessible à l'adresse `www.longpaddock.qld.gov.au/silo/`), qui fournit des chroniques journalières sur une grille de résolution 0,05°, soit 5 km environ, couvrant l'ensemble de l'Australie. Ces chroniques ont été calculées à partir de l'interpolation de mesures ponctuelles du *BoM*. Les données ont été agrégées par le *CSIRO Land&Water* sur les bassins versants de l'échantillon. Ne disposant pas des observations nécessaires pour les calculer nous-mêmes, nous nous sommes contentés des séries d'évaporation potentielle calculées par le *BoM*, obtenues avec les formules d'Oudin (Oudin *et al.*, 2005a) et de Morton (Morton, 1983), décrites en Section 3.3.

La *Millenium Drought* évoquée précédemment a été caractérisée dans le Sud-Est de l'Australie par une diminution importante des précipitations, accompagnée par une augmentation de la température (Chiew *et al.*, 2014). La diminution des précipitations, accumulée plusieurs années de suites a causé un déficit important de précipitations, et s'est repercutée rapidement sur les réserves d'eau de surface (Leblanc *et al.*, 2009). Une tendance décroissante marquée s'observe sur les niveaux d'eaux souterraines au moins jusqu'en 2009.

On notera qu'en Australie, par convention, les années hydrologiques commencent le 1er juillet et terminent le 30 juin de l'année calendaire suivante.[2] Il convient de noter que dans l'ensemble du manuscrit, sauf mention contraire, les années seront comptées d'après le calendrier hydrologique.

3.1.3. Caractéristiques des bassins versants de l'échantillon

Pour décrire les bassins versants de l'échantillon et étudier leur diversité, nous avons détaillé les caractéristiques selon trois catégories d'indicateurs : topographiques, climatiques et hydrologiques. Les distributions des deux sous-échantillons français et australiens ont été séparées afin d'identifier clairement leurs différences.

[2] `http://www.bom.gov.au/water/awid/id-644.shtml`

3.1.3.a. Caractéristiques topographiques

La topographie des bassins utilisés dans la thèse est décrite en Figure 3.2 au moyen de six caractéristiques :

- Surface des bassins versants topographiques ;
- Longueur hydraulique, calculée comme la distance maximale parcourue par une goutte d'eau tombée en un point d'un bassin jusqu'à l'exutoire ;
- Densité de drainage, qui est le rapport entre la longueur hydraulique et la surface d'un bassin. Plus cette valeur est faible, plus la surface de drainage contribuant à l'alimentation d'un tronçon de rivière est grande ;
- Altitude moyenne ;
- Pente moyenne ;
- Indice topographique moyen calculé avec la méthode de Beven et Kirkby (1979), qui est couramment employée pour rendre compte de l'influence de la topographie sur les processus hydrologiques.

D'après les distributions observées sur la Figure 3.2, les bassins de l'échantillon sont généralement de petite taille, dans la mesure où les exutoires ont été choisis assez proches de la source des cours d'eau pour limiter l'influence des activités humaines. Comparés aux bassins français, les bassins australiens drainent nettement moins de surface au regard de la taille des cours d'eau. Ils sont également plus hauts et un peu plus pentus que les bassins français. Les bassins français d'altitude les plus enneigés ayant été rétirés de la base de données, les bassins français sont plus bas que les bassins australiens, où il ne neige presque jamais quelle que soit leur altitude. Nous ne disposions pas d'assez d'informations pour calculer l'indice topographique des bassins australiens.

3.1.3.b. Caractéristiques climatiques

La description des caractéristiques climatiques des bassins versants a été réalisée à partir de six indices (Figure 3.3) :

- Précipitations (notées P) annuelles moyennes ;
- Évaporation potentielle (notée EP) annuelle moyenne ;
- Indice d'humidité (notée IH), calculé comme le rapport des deux quantités précédentes. Les bassins dont l'indice d'humidité est inférieur à 1 sont limités en eau, tandis que les bassins dont l'indice d'humidité est supérieur à 1 sont limités en énergie ;
- Coefficient de variation des précipitations journalières ;
- Indice d'intensité des précipitations journalières, obtenu en rapportant le quantile 99% des précipitations journalières aux précipitations journalières moyennes ;
- Taux annuel moyen de jours sans précipitations.

D'après les distributions observées sur la Figure 3.3, les précipitations moyennes subies par les bassins de l'échantillon sont échelonnées sur une gamme assez large allant de 500 à 1200 mm par an, avec une petite partie des bassins marqués par des précipitations très abondantes. Bien que les précipitations moyennes des deux sous-échantillons soient centrées autour des mêmes ordres de grandeur, des différences très marquées sont observables dans leur distribution temporelle. Le régime australien est en effet particulièrement variable et est caractérisé par des précipitations très intenses. Les taux de jours sans précipitations des échantillons sont néanmoins équivalents, ce qui suggère que les bassins australiens se démarquent surtout par une grande variabilité interannuelle des précipitations plutôt que par une saisonnalité beaucoup plus marquée que celle des bassins français. La demande évaporatoire sur les bassins français est spatialement assez homogène, et tranche nettement avec les conditions australiennes plus arides et spatialement hétérogènes. De ce fait, une grande partie des bassins français est limitée en énergie tandis qu'un peu plus de la moitié des bassins australiens est limitée en eau.

3.1.3.c. Caractéristiques hydrologiques

Six signatures hydrologiques ont servi pour la caractérisation des écoulements des bassins de l'échantillon (Figure 3.4) :

- Débit journalier moyen ;
- Coefficient de variation du débit journalier ;
- Indice de débit de base, estimant la part du débit de base dans les écoulements, en opposition au débit provenant du ruissellement (calculé avec la méthode de Pelletier et Andréassian, 2020; Pelletier *et al.*, 2020) ;
- Moyenne des débits journaliers maximaux annuels, qui renseigne sur l'intensité des crues ;
- Moyenne du débit mensuel minimal annuel (QMNA), indiquant le niveau de sévérité des étiages ;
- Rendement des précipitations, calculé comme le rapport du débit journalier moyen sur les précipitations journalières moyennes.

D'après les distributions observées sur la Figure 3.4, les écoulements moyens des bassins de l'échantillon sont répartis sur un spectre assez large, les bassins français étant en moyenne largement moins secs que les bassins australiens où les débits moyens observés sont parfois très faibles. De même que celui des précipitations, le régime des débits australiens est très nettement plus variable, notamment en raison de l'incidence majeure qu'a eue la *Millenium Drought* sur les débits des années les plus récentes. Les débits d'étiage des bassins australiens sont d'ailleurs faibles en comparaison des bassins français. La part du débit de base dans les écoulements des bassins français est répartie sur une gamme de grande envergure et témoigne de configurations hydrogéologiques très variées. Selon ce critère, le régime des bassins australiens semble en revanche plus uniforme et davantage contrôlé par le ruissellement. Les crues de la plupart des bassins étudiés sont plutôt modérées, mais peuvent être violentes pour une petite partie d'entre eux. Enfin, le rendement des précipitations est nettement plus faible dans le bassins australiens, ce qui s'explique par la demande évaporatoire plus forte précédemment observée.

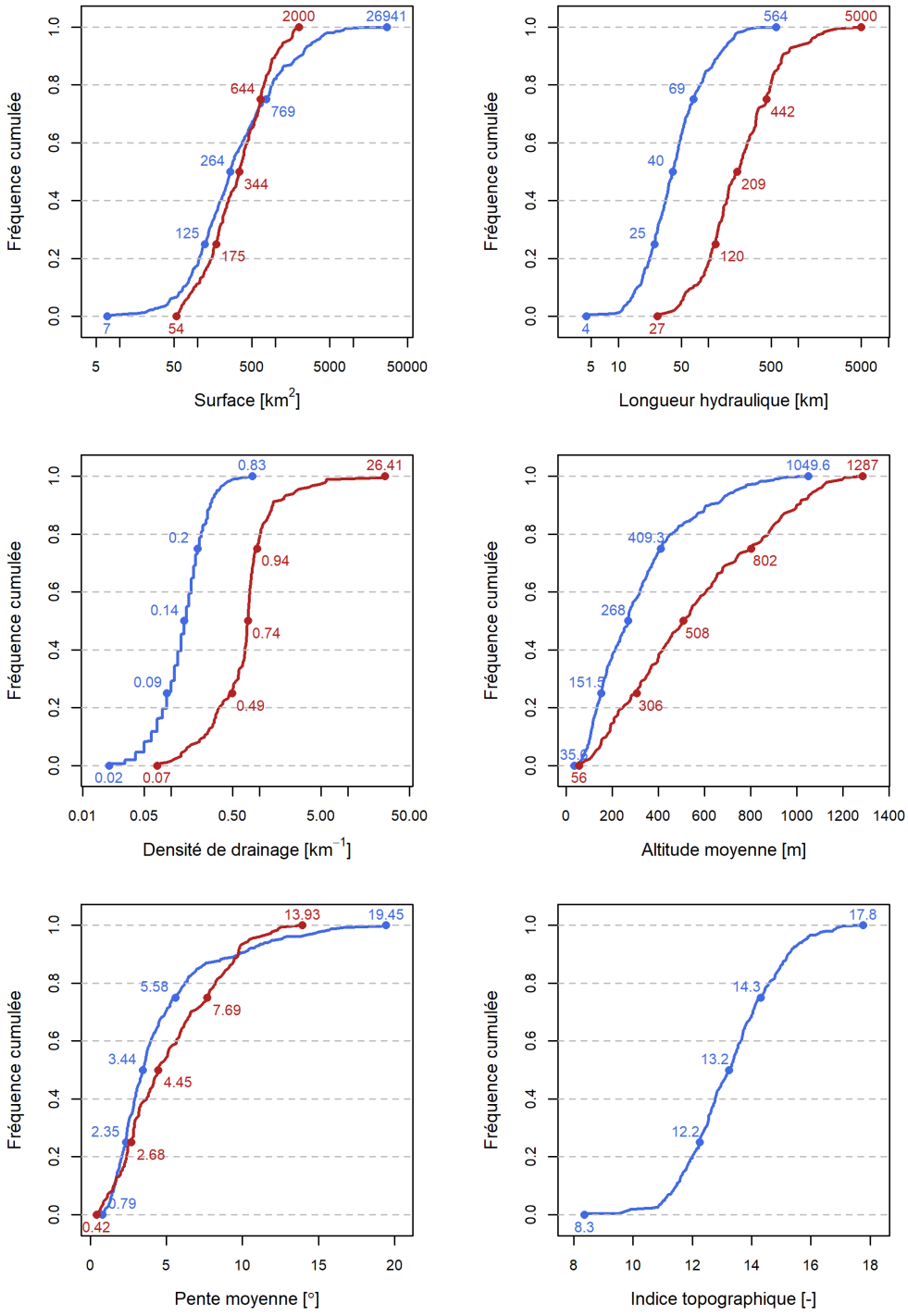


FIGURE 3.2 : Distributions des six indices topographiques calculés sur l'échantillon des 377 bassins versants français (en bleu) et 181 bassins versants australiens (en rouge foncé). Les quartiles et extrema de chaque distribution sont représentés par des points.

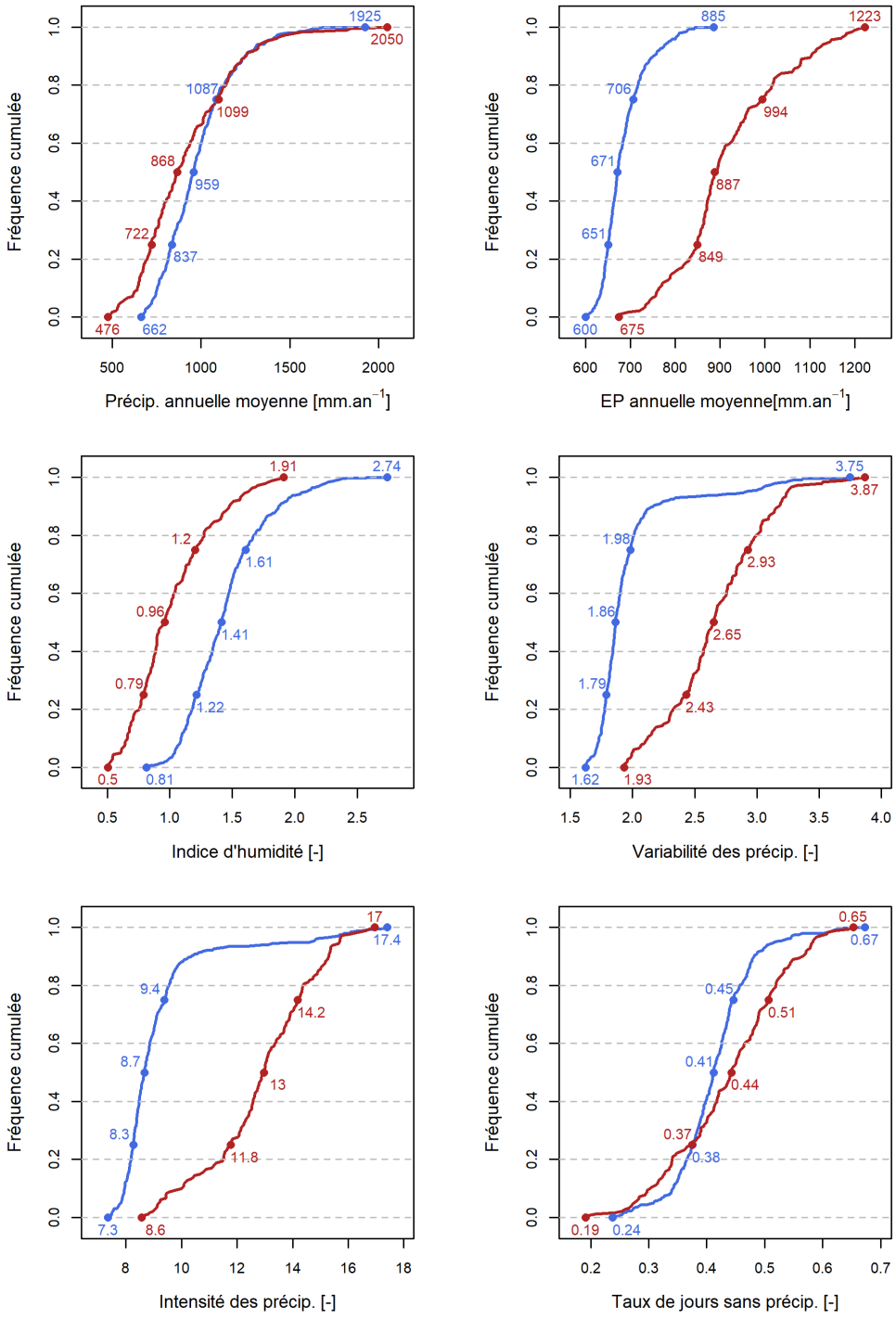


FIGURE 3.3 : Distributions des six indices climatiques calculés sur l'échantillon des 377 bassins versants français (en bleu) et 181 bassins versants australiens (en rouge foncé). Les quartiles et extrema de chaque distribution sont représentés par des points.

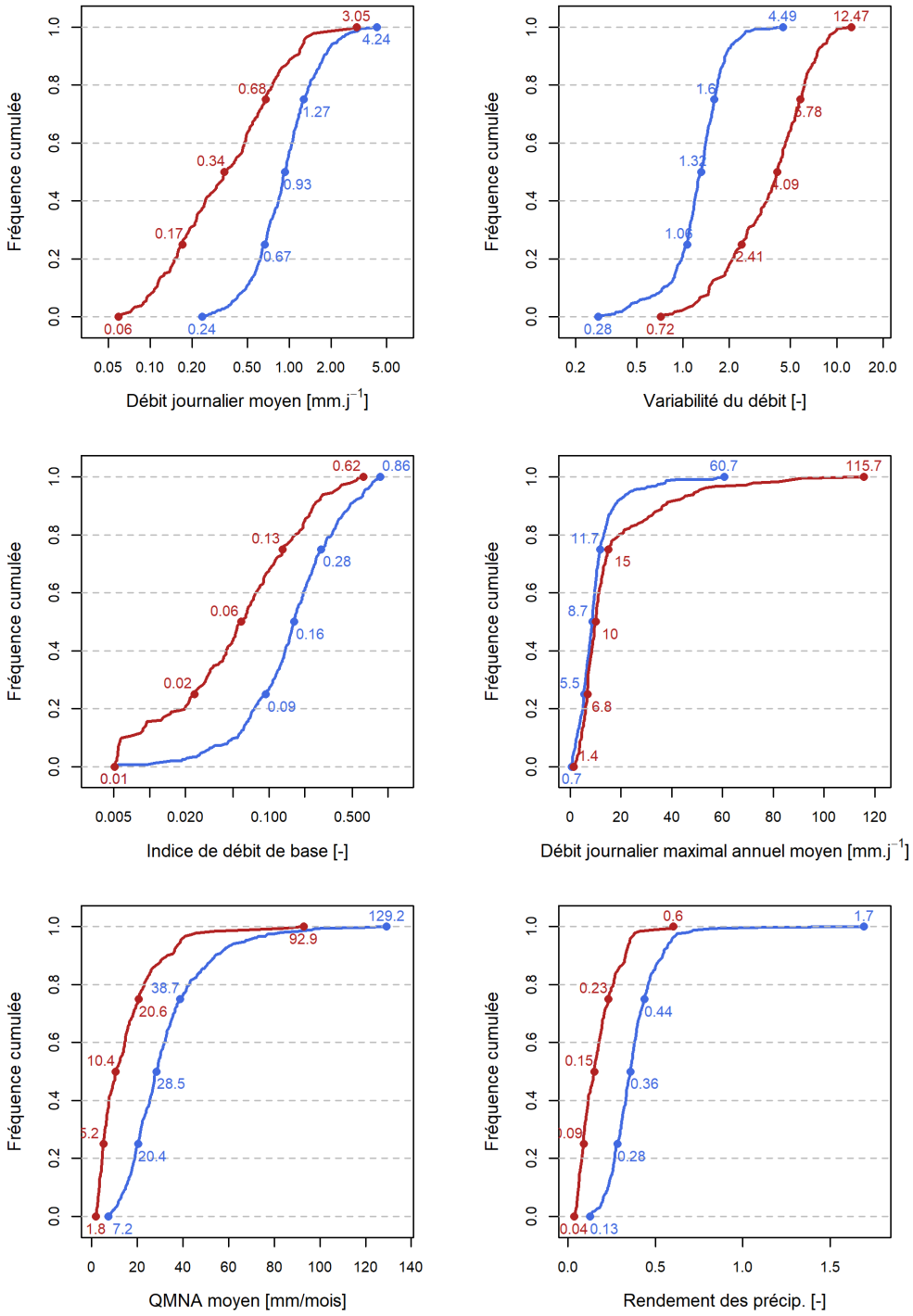


FIGURE 3.4 : Distributions des six indices hydrologiques calculés sur l'échantillon des 377 bassins versants français (en bleu) et 181 bassins versants australiens (en rouge foncé). Les quartiles et extrema de chaque distribution sont représentés par des points.

3.2. Modèles hydrologiques

L'évaluation de la robustesse des modèles hydrologiques a été réalisée à partir de trois modèles conceptuels : GR4J (Perrin *et al.*, 2003), TOPMO (Edijatno *et al.*, 1999; Michel *et al.*, 2003) et SimHyd (Chiew *et al.*, 2002a; Tan *et al.*, 2005). Ces trois modèles sont tous relativement simples (jusqu'à huit paramètres libres et quatre réservoirs) mais néanmoins de complexités différentes. Ils ont de plus été conceptualisés par des équipes de recherche de trois pays marqués par des topographies et des climats différents (France, Royaume-Uni, Australie). Les modèles ont été utilisés au moyen du package R *airGR* (Coron *et al.*, 2017, 2018) disponible sur le CRAN (`cran.r-project.org/web/packages/airGR/`), et complété par le package R *airGRplus* (Coron, 2020), qui est encore en développement, pour les modèles TOPMO et SimHyd. Peu d'autres modèles hydrologiques codés dans le package *airGRplus* étaient fonctionnels au moment de la sélection. Au début de la thèse, le modèle HBV faisait partie de la liste des modèles utilisés, mais il a finalement été mis à l'écart afin de réduire les temps de calcul.
Ces modèles ont été couplés à un module de neige simple, le modèle CemaNeige (Valéry, 2010), pour calculer la partition des précipitations solides et liquides et l'évolution du stock de neige. Le module fait également partie du package *airGR*.

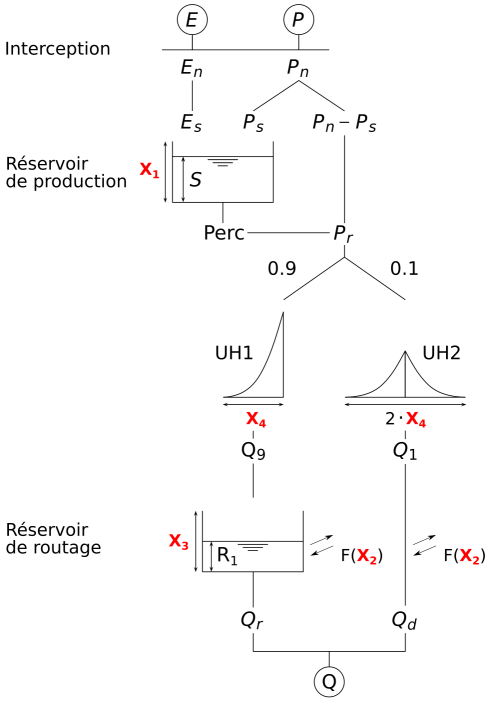
3.2.1. GR4J

GR4J est un modèle parcimonieux à quatre paramètres, développé par Perrin *et al.* (2003). Son architecture et ses équations sont les fruits d'une comparaison des performances de multiples composants de modèles, menée de sorte à dégager une structure à la fois simple et efficace. GR4J a été utilisé dans de nombreuses études de modélisation dans des contextes différents (e.g. Coron *et al.*, 2012; Fowler *et al.*, 2016; Mathevet *et al.*, 2020). La structure et les équations du modèle sont décrites sur la Figure 3.5, et ses paramètres sont répertoriés dans le Tableau 3.1.

Le bilan hydrique du modèle est contrôlé par des processus liés à l'évaporation, comprenant interception (parfois qualifée de « neutralisation ») et transpiration, et par une fonction d'échange souterrain simulant les pertes ou les gains avec les bassins versants voisins. La fonction de neutralisation permet de prendre en compte le phénomène d'interception sans avoir recours à un réservoir spécifique. L'évaporation réelle et les précipitations nettes sont calculées en fonction du remplissage du réservoir de production. Les précipitations nettes sont divisées en deux parties, la plus faible étant orientée dans un hydrogramme unitaire symétrique à partir duquel elle s'écoule directement comme composante directe du débit. L'autre partie passe par un hydrogramme unitaire asymétrique, dont la sortie alimente un réservoir de routage. Le remplissage de ce réservoir détermine la quantité d'eau échangée avec l'extérieur (gain ou perte selon le signe du paramètre X_2) ainsi que la quantité d'eau vidangée, qui représente la composante indirecte du débit simulé.

TABLEAU 3.1 : Paramètres du modèle GR4J.

Nom	Description	Unité	Min	Max
X_1	capacité maximale du réservoir de production	mm	0	2×10^4
X_2	coefficient d'échanges souterrains	mm j^{-1}	-1×10^4	1×10^4
X_3	capacité journalière du réservoir de routage	mm	0	2×10^4
X_4	temps de base de l'hydrogramme unitaire	j	0,5	20



$$Si\ P \geq E,\ P_n = P - E,\ E_n = 0$$
$$Si\ P < E,\ E_n = E - P,\ P_n = 0$$

$$P_s = \frac{P_n\left(1 - \left(\frac{S}{X_1}\right)^2\right)}{1 + \frac{P_n}{X_1}\left(1 + \frac{S}{X_1}\right)} \qquad E_s = \frac{E_n \frac{S}{X_1}\left(2 - \frac{S}{X_1}\right)}{1 + \frac{E_n}{X_1}\left(2 - \frac{S}{X_1}\right)}$$

$$Perc = S - \left(S^{-4} + \left(\frac{9}{4}X_1\right)^{-4}\right)^{-\frac{1}{4}} \qquad S = S - Perc$$

$$F = X_2\left(\frac{R}{X_3}\right)^{\frac{7}{2}} \qquad R = \max\left(0,\ R + Q_9 + F\right)$$

$$Q_r = R - \left(R^{-4} + X_3^{-4}\right)^{-\frac{1}{4}} \qquad R = R - Q_r$$
$$Q_d = \max\left(0,\ Q_1 + F\right)$$

FIGURE 3.5 : Schéma et équations du modèle GR4J.

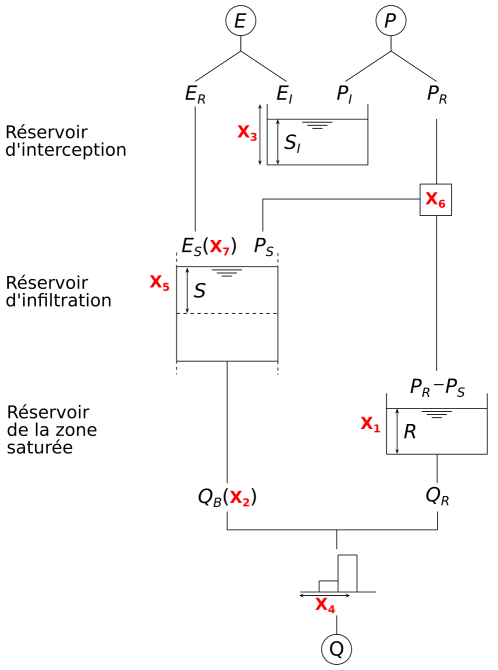
3.2.2. TOPMO

Le modèle TOPMO est une adaptation du modèle distribué TOPMODEL de Beven et Kirkby (1979), dont la distribution des indices topographiques a été remplacée par une équation paramétrée afin de pallier l'absence des données topographiques requises pour utiliser le modèle original (Edijatno *et al.*, 1999; Michel *et al.*, 2003). Ne possédant pas ces informations pour le sous-échantillon de bassins australiens, il était naturel de préférer TOPMO à TOPMODEL. Le modèle possède sept paramètres libres, décrits dans le Tableau 3.2, et a été utilisé dans des études impliquant des bassins de différentes régions du monde (e.g. Oudin *et al.*, 2006a; Brigode *et al.*, 2013). Sa structure est détaillée sur la Figure 3.6.

Le bilan hydrique du modèle est simplement contrôlé par les processus liés à l'évaporation. Le réservoir d'interception peut retenir une partie des précipitations non évaporées et la rendre disponible à l'évaporation à taux potentiel des jours suivants. Une partie des précipitations non interceptées rechargent le réservoir d'infiltration selon son remplissage, la partie restante alimentant le réservoir de la zone saturée. Le réservoir d'infiltration est également soumis à l'évaporation non interceptée selon une loi exponentielle dépendant de son remplissage. Ce réservoir contribue au débit simulé en générant un écoulement par excès de saturation. Le réservoir de la zone saturée alimente le débit de base simulé. L'écoulement de la somme de ces deux composantes se produit enfin avec un décalage temporel simple.

TABLEAU 3.2 : Paramètres du modèle TOPMO.

Nom	Description	Unité	Min	Max
X_1	capacité du réservoir de la zone saturée	mm	0	2×10^4
X_2	paramètre d'écoulement par excès de la zone saturée	mm j^{-1}	0	2×10^4
X_3	capacité du réservoir d'interception	mm	0	2×10^4
X_4	décalage temporel	j	0	2×10^4
X_5	capacité du réservoir d'infiltration	mm	1	20
X_6	paramètre du taux de recharge	–	0	2×10^4
X_7	paramètre du taux d'évaporation	–	-1×10^4	1×10^4



$$Si\ E > T + P,\ P_R = 0,\ E_R = E - (T + P)\,,\ E_I = T + P$$
$$Si\ P > X_3 + E - T,\ P_R = P + T - (E + X_3)\,,\ E_R = 0,\ E_I = E$$
$$Si\ E - P \leq T \geq X_3 + E - T,\ P_R = 0,\ E_R = 0,\ E_I = E$$

$$P_S = \frac{P_R}{1 + \exp\left(\frac{S}{X_5} - X_6\right)} \qquad S = S + P_S$$
$$E_S = \frac{E_R}{1 + \exp\left(X_7 - \frac{S}{X_5}\right)} \qquad S = S - E_S$$
$$Q_B = X_2 \log\left(1 + \exp\left(\frac{S}{X_2}\right)\right) \qquad S = S - Q_B$$

$$R = R + P_R - P_S \qquad Q_R = \frac{R^2}{R + X_1}$$

FIGURE 3.6 : Schéma et équations du modèle TOPMO.

3.2.3. SimHyd

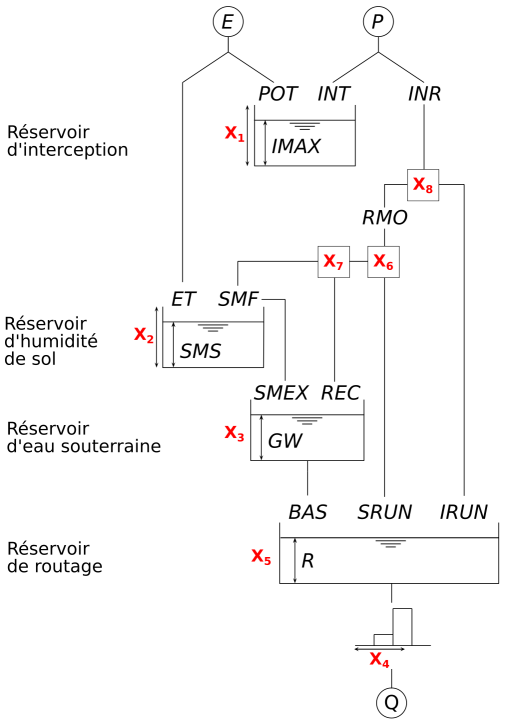
Le modèle SimHyd a été développé par Chiew *et al.* (2002a) et simplifié par Tan *et al.* (2005). Il a été souvent employé pour la simulation du débit de bassins versants australiens (e.g. Chiew *et al.*, 2009; Coron *et al.*, 2012). Il possède sept paramètres dans la version originale. Dans la version du package *airGRplus* développée par Coron (2020), un de ces paramètres a été fixé à sa valeur médiane ($SQ = 2$) pour faciliter le calage du modèle, et deux paramètres ont été ajoutés par l'addition d'une fonction de routage au modèle (voir Tableau 3.3). Le modèle utilisé dans cette thèse possède donc huit paramètres. La structure du modèle est exposée sur la Figure 3.7.

Le bilan hydrique du modèle est simplement contrôlé par les processus liés à l'évaporation. Le réservoir d'interception se comporte comme une fonction de neutralisation à ceci près que la quantité d'eau interceptée possède une limite supérieure. Un partie des précipitations non interceptées s'écoule directement vers le routage, sans s'infiltrer. Le reste est divisé entre les infiltrations vers le

TABLEAU 3.3 : Paramètres du modèle SimHyd. Les noms des paramètres entre parenthèses correspondent au nom des paramètres du modèle original.

Nom	Description	Unité	Min	Max
X_1 $(INSC)$	capacité du réservoir d'interception	mm	0	$2{\times}10^4$
X_2 $(SMSC)$	capacité du réservoir d'humidité du sol	mm	0	$2{\times}10^4$
X_3 (K)	paramètre de récession linéaire du débit de base	mm	1	$2{\times}10^4$
X_4 (TD)	décalage temporel	j	1	20
X_5 (RC)	capacité du réservoir de routage	mm	1	$2{\times}10^4$
X_6 (SUB)	coefficient de proportionnalité de l'écoulement hypodermique	–	1	$2{\times}10^4$
X_7 $(CRAK)$	coefficient de proportionnalité de la recharge souterraine	–	1	$2{\times}10^4$
X_8 $(COEFF)$	pertes maximales par infiltration	mm	0	$2{\times}10^4$

réservoir d'humidité du sol et la recharge du réservoir souterrain. Après avoir été alimenté par les infiltrations et attaqué par un reliquat de demande évaporatoire, l'excédent du réservoir d'humidité est dirigé vers le réservoir souterrain. Le débit de base généré par le réservoir souterrain s'écoule dans le réservoir de routage. Le débit simulé par le modèle correspond à l'écoulement linéaire généré par le réservoir de routage avec un décalage temporel simple.



$$IMAX = \min(X_1,\ E) \qquad INT = \min(IMAX,\ P)$$
$$INR = P - INT \qquad POT = E - INT$$

$$RMO = \min\left(X_8 e^{-2\frac{SMS}{X_2}},\ INR\right) \qquad IRUN = INR - RMO$$

$$SRUN = \frac{1}{X_6}\frac{SMS}{X_2}RMO$$
$$REC = \frac{1}{X_7}\frac{SMS}{X_2}(RMO - SRUN)$$
$$SMF = RMO - SRUN - REC$$
$$ET = \min\left(10\frac{SMS}{X_2},\ POT\right)$$
$$SMEX = \max(SMS - X_2,\ 0)$$

$$GW = GW + REC + SMEX$$
$$BAS = \frac{1}{X_5}\frac{GW}{X_3} \qquad GW = GW - BAS$$

$$R = R + IRUN + SRUN + BAS$$
$$Q_R = \frac{R}{X_5} \qquad R = R - Q_R$$

FIGURE 3.7 : Schéma et équations du modèle SimHyd.

3.2.4. CemaNeige

Le module de neige *CemaNeige* est un modèle degré-jour à réservoir conçu par Valéry *et al.* (2014) pour représenter l'évolution du volume du manteau neigeux en amont de la modélisation hydrologique. Le modèle calcule à chaque pas de temps la fraction solide des précipitations alimentant le manteau neigeux, ainsi que la lame d'eau de fonte qui est ajoutée aux précipitations liquides en entrée des modèles hydrologiques. Le module est parcimonieux et possède 2 paramètres. Sa

structure et ses équations sont décrites sur la Figure 3.8.

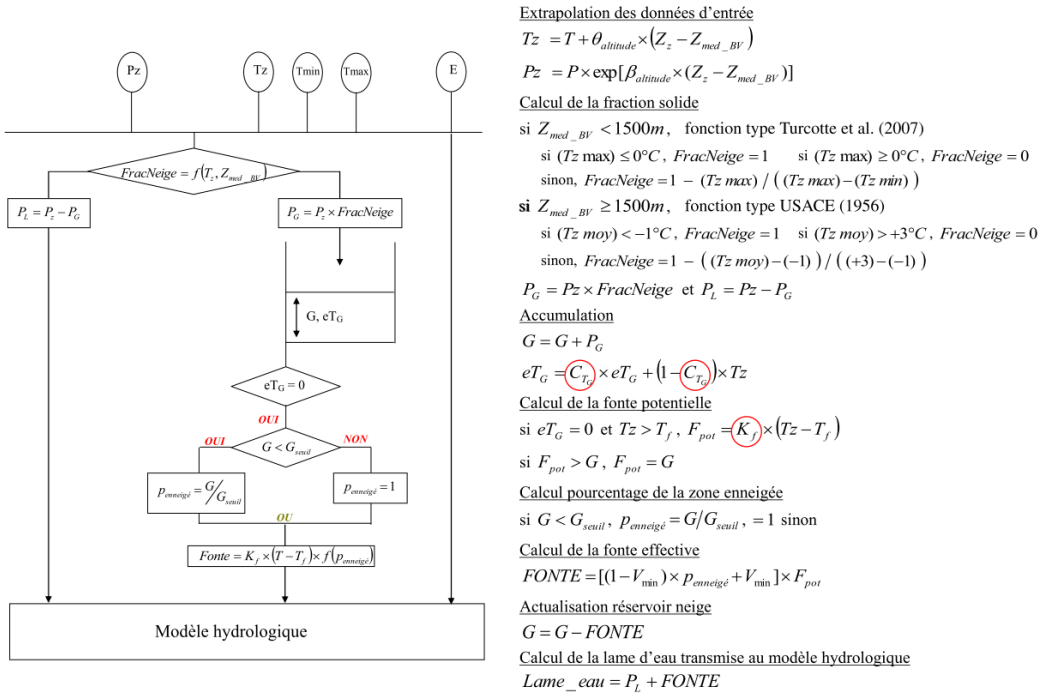


FIGURE 3.8 : Schéma et équations du modèle *CemaNeige*. Tiré de Valéry (2010).

CemaNeige peut être implémenté en parallèle sur plusieurs couches d'altitude de surfaces égales afin de représenter la distribution altitudinale des précipitations solides et donc de modéliser des évolutions différenciées du manteau neigeux. Néanmoins, nous avons choisi de ne l'utiliser ici que de manière globale puisque les bassins de l'étude sont peu influencés par la neige et ne nécessitent a priori pas ce niveau de détail. Par ailleurs, les deux paramètres libres du module ont été fixés aux valeurs médianes identifiées par Valéry (2010) afin d'éviter qu'ils soient sous-contraints en calage sur des périodes peu enneigées. La transférabilité climatique des paramètres d'un module neigeux est un sujet d'étude important pour la qualité des projections hydrologiques dans les bassins nivaux, mais ne rentre pas dans les objectifs fixés pour cette thèse.

3.3. Formules d'évaporation potentielle

Les formulations d'évaporation potentielle évaluées dans cet article sont celles d'Oudin (Oudin *et al.*, 2005a), de Penman (Penman, 1948), de Penman-Monteith (Monteith, 1965) et de Morton (Morton, 1983). Les formules d'Oudin et de Penman-Monteith sont très couramment utilisées en modélisation hydrologique en France (e.g. Perrin *et al.*, 2008; Nicolle *et al.*, 2014), tandis que celle de Morton est populaire en Australie (e.g. Vaze *et al.*, 2010; Fowler *et al.*, 2016). La formule de Penman correspond à un modèle antérieur à celui proposé par Monteith (1965), rarement utilisé mais néanmoins intéressant dans le cadre de la modélisation hydrologique. Mis à part la grande similarité entre les formules de Penman et de Penman-Monteith, ces formules correspondent à des conceptions diverses de l'évaporation potentielle et s'échelonnent sur plusieurs niveaux de complexité. La formule d'Oudin, la plus simple, est une formule empirique fondée uniquement sur la température. La formule de Morton dérive d'une approche radiative, liée à la représentation du bilan d'énergie. Les formules de Penman et de Penman-Monteith sont les plus complexes et sont

issues d'une représentation combinant une approche radiative et une approche aérodynamique.
Les données nécessaires au calcul des formules de Penman et de Penman-Monteith n'étant pas disponibles pour les bassins australiens de la thèse, seules les formules d'Oudin et de Morton y ont été utilisées.
Nous présentons dans la suite les différents modèles d'évaporation potentielle de façon succincte. Les théories et conceptions sur lesquelles ces modèles se fondent sont suffisamment complexes pour que nous ne soyons pas en mesure de les détailler, et constituent des sujets de recherche à part entière.

3.3.1. Oudin

La formule d'Oudin résulte d'une conception empirique et volontairement simple de l'évaporation potentielle, et est fondé sur le constat formulé par Oudin *et al.* (2005a) à partir d'un échantillon de 27 formulations différentes que l'utilisation des formules complexes n'améliore pas les performances des modèles hydrologiques. La formule d'Oudin est basée uniquement sur les observations de température journalière, dans la mesure où le rayonnement est calculé à partir d'une formule théorique définissant un cycle annuel répété chaque année. La lame d'évaporation potentielle exprimée en $\mathrm{m\,s^{-1}}$ est estimée par la formule suivante :

$$E_{Ou} = \frac{R_e}{\lambda\rho}\frac{T+5}{100} \tag{3.1}$$

La conversion en $\mathrm{mm\,j^{-1}}$ est trivialement réalisée en factorisant par $8{,}64\times10^{7}$. La température journalière moyenne de l'air est matérialisée par le symbole T et est mesurée en °C. Les constantes λ et ρ représentent respectivement l'enthalpie de vaporisation de l'eau (en $\mathrm{J\,kg^{-1}}$) et la masse volumique de l'eau (en $\mathrm{kg\,m^{-3}}$). Enfin, R_e désigne le rayonnement solaire extraterrestre incident au bassin versant avant traversée de l'atmosphère. Il est calculé au moyen de la formule suivante :

$$R_e = \frac{1400}{\eta^2}\frac{\omega}{\pi}\cos z \tag{3.2}$$

Où $1400\,\mathrm{W\,m^{-2}}$ correspond approximativement au rayonnement reçu par la plus haute couche de l'atmosphère perpendiculaire aux rayons du Soleil. L'angle ω est l'angle en radians parcouru par la Terre entre le lever du soleil et son point culminant, tel que rapporté à π il dénote la fraction du jour pendant laquelle le Soleil est visible. Le facteur $\cos z$ corrige le calcul en fonction de la distance angulaire z en radians entre le zénith et le point culminant atteint par le Soleil. Enfin η correspond au rayon vecteur du Soleil (sans unité) et module l'ensoleillement par rapport à la distance entre le Soleil et la Terre, qui varie au cours de l'année. Les variables ω, $\cos z$ et η sont déterminées en fonction du jour julien par des calculs présentés dans l'annexe C des travaux de Morton (1983). Le rayonnement R_e ne dépend donc que de la date du calendrier et suit un cycle identique chaque année.

3.3.2. Penman-Monteith

La formule de Penman provient d'une des théories pionnières de l'évaporation potentielle. Elle combine un terme aérodynamique lié à l'humidité observée et un terme énergétique lié au rayonnement reçu par la surface :

$$E_{Pe} = \frac{\Delta}{\Delta + \gamma} \frac{R_n}{\lambda \rho} + \left(1 - \frac{\Delta}{\Delta + \gamma}\right) \frac{1}{\lambda \rho} \frac{\rho_a c_p}{r_a} (\nu_s - \nu) \tag{3.3}$$

Où ν, ν_s et Δ représentent respectivement la pression de vapeur partielle de l'air (en kPa), la pression de vapeur saturante à la température T (en kPa) et sa pente dans le diagramme de Clausius-Clapeyron (en kPa K^{-1}). La constante psychrométrique γ (en kPa K^{-1}) dépend linéairement de la pression atmosphérique, calculée à partir de l'altitude moyenne du bassin. La quantité R_n représente le rayonnement net du sol (en W m^{-2}), c'est-à-dire le bilan entre l'énergie reçue et l'énergie perdue par rayonnement. Enfin ρ_a et c_p sont la masse volumique (en kg m^{-3}) et la capacité thermique spécifique de l'air (en J K^{-1} kg^{-1}) et r_a matérialise une résistance aérodynamique (en s m^{-1}).

La formule de Penman représente l'évaporation comme la résultante d'un arbitrage entre un terme radiatif et un terme convectif. Il convient de noter qu'à l'équilibre, à apport énergétique constant et sans advection d'air en dehors du système, ces deux termes convergent l'un vers l'autre. En effet, si le terme radiatif est supérieur au terme convectif, alors l'évaporation est limitée par le transfert de vapeur entre le sol et l'air ambiant, ce qui signifie que le sol n'évacue pas suffisamment la chaleur par les changements d'état et donc s'échauffe. Son échauffement a pour double effet d'augmenter son rayonnement de corps noir, et donc de diminuer le rayonnement net, ainsi que d'augmenter la pression de vapeur saturante de l'air ambiant par transfert de chaleur sensible et donc d'augmenter le terme convectif. Si le terme convectif est supérieur au terme radiatif, alors l'évaporation est limitée en énergie. Le transfert de vapeur entre le sol et l'air a dans ce cas pour effet de refroidir le sol lors de la vaporisation de l'eau. Ce refroidissement augmente le terme radiatif et diminue le terme convectif pour les mêmes raisons que son réchauffement provoquait la diminution du terme radiatif et l'augmentation du terme convectif.

À partir des approximations proposées par Allen *et al.* (1998), le calcul de l'évaporation potentielle de Penman nécessite la mesure de quatre variables météorologiques : la température moyenne de l'air, l'humidité spécifique moyenne de l'air, les rayonnements incidents reçus par la surface et la vitesse du vent à 2 m au-dessus du sol. Si les deux premières sont généralement mesurées telles quelles, les deux autres sont souvent dérivées de modèles à partir d'autres observations, comme la durée d'ensoleillement quotidienne ou la vitesse du vent à 10 m. Pour le calcul du rayonnement net R_n, les données dont nous disposons par Météo-France incluent directement le rayonnement incident de hautes fréquences R_S et le rayonnement de basses fréquences R_a et permettent d'estimer le rayonnement net comme suit :

$$R_n = (1 - a) R_S + R_a - \epsilon \sigma T^4 \tag{3.4}$$

Le terme en T^4 est lié à la perte d'énergie du sol par rayonnement vers l'atmosphère, avec σ la constante de Boltzmann (en J K^{-1}), et ϵ l'émissivité de la surface qui est fixée à 0,95 (sans unité) en considérant que le sol est constitué d'herbe verte. La valeur de l'albédo utilisée dans ces deux formules, prise par défaut égale à 0,23 (herbe verte), a été corrigée au moyen du modèle degré-jour d'accumulation et de fonte de la neige CemaNeige, en suivant une loi de décroissance exponentielle de l'albédo vers 0,23 après les dernières chutes de neige.
La résistance aérodynamique est approximée par Allen *et al.* (1998) par une simple relation inverse :

$$r_a = \frac{208}{u_2} \qquad \text{avec} \quad u_2 = u_z \frac{4.87}{\ln(67.8\ z - 5.42)} \tag{3.5}$$

Où u_2 et u_z sont les vitesses du vent à 2 m et à $z = 10$ m.
Enfin, la pression de vapeur saturante et la pente de la courbe sont calculées avec l'approximation d'Auguste-Roche-Magnus :

$$\nu_s = 0.6108 \exp\left(\frac{17.27\,T}{T + 237.3}\right) \tag{3.6}$$

$$\Delta = \frac{4098.2\,\nu_s}{(T + 237.3)^2} \tag{3.7}$$

Le modèle de Penman est rarement utilisé tel quel, mais plutôt corrigé en suivant les recommandations de Monteith (1965). Ce dernier suggère d'introduire un terme de résistance des plantes à l'évaporation, lié à l'ouverture de leurs stomates, selon leur développement ou les conditions d'aridité. La formule qui en résulte, communément appelée formule de Penman-Monteith, est construite comme suit :

$$E_{PM} = \frac{\Delta + \gamma}{\Delta + \gamma\left(1 + \frac{r_s}{r_a}\right)} E_{Pe} \tag{3.8}$$

La résistante r_s est censée varier au cours de l'année. Néanmoins, Allen *et al.* (1998) en propose un paramétrage fixe pour une herbe verte bien irriguée tel que $r_s = 70$ s m^{-1}. L'hypothèse selon laquelle la végétation est bien irriguée rend cette définition de l'évaporation conforme à celle de l'évaporation potentielle pour la modélisation hydrologique. La série d'hypothèses sur le type de végétation recouvrant la surface (l'herbe verte) permet en théorie d'utiliser la formule dans des climats tempérés sans que ces hypothèses soient trop irréalistes. En pratique, il faudrait adapter les caractéristiques de la surface évaporante selon le couvert végétal majoritaire des bassins versants étudiés, voire intégrer les différences spatiales au sein du bassin. Toutefois, en considérant les constats mentionnés précédemment sur la faible sensibilité des modèles hydrologiques globaux aux formules d'évaporation potentielle, il n'a pas semblé pertinent de calculer l'évaporation avec un tel niveau de détail.

Nous avons fait le choix d'utiliser les modèles de Penman et de Penman-Monteith plutôt que seulement le modèle de Penman-Monteith, afin de vérifier l'influence d'une résistance stomatique non nulle sur l'évaporation et le débit simulés. Ce choix s'explique en outre par l'hypothèse que les fonctions d'interception ou de neutralisation des modèles hydrologiques utilisés représentent l'évaporation de l'eau liquide présente en gouttes ou en flaques à la surface du sol, qui n'est donc pas soumise à la résistance des plantes mais uniquement à la résistance aérodynamique.

3.3.3. Morton

Morton a proposé une série de formules pour calculer l'évaporation dans ses travaux. La formule retenue ici correspond à la formule dite « d'évaporation réelle en milieu humide » (Morton, 1983), qui est plus adaptée pour être utilisée comme donnée d'entrée d'un modèle hydrologique que la formule d'évaporation potentielle qu'il propose (Chiew et McMahon, 1991).

La théorie développée par Morton (1983) repose sur l'hypothèse de Bouchet (1963), qui suggère que l'évaporation potentielle est affectée par la quantité d'eau disponible à l'évaporation, de sorte qu'elle diminue d'autant que l'évaporation réelle augmente lorsque les conditions d'humidité du sol se rapprochent de la saturation en eau. Schématiquement, on peut considérer que les seuls lieux où l'évaporation potentielle n'est pas affectée par la disponibilité en eau sont les oasis balayées par des vents secs, où l'air chargé en humidité est sans cesse chassé. Au contraire, dans les situations où l'eau disponible à l'évaporation couvre une surface suffisamment grande, par exemple dans un bassin versant de plusieurs dizaines de kilomètres carrés, les conditions de l'air balayant une parcelle sont dépendantes de l'évaporation des surfaces situées en amont du flux d'air. Dans ces

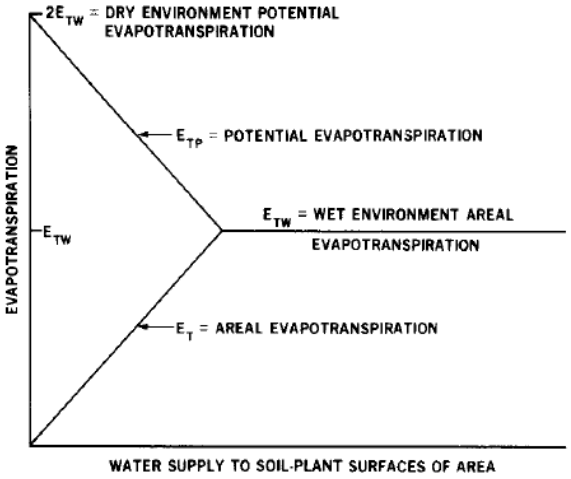


FIGURE 3.9 : Représentation schématique de l'hypothèse de Bouchet (1963) reliant l'évaporation réelle et l'évaporation potentielle à apport énergétique constant (tiré de Morton, 1983).

cas-là, la demande évaporatoire varie en fonction de la disponibilité en eau. L'hypothèse de Bouchet (1963), basée sur des considérations physiques discutées par Morton (1983), consiste à poser que la perte de demande évaporatoire correspond exactement à l'augmentation de l'évaporation réelle (voir Figure 3.9). Morton (1983) propose de calculer l'évaporation obtenue lorsque les conditions sont suffisamment humides pour que l'évaporation réelle et l'évaporation potentielle convergent vers une même valeur en adaptant le modèle de Priestley et Taylor (1972) :

$$E_{Mo} = b_1 + b_2 \frac{\Delta_p}{\Delta_p + \gamma} R_{T_p} \tag{3.9}$$

$$\frac{1}{2} E_{Mo,pot} \leq E_{Mo} \leq E_{Mo,pot} \tag{3.10}$$

$$E_{Mo,pot} = f_T \left(\nu_{s,p} - \nu \right) \tag{3.11}$$

Les constantes $b_1 = 14$ W m^{-2} et $b_2 = 1.2$ (sans unité) ont été calculées par Morton (1983) à partir de chroniques de mesures dans des lieux désertiques où l'évaporation « potentielle » $E_{Mo,pot}$ vaut le double de l'évaporation « réelle en milieu humide » E_{Mo}. L'évaporation E_{Mo} obtenue est bornée pour respecter l'hypothèse de Bouchet, et le terme $E_{Mo,pot}$ correspond à l'évaporation potentielle qui serait mesurée à l'équilibre à apport énergétique constant et sans advection, c'est-à-dire à la température d'équilibre T_p telle que les termes radiatif et convectif de la formule de Penman seraient égaux. Le terme E_{Mo} correspond à la définition de l'évaporation potentielle que nous avons retenue ici pour la modélisation hydrologique, puisqu'il correspond à l'évaporation réelle maximale observable sur une surface saturée en eau.
Les termes R_{T_p} (W), $\nu_{s,p}$ (Pa) et Δ_p (K kPa^{-1}) indiquent le rayonnement net, la pression de vapeur saturante et la pente de la courbe de pression de vapeur saturante à la température d'équilibre T_p. Le terme f_T correspond à un coefficient de transfert de vapeur calculé de la façon suivante :

$$f_T = \sqrt{\frac{p_s}{p}} \frac{f_z}{\zeta} \tag{3.12}$$

$$\frac{1}{\zeta} = 0.28 \left(1 + \frac{\nu}{\nu_s} \right) + \frac{\Delta R_n}{\gamma \sqrt{\frac{p_s}{p}} f_z \left(\nu - \nu_s \right)} \tag{3.13}$$

$$\frac{1}{\zeta} \leq 1 \tag{3.14}$$

Où p_s représente la pression atmosphérique au niveau de la mer et p la pression atmosphérique du bassin versant (calculée d'après son altitude moyenne). La constante f_z est fixée par Morton (1983) à 280 W m^{-2} kPa^{-1}. On notera qu'au contraire de la formule de Penman, le transfert de vapeur n'est pas relié à la vitesse du vent. Morton (1983) propose de considérer que le coefficient de transfert de vapeur est indépendant de la vitesse du vent pour plusieurs raisons : (i) le coefficient de transfert de vapeur augmente à la fois avec la rugosité de surface et la vitesse du vent, or la vitesse du vent diminue avec la rugosité de surface, ce qui tend à annuler ces effets, (ii) le coefficient de transfert de vapeur augmente lorsque l'atmosphère est plus instable, d'autant plus que le vent est faible, ce qui amène encore une fois à une compensation des effets, (iii) enfin les mesures de vitesse du vent sont généralement très incertaines du fait de leur grande sensibilité à la position des capteurs.
L'estimation de la température d'équilibre est réalisée par un processus itératif proposé par Morton (1983) à partir d'un développement limité pour résoudre l'égalité entre les termes radiatif et convectif de la formule de Penman :

$$\delta T_p = \frac{\frac{R_n}{f_T} + \nu - \nu'_{s,p} + \kappa \left(T - T'_p\right)}{\Delta'_p + \kappa} \tag{3.15}$$

$$\kappa = \gamma + 4\epsilon\sigma \left(T + 273\right)^3 \frac{1}{f_T} \tag{3.16}$$

En partant de $T'_p = T$ et en recalculant à chaque pas les termes $\nu'_{s,p}$ et Δ'_p, le processus converge en moins de quatre itérations.

Il convient de noter que certaines notations choisies par Morton (1983) ont été changées pour faciliter la compréhension et harmoniser avec les notations précédentes. Par ailleurs, nous avons utilisé les valeurs des paramètres b_1, b_2 et f_z calculées par Morton (1983) pour les bassins français, mais pour les bassins australiens les chroniques d'évaporation à notre disposition ont été produites avec des valeurs différentes par les fournisseurs de données. Ces paramètres ont en effet fait l'objet d'un recalage par Chiew *et al.* (2002b) pour mieux correspondre aux conditions rencontrées en Australie ($b_1 = 13.4$ W m^{-2}, $b_2 = 1.13$ et $f_z = 292$ W m^{-2} kPa^{-1}). Nous avons tenté de procéder pendant la thèse à un recalage spécifique pour les bassins français, mais ce travail n'a pas été jugé concluant. Enfin, Morton (1983) propose un calcul de l'albédo journalier à partir des mesures de rayonnements lumineux basses et hautes fréquences incidents, que nous avons choisi de ne pas intégrer au calcul de l'évaporation potentielle. L'albédo et les autres caractéristiques de surface pour le calcul de la formule de Morton ont été calculés ou fixés de la même manière que pour les formules de Penman et Penman-Monteith. Dans toute la suite de la thèse, l'évaporation potentielle de Morton fait référence à la formule d'évaporation « réelle en milieu humide » de Morton E_{Mo}, et non à la formule d'évaporation « potentielle » $E_{pot,Mo}$.

3.3.4. Comparaison des chroniques d'évaporation potentielle

Les variables nécessaires au calcul de chaque modèle d'évaporation potentielle ont été récapitulées dans le Tableau 3.4. Les modèles de Penman et Penman-Monteith nécessitent le plus de données et sont donc les plus complexes en ce sens, bien que la théorie de Morton soit en réalité la plus difficile à comprendre et à appliquer. Le fait que cette dernière théorie soit populaire en Australie tient au fait qu'elle ne requiert pas de données de vitesse du vent, qui n'étaient pas disponibles jusqu'aux travaux de réanalyse de McVicar *et al.* (2008). Ne disposant pas de ces données ni de mesures de rayonnement ou de durée d'ensoleillement pour les bassins australiens, nous n'avons pas pu y calculer les évaporations de Penman et Penman-Monteith.

Dans la suite, nous comparons brièvement les caractéristiques des évaporations potentielles obtenues au moyen des différents modèles, en séparant les sous-échantillons de bassins français et australiens.

TABLEAU 3.4 : Variables nécessaires au calcul des quatre formules d'évaporation potentielle utilisées dans la thèse. T correspond à la température, H à l'humidité spécifique, R au rayonnement et U à la vitesse du vent.

Nom	T	H	R	U
Oudin (Ou)	X			
Morton (Mo)	X	X	X	
Penman (Pe)	X	X	X	X
Penman-Monteith (PM)	X	X	X	X

3.3.4.a. Comparaison sur les bassins français

Les caractéristiques des évaporations potentielles calculées avec les différents modèles sur l'échantillon français sont représentées sur la Figure 3.10. Des différences entre les moyennes mensuelles sont particulièrement visibles pendant les mois d'été. Les évaporations les plus faibles sont produites par le modèle de Penman-Monteith, suivi par celui d'Oudin, puis de Penman. La formulation de Morton estime la demande évaporatoire la plus forte en été. Les différences s'atténuent en hiver et certaines relations d'ordre s'inversent. La formule d'Oudin donne les évaporations les plus faibles durant la saison froide. La distribution des évaporations potentielles moyennes des bassins montre que les formules d'évaporation se classent plutôt par paires : Oudin et Penman-Monteith estiment les demandes évaporatoires les plus faibles, tandis que Morton et Penman estiment les demandes évaporatoires les plus hautes.

La taille des boîtes à moustache renseigne sur les disparités spatiales entre bassins versants, ainsi que l'amplitude du graphe de répartition en fréquence cumulée. De ce point de vue, la formule d'Oudin aboutit à une distribution spatiale assez homogène de l'évaporation, à l'inverse des formules de Penman ou de Morton. La faible variabilité spatiale tient vraisemblablement du faible nombre de degrés de liberté du modèle d'Oudin, qui ne dépend que de la température, en comparaison des autres modèles. Néanmoins, les différences de degrés de liberté n'expliquent pas complètement les comportements des modèles d'évaporation potentielle, dans la mesure où les formules de Penman et de Penman-Monteith présentent des différences de variabilité temporelle (Penman-Monteith étant la plus variable) alors que les variables utilisées dans les calculs des deux formules sont les mêmes. Il faut noter toutefois que les coefficients de variation analysés masquent des différences dans la variabilité interannuelle de l'évaporation moyenne sur les bassins. À cette échelle temporelle, la formule d'Oudin apparaît comme très peu variable, tandis que les autres modèles montrent des évolutions plus marquées très similaires entre elles. L'augmentation observée est vraisemblablement un signe des effets du changement climatique en cours.

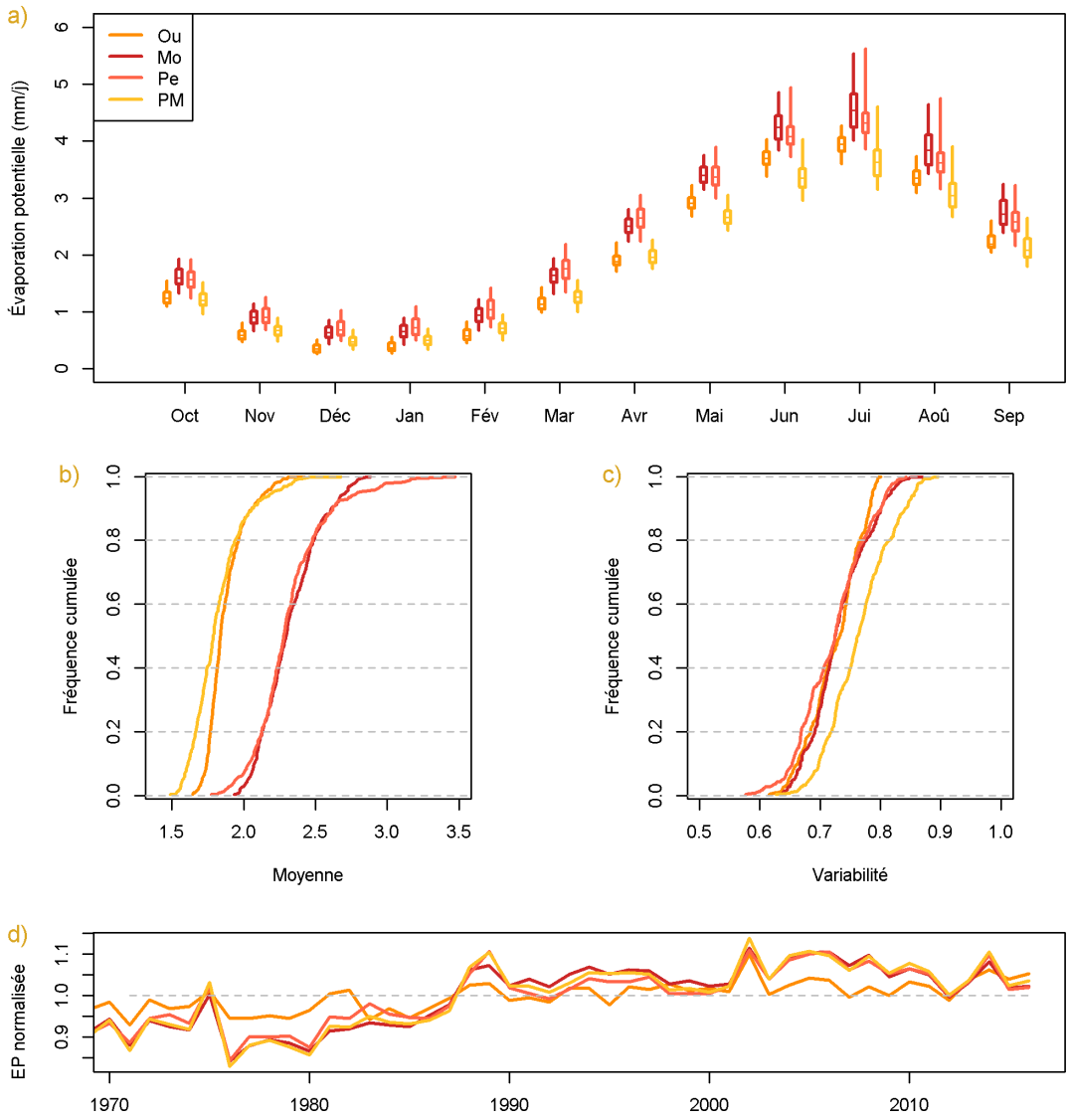


FIGURE 3.10 : a) Distribution sur les bassins français des moyennes mensuelles de l'évaporation potentielle calculée par les différentes formules, entre 1971 et 2016. Les couleurs du jaune au rouge foncé ont été choisies de telle sorte à indiquer les évaporations moyennes dans l'ordre croissant. b) Distribution des évaporations moyennes sur les bassins français sous forme de fréquence cumulée. c) Distribution des coefficients de variation sur les bassins français des évaporations sous forme de fréquence cumulée. d) Moyennes annuelles sur l'ensemble des bassins des évaporations potentielles normalisées par leurs moyennes respectives.

3.3.4.b. Comparaison sur les bassins australiens

En Australie, l'été et l'hiver sont inversés par rapport à la France, ce qui explique un cycle annuel de l'évaporation en opposition de phase en comparaison du cycle français (Figure 3.11).

La relation d'ordre entre les formules d'Oudin et de Morton est la même qu'en France : la formule d'Oudin produit l'évaporation potentielle la plus faible. En revanche, les disparités spatiales et temporelles de l'évaporation potentielle sont plus fortes lorsque celle-ci est calculée avec la formule d'Oudin qu'avec la formule de Morton. Ces disparités s'expliquent vraisemblablement par la diversité des altitudes moyennes des bassins australiens, distribuées de façon assez homogène sur une plage assez large de valeurs, qui étale la distribution des températures moyennes de l'échantillon.

L'approche radiative de la formule de Morton la rend de fait peu sensible à cette répartition, au contraire de la formule d'Oudin qui ne dépend que de la température moyenne.
On notera que l'évolution interannuelle de l'évaporation potentielle ne montre pas d'augmentation sensible à partir de 1997, date du début de la *Millenium Drought* en Australie.

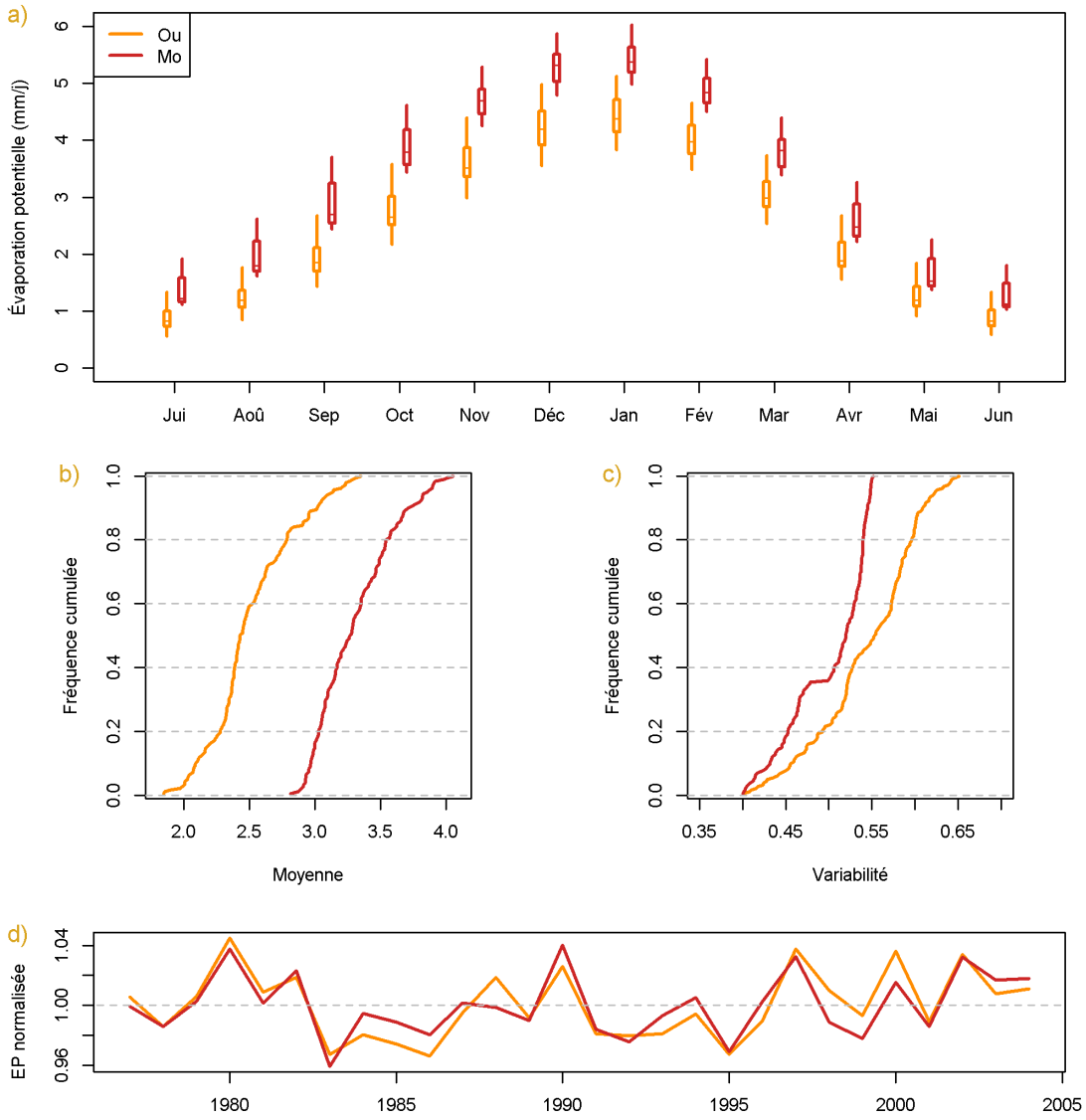


FIGURE 3.11 : a) Distribution sur les bassins australiens des moyennes mensuelles de l'évaporation potentielle calculée par les deux formules, entre 1977 et 2004. Les couleurs choisies pour l'évaporation potentielle d'Oudin (orange) et de Morton (rouge foncé) sont les mêmes que leur équivalent dans l'échantillon français. b) Distribution des évaporations moyennes sur les bassins australiens sous forme de fréquence cumulée. c) Distribution des coefficients de variation sur les bassins australiens des évaporations sous forme de fréquence cumulée. d) Moyennes annuelles sur l'ensemble des bassins des évaporations potentielles normalisées par leurs moyennes respectives.

Les différences observées entre les évaporations potentielles produites par les quatre formules comparées dans la thèse sont parfois assez marquées et concernent des caractéristiques varieés. On peut soupçonner que ces différences se répercutent dans l'estimation des paramètres en calage, par exemple pour boucler le bilan hydrique simulé.

3.4. Calage des modèles hydrologiques

3.4.1. Fonctions objectif utilisées

Les fonctions objectif employées pour caler les modèles hydrologiques ont, de même que les modèles d'évaporation potentielle, une influence sur l'estimation des paramètres. Nous avons déjà mentionné le fait qu'il existait de très nombreux critères d'erreur pour évaluer les performances des modèles hydrologiques, et donc de très nombreuses fonctions objectif potentiellement utilisables. Nous avons fait le choix d'en sélectionner quatre en vue d'une comparaison à l'aune de la robustesse des modèles. Ces fonctions objectif se distinguent par leur popularité ou par le potentiel qui a pu leur être prêté dans la littérature.

Cette liste courte est caractérisée par un dénominateur commun à la base de la construction de chacune des fonctions objectif utilisées. Elles sont en effet construites à partir du même critère de départ, très simple et fréquemment utilisé en modélisation hydrologique : le KGE (Gupta *et al.*, 2009). Celui-ci est décrit par l'équation suivante :

$$KGE\left[Q\right] = 1 - \sqrt{\left(r\left[Q\right] - 1\right)^2 + \left(\frac{\sigma_s\left[Q\right]}{\sigma_o\left[Q\right]} - 1\right)^2 + \left(\frac{\mu_s\left[Q\right]}{\mu_o\left[Q\right]} - 1\right)^2} \tag{3.17}$$

Où $r\left[Q\right]$ correspond à la corrélation de Pearson entre les débits simulés et observés, $\sigma_s\left[Q\right]$ et $\sigma_o\left[Q\right]$ sont les écarts-types des débits simulés et observés, et $\mu_s\left[Q\right]$ et $\mu_o\left[Q\right]$ sont les moyennes des débits simulés et observés. On nomme parfois α le rapport des écarts-types simulés et observés, et β le rapport des moyennes. Kling *et al.* (2012) ont proposé une version alternative du KGE en remplaçant le ratio des écarts-types par celui des coefficients de variation, que nous avons décidé d'écarter afin que les composantes du critère soient strictement indépendantes entre elles.
Le KGE est la première des quatres fonctions objectif utilisées dans la thèse.

Les raisons pour lesquelles le KGE est le dénominateur commun des autres fonctions objectif sont multiples. D'une part, il a déjà été constaté que le KGE était de loin le critère le plus employé en calage dans les études sur la robustesse des modèles hydrologiques. Ce choix est parfois justifié dans ces études (e.g. Fowler *et al.*, 2016) par une volonté de ne pas recourir au NSE (Nash et Sutcliffe, 1970) étant donné les nombreux problèmes qu'il soulève, déjà évoqués au Chapitre 2. Il apparaît crucial de vérifier la pertinence du KGE afin d'envisager de reconsidérer les résultats de ces études à l'aune du choix de la fonction objectif.
Le KGE possède par ailleurs l'avantage majeur d'être interprétable facilement dans la mesure où il est calculé comme une distance euclidienne dans un espace à trois dimensions, chacune des dimensions représentant l'erreur du modèle au regard d'un moment statistique simple. De ce fait, la valeur du KGE indique toujours au moins l'écart entre le modèle évalué et le modèle parfait pour la composante où cet écart est maximal. De ce point de vue, se servir d'un critère simple et interprétable comme base commune pour les autres fonctions objectif est censé faciliter l'analyse des résultats.
Dans le même ordre d'idée, nous estimons que la comparaison de métriques proches et construites comme des variations d'un motif unique est plus aisée que la comparaison de critères tous très différents. En effet, dans une telle situation il serait difficile d'identifier un unique facteur expliquant la meilleure robustesse observée en calant les modèles avec l'une ou l'autre des fonctions objectif, tant ceux-ci seraient nombreux *a priori*.

La deuxième fonction objectif utilisée dans le travail de thèse est le KGE calculé sur la racine carrée des débits. Le choix de la racine carrée pour la transformation des débits est justifié par le fait qu'elle permet de diminuer « en douceur » le poids des erreurs sur les hauts débits et donc l'hétéroscédasticité des erreurs sans s'éloigner du comportement du critère calculé sur les débits non transformés (e.g. Le Moine, 2008; Pushpalatha *et al.*, 2012; Lebecherel, 2015). Cette transformation a par ailleurs déjà été appliquée sur le KGE dans l'étude de Coron *et al.* (2014) portant sur la robustesse des modèles. Enfin, les résultats de Fowler *et al.* (2018b) laissent supposer que cette transformation associée à un critère aux moindres carrés (le NSE) permet d'obtenir des performances très nettement meilleures en DSST en comparaison du critère appliqué aux débits non transformés. Ces travaux n'incluent pas de tests de la racine carrée sur le KGE.
La deuxième fonction objectif, que l'on nommera par la suite sobrement « KGE racine » par souci de concision est donc calculée ainsi :

$$KGE\left[\sqrt{Q}\right] = 1 - \sqrt{\left(r\left[\sqrt{Q}\right] - 1\right)^2 + \left(\frac{\sigma_s\left[\sqrt{Q}\right]}{\sigma_o\left[\sqrt{Q}\right]} - 1\right)^2 + \left(\frac{\mu_s\left[\sqrt{Q}\right]}{\mu_o\left[\sqrt{Q}\right]} - 1\right)^2} \tag{3.18}$$

Il convient de noter que les sous-critères du KGE racine ne sont pas aussi facilement interprétables que ceux du KGE, puisque calculés sur la racine des débits.

Enfin, il se trouve qu'une variante du KGE semble justement montrer un potentiel intéressant dans la seule étude parue à la connaissance de l'auteur au sujet du choix de la fonction objectif pour améliorer la robustesse des modèles hydrologiques (Fowler *et al.*, 2018b). Nous avons donc jugé pertinent de sélectionner cette variante dans notre liste de fonctions objectif pour confirmer, ou non, ce potentiel.
Cette troisième fonction objectif choisie est le *Split* KGE, ou SKGE, proposé par Fowler *et al.* (2018b). Avec le *Refined index of Agreement* de Willmott *et al.* (2012), le SKGE est la meilleure fonction objectif identifiée par ces auteurs pour améliorer la robustesse de modèles hydrologiques sur des bassins australiens sur des périodes très sèches marquées par la *Millenium Drought*. Des deux fonctions objectif, nous avons choisi de ne retenir que le SKGE à la suite d'une discussion avec K. Fowler, venu en France présenter ses travaux, simplement parce qu'elle présentait l'avantage d'être une variante du KGE et parce qu'elle n'avait jamais été testée auparavant sur un grand échantillon de bassins versants.
Le SKGE est contruite comme la moyenne des KGEs interannuels, c'est-à-dire calculés séparément sur chaque année hydrologique présentant suffisamment d'observations de débit disponibles (i.e. moins de 10% de lacunes) :

$$SKGE\left[Q\right] = \frac{1}{N}\sum_{i=1}^{N} KGE_i\left[Q\right] \tag{3.19}$$

Où N est le nombre d'années hydrologiques disponibles et $KGE_i\left[Q\right]$ est le KGE calculé sur les débits non transformés de l'année i.
L'idée qui sous-tend cette formulation est la meilleure prise en compte des années sèches, où l'erreur a typiquement moins de poids dans le calcul du KGE. Le SKGE donne un poids équivalent à chaque année hydrologique dans le calcul des performances du modèle et est donc supposé améliorer la robustesse des modèles hydrologiques lorsque transférés de conditions hydro-climatiques humides vers des conditions sèches.

La quatrième fonction objectif est naturellement le SKGE calculé sur la racine carrée des débits, afin de combiner les deux variantes du KGE précédentes. Ce critère est donc supposé à la fois diminuer l'influence des années les plus humides et l'hétéroscédasticité intra-annuelle des erreurs en dimi-

nuant l'influence des périodes de hautes eaux. De ce fait, c'est la fonction objectif qui vise le plus explicitement à améliorer les performances du modèle pendant les basses eaux et les sécheresses. De même que pour le KGE racine, on désignera cette fonction comme le SKGE racine :

$$SKGE\left[\sqrt{Q}\right] = \frac{1}{N}\sum_{i=1}^{N} KGE_i\left[\sqrt{Q}\right] \tag{3.20}$$

Le tableau 3.5 décrit succinctement les avantages et inconvénients des quatre fonctions objectif comparées dans la suite de la thèse.

TABLEAU 3.5 : Récapitulatif des fonctions objectif utilisées.

Nom	Avantages supposés	Inconvénients supposés
KGE	- populaire - facilement interprétable	- associée à un constat général de manque de robustesse
KGE racine	- parfois utilisée - limite l'hétéroscédasticité des erreurs - facilement interprétable	- associée à un constat de manque de robustesse
SKGE	- potentiel à confirmer - répartition homogène du poids des années - plutôt interprétable	- non connus
SKGE racine	- limite l'influence des erreurs en crues - plutôt interprétable	- non connus

3.4.2. Algorithme de calage

L'optimisation des fonctions objectif de calage est réalisée avec l'algorithme « pas-à-pas » développé par Michel (1989), et inclus dans le package *airGR* (Coron *et al.*, 2017, 2018). Cette méthode est traditionnellement utilisée pour le calage des modèles du Génie Rural, dont GR4J fait partie, et a déjà servi avec succès pour le calage des modèles TOPMO (e.g. Mathevet, 2005; Oudin *et al.*, 2006a) et SimHyd (e.g. Coron, 2013).
La méthode « pas-à-pas » combine un algorithme de recherche global et un algorithme de recherche local.

L'algorithme de recherche global examine les performances des modèles à partir d'une grille de jeux de paramètres pré-définie à partir des valeurs les plus probables des paramètres des modèles. Cette grille est de taille 3^n, avec n le nombre de paramètres libres du modèle à caler. Le jeu de paramètres de la grille ayant obtenu les meilleures performances au regard de la fonction objectif est sélectionné à l'issue de cette étape de « pré-filtrage » pour servir de point de départ à la recherche locale.
L'algorithme de recherche local teste les performances du modèle dans le voisinage de chaque paramètre, à partir d'un pas ajouté ou soustrait à la valeur des paramètres pour définir les jeux de paramètres candidats. Les paramètres du modèle subissent chacun une transformation de telle sorte à adapter le pas de recherche à chaque paramètre. À chaque itération, le meilleur jeu de paramètres parmi la liste ainsi définie est sélectionné comme point de départ de l'itération suivante. Le pas est augmenté ou diminué pour chaque paramètre selon le gradient d'augmentation des performances, pour accélerer la convergence vers l'optimum, par exemple lorsque des itérations

successives dans une direction ont apporté des améliorations. Lorsque le pas atteint une valeur faible en dessous d'un seuil pré-défini, la recherche s'arrête : le modèle est calé.

L'intérêt majeur de l'algorithme global est la garantie que la recherche de l'optimum n'est pas confiné dans une zone resteinte de l'espace des paramètres. Ainsi, il limite les risques de convergence vers un optimum secondaire. La faiblesse principale de l'algorithme global est son manque d'efficience dans la mesure où la recherche doit être réalisée sur une grille suffisamment fine pour rendre compte correctement du relief de la surface de réponse du modèle. Les algorithmes locaux sont au contraire réputés pour être efficients, mais sujets à la convergence vers des optima locaux ce qui peut réduire leur efficacité. La combinaison des deux modes de recherche, global et local, permet de bénéficier des avantages des deux.
Dans la mesure où l'ambition de la thèse suppose d'effectuer de très nombreux calages dans le cadre du DSST et surtout du GSST, la rapidité de l'algorithme combinée à son efficacité de Michel (1989) est particulièrement appréciable.

3.5. Critères d'évaluation des performances des modèles

Considérant la diversité des aspects du régime hydrologique qu'il est nécessaire que les modèles soient capables de représenter, nous avons choisi d'utiliser plusieurs indicateurs des performances des modèles hydrologiques en plus du KGE et de ses composantes introduits précédemment. Nous nous sommes efforcés de choisir ces critères de telle sorte à évaluer un large spectre d'aspects hydrologiques tout en restant aussi synthétiques que possible, afin de faciliter la lecture et l'analyse des résultats. La liste résultant de ce processus est composée d'une petite dizaine de métriques, qui seront impliquées à différentes occasions au gré des travaux présentés.

Un premier critère incontournable de l'évaluation des modèles hydrologiques est le critère de Nash et Sutcliffe (1970), appelé NSE. Certains défauts du NSE ont été relevés dans le Chapitre 2 précédent, néanmoins sa popularité en fait un critère facilement interprétable pour la majorité des hydrologues. L'un des intérêts du NSE est également l'utilisation abondante qui en a été faite en association avec des transformations des débits. Le Moine (2008) propose une description des pas de temps contributifs de l'erreur calculée avec le NSE associé à plusieurs transformations, telles que les fonctions inverse, logarithme ou racine carrée. Le NSE s'utilise notamment avec toute la gamme des transformations puissance, exposants négatifs inclus (e.g. Pushpalatha *et al.*, 2012). Le NSE associé à une transformation puissance se calcule ainsi :

$$NSE_{\lambda}\left[Q\right] = 1 - \frac{\sum_{i=1}^{N}\left(Q_{s,i}^{\lambda} - Q_{o,i}^{\lambda}\right)^{2}}{\sum_{i=1}^{N}\left(\overline{Q_{o}^{\lambda}} - Q_{o,i}^{\lambda}\right)^{2}} \tag{3.21}$$

Lorsque λ tend vers 0, le comportement de NSE_{λ} se rapproche du comportement de $NSE_{\log}$. Dans le manuscrit, nous utiliserons les transformations inverse, logarithme, racine carrée, et carré. Il convient de noter que l'un des avantages du NSE par rapport au KGE pour l'association avec des transformations des débits est la possibilité d'utiliser le logarithme, ce qui est fortement déconseillé avec le KGE (Santos *et al.*, 2018b). Le logarithme présente notamment l'intérêt de balayer assez équitablement le spectre des débits (e.g. Pushpalatha *et al.*, 2012).

Pour l'évaluation des simulations en étiage, nous avons parfois préféré l'utilisation d'un autre critère que le NSE associé à la fonction inverse ou au logarithme. En effet, le logarithme ne cible pas suffisamment les étiages, tandis que la fonction inverse peut avoir un comportement chaotique lorsque

les débits sont très faibles, ce qui ne facilite pas son interprétation. Nous utiliserons donc le MARE, pour *Mean Absolute Relative Error*, qui a la caractéristique d'être bien corrélé aux jugements produits par les hydrologues en évaluation des simulations d'étiage (Crochemore *et al.*, 2015). De ce fait, on peut considérer que les valeurs prises par le MARE sont parlantes dans la mesure où elles sont en adéquation avec l'évaluation qui serait rendue si nous devions comparer visuellement les étiages simulés aux observations. Le MARE se calcule simplement :

$$MARE = \frac{1}{N}\sum_{i=1}^{N}\frac{|Q_{s,i} - Q_{o,i}|}{Q_{o,i}} \tag{3.22}$$

Dans le but d'apprécier la capacité des modèles à représenter les volumes écoulés en conditions climatiques contrastées, nous avons également utilisé le biais du modèle, qui est une composante du KGE. Dans certaines situations, il a été choisi de normaliser le biais pour calculer l'erreur volumique sous forme de pourcentage en valeur absolue :

$$\text{Biais absolu} = \left|\frac{\mu_s}{\mu_o} - 1\right| \tag{3.23}$$

L'analyse de la capacité des modèles à représenter la variabilité du débit simulé sera réalisée à partir des chroniques mensuelles de rendement des précipitations simulées et observées, au moyen de la corrélation et du rapport des écarts-types (adaptées des composantes du KGE). Pour rappel, le rendement des précipitations est calculé comme le rapport du débit moyen sur les précipitations moyennes.
L'intérêt de l'agrégation au pas de temps mensuel permet de ne pas considérer les variations journalières du débit, qui sont déjà évaluées dans les critères comme le NSE ou le MARE, et de se concentrer sur les modes de variabilité caractérisés par des temps caractéristiques saisonniers voire interannuels. C'est aussi une manière de réduire l'influence des crues sur le calcul du rapport des écarts-types.
D'autre part, le rendement des précipitations a été préféré au simple débit dans l'objectif d'évaluer le spectre simulé des réponses du bassin versant aux forçages météorologiques. En effet, les fortes valeurs de rendement des précipitations sont atteintes lorsque le bassin produit beaucoup d'écoulement par rapport aux précipitations, c'est-à-dire généralement lorsque ses réserves d'eau sont élevées (en surface, en sous-sol, ou les deux à la fois). Inversement, les faibles valeurs de rendement sont atteintes lorsque les précipitations ne sont pas efficaces car stockées ou perdues. Ainsi, analyser la capacité d'un modèle à reproduire les variations du rendement des précipitations permet en un sens de juger des processus simulés intervenant dans le bilan hydrique des bassins versants. Ainsi que souligné au chapitre précédent, ces processus semblent justement faire défaut aux modèles actuels.
Les critères ad hoc introduits se formulent de la manière suivante :

$$\alpha_{RR} = \frac{\sigma\left[\frac{Q_{s,m}}{P_m}\right]}{\sigma\left[\frac{Q_{o,m}}{P_m}\right]} \tag{3.24}$$

$$\rho_{RR} = \text{cor}\left(\frac{Q_{o,m}}{P_m}, \frac{Q_{s,m}}{P_m}\right) \tag{3.25}$$

Où Q_m et P_m sont les débits et précipitations agrégés au pas de temps mensuel. L'indice RR signifie *Runoff Ratio*, terme employé en anglais pour désigner le rendement des précipitations, et permet de marquer la distinction entre ces métriques et les composantes du KGE.

Une autre métrique conçue spécifiquement lors de cette thèse pour évaluer la robustesse des modèles hydrologiques est présentée au chapitre suivant.

Lors de l'analyse des performances des modèles sur les bassins de l'échantillon, nous avons parfois recouru à des tests statistiques. Ceux-ci ont majoritairement servi à évaluer le rôle joué par le hasard dans les écarts constatés entre différentes distributions de performances ou de caractéristiques physiques. Pour cela, nous avons appliqué le test de Wilcoxon-Mann-Whitney (W, 1945; Mann et Whitney, 1947), un test non paramétrique permettant de tester la proximité des médianes de deux distributions. C'est le test standard auquel nous avons recouru pour la comparaison de deux groupes. Le seuil de significativité a été adapté aux occurences des analyses menées au cours de la thèse. Pour des tests réalisés moins de quelques dizaines de fois, le seuil a été fixé à 1%. C'est par exemple le cas des tests entre plusieurs distributions de performances réalisés pour une poignée de métriques.
L'évaluation de la significativité des corrélations entre deux séries a été réalisée au moyen du test de Student. Typiquement, les valeurs p des tests de significativité de corrélation pour chaque bassin de l'échantillon ont été fixées à 0.1% pour tenir compte de la multiplication du risque d'obtenir des faux positifs en appliquant un test plusieurs centaines de fois (le nombre moyen de bassins faux positifs est ainsi inférieur à 1). Les seuils de significativité ont été indiqués toutes les fois où cela a été jugé nécessaire, afin d'éviter les jugements arbitraires.

3.6. Synthèse

Base de données

Les problématiques de la thèse sont évaluées sur un échantillon de bassins français et australiens couvrant un large spectre de spécificités topographiques et hydro-climatiques. Les bassins ont été choisis de telle sorte à être faiblement influencés par la neige et le moins possible par les activités humaines.

Modèles hydrologiques

Les simulations de trois modèles hydrologiques sont étudiés sur cet échantillon. Les modèles GR4J, TOPMO et SimHyd sont conceptuels mais de complexité variées, le plus simple étant le modèle GR4J. Leur association avec un module de neige aux paramètres pré-fixés permet de tenir compte des précipitations solides sans augmenter le nombre de paramètres libres à caler.

Formules d'évaporation potentielle

Les modèles hydrologiques sont forcés par des entrées d'évaporation potentielle estimées à partir de quatre formules dont les complexités s'échelonnent du modèle empirique simple (Oudin) au modèle théorique de référence (Penman et Penman-Monteith), en passant par un modèle intermédiaire (Morton). Ces formules ont été fréquemment utilisées avec des modèles hydrologiques dans les deux régions d'étude.

Fonctions objectif

Plusieurs fonctions objectif seront testées pour le calage des modèles, sélectionnées parmi les fonctions objectif de référence et/ou celles montrant un potentiel intéressant pour la transférabilité des paramètres. Toutes ces fonctions objectif ont été développées à parti du KGE. D'autres métriques seront également employées pour l'évaluation des performances du modèle, dont une présentée au chapitre suivant.

Partie II

Diagnostic de la robustesse des modèles hydrologiques

Chapitre 4.

Procédures d'évaluation de la robustesse des modèles hydrologiques

Introduction

Ce chapitre présente les procédures utilisées pour évaluer la robustesse des modèles hydrologiques.
La robustesse n'est pas une grandeur en soi, et il n'existe pas de méthode claire pour la calculer : elle correspond plutôt à l'écart entre les performances d'un modèle sur une période de contrôle par rapport à ses performances sur cette même période si celle-ci avait servi au calage des paramètres du modèle. L'estimation de la robustesse est donc intimement liée au calcul des performances d'un modèle, mais correspond à un cadre d'évaluation spécifique de celles-ci. Les critères de performance utilisés dans la thèse ayant été définis au chapitre précédent, ce chapitre décrit dans un premier temps les procédures de calage-contrôle implémentées pour l'analyse de la robustesse des modèles hydrologiques. La suite du chapitre est consacrée à l'introduction d'un critère de performance spécifiquement construit pour l'estimation directe de la robustesse du modèle.

4.1. Quelle stratégie de test pour quel objectif ?

4.1.1. Le *Generalized Split-Sample Test* : stratégie explorative en extrapolation

Le *Differential Split-Sample Test* est la procédure de calage-contrôle de référence pour l'évaluation de la robustesse des modèles hydrologiques. La plupart des procédures alternatives sont fondées sur le même principe, à savoir un examen des performances d'un modèle dans des conditions climatiques différentes de celles rencontrées en calage. En d'autres termes, son principe repose sur l'évaluation de la capacité d'un modèle à extrapoler le comportement d'un bassin versant, et ainsi à la détermination du domaine de validité de ce modèle.

Dans le cadre de nos travaux de comparaison des choix de modélisation, comme ceux des méthodes de calage ou des formulations de l'évaporation potentielle, nous avons toutefois évité l'emploi du DSST, pour plusieurs raisons.
Le DSST suppose la sélection de variables hydrologiques ou climatiques servant à définir des périodes distinctes pour le calage-contrôle. Dans un contexte d'évaluation des modèles en vue de préparer une étude des conséquences du changement climatique sur le régime hydrologique des bassins d'une région, cette approche est tout à fait justifiée. En effet, les résultats des projections climatiques peuvent indiquer des tendances sur l'évolution du contexte climatique de la région d'étude et désigner le domaine de variations hydro-climatiques dans lequel il est nécessaire que les modèles employés soient robustes. Dans cette thèse néanmoins, l'objectif poursuivi est davantage de proposer une évaluation générale des modèles hydrologiques plutôt que d'estimer leur robustesse face à certaines catégories de changement. À cet égard, la sélection de variables hydro-climatiques spécifiques pour définir les périodes du DSST aurait pour conséquence de réduire l'information exploitable des expérimentations réalisées.
Le contraste géographique évident entre les sous-échantillons de bassins français et australiens réduit également la pertinence du choix d'un ensemble de variables caractérisant des évolutions possibles des conditions hydro-climatiques communes à tous les bassins versants de l'échantillon. Au sein de ces sous-échantillons, des différences intra-régionales sont susceptibles d'exister, particulièrement en France étant donné la diversité des régimes climatiques qui caractérisent le

territoire, et accentuent la difficulté de ce choix.
Enfin, il peut être intéressant de se placer dans un cadre d'analyse neutre, sans formuler d'hypothèses *a priori* sur les conditions limitant la transférabilité des modèles hydrologiques. Une étude *a posteriori* pourrait dans ce cadre permettre de proposer un diagnostic plus fin de la robustesse des modèles.

À la place du DSST traditionnel, nous avons mis en œuvre la variante du *Generalized Split-Sample Test* introduite par Coron *et al.* (2014). La procédure est identique à celle employée pour la construction des courbes de biais glissants (voir 2.1.1.c).

- Définition de sous-périodes de 10 ans successives à partir du passage itératif d'une fenêtre temporelle glissant d'un an à chaque itération
- Calages successifs d'un modèle hydrologique sur les sous-périodes ainsi définies
- Évaluation du modèle muni de chacun des jeux de paramètres obtenus sur l'ensemble des données disponibles (période « totale »)

Ce procédé possède plusieurs avantages. Il permet d'une part de réaliser un grand nombre de calages-contrôles et d'évaluer de nombreuses configurations de transfert des paramètres, et offre ainsi une évaluation aussi exhaustive que possible de la robustesse d'un modèle pour la simulation des débits d'un bassin versant. D'autre part, l'unicité de la période de contrôle, qui est la même quelle que soit la période de calage, autorise la comparaison des performances obtenues en calant un modèle sur différentes périodes de calage. De ce fait, il est possible de calculer des valeurs moyennes de ces performances pour l'ensemble des calages-contrôles du GSST, ou bien de comparer les performances du modèle pour différents types de changements des conditions hydroclimatiques. Enfin, la période de contrôle correspondant à l'ensemble des données disponibles, le calcul des performances intègre les erreurs d'un modèle dans des conditions de simulation variées et permet d'estimer plus fidèlement la robustesse réelle du modèle. Ainsi, cette estimation dépend peu des effets particuliers de certaines conditions sur l'erreur du modèle.

Malgré les opportunités qu'il offre pour un diagnostic affiné de la robustesse des modèles hydrologiques, l'emploi du GSST présente malgré tout le défaut de ne pas séparer strictement les périodes de calage de la période de contrôle. La non-indépendance des périodes de calage et de contrôle introduit un risque d'atténuation de l'exigence du test dans la mesure où les performances des modèles testés sont nécessairement optimales sur une partie de la période de contrôle. Toutefois, cette optimalité reste à nuancer dans la mesure où elle concerne la fonction objectif utilisée pour le calage, mais pas obligatoirement les performances d'un modèle jugées avec d'autres métriques. Cette raison nous a conduit à impliquer un grand nombre de critères de performance (voir Chapitre 3), différents des fonctions objectif de calage, pour l'évaluation de la robustesse des modèles hydrologiques réalisée aux Chapitres 5 et 6 par le biais du GSST.

4.1.2. Intérêt d'une stratégie de test en interpolation

Dans le cadre d'études de la robustesse de modèles conceptuels ou de modèles à base physique parcimonieux, les stratégies d'évaluation en extrapolation fondées sur la partition des données pour le calage et le contrôle apparaissent comme les plus pertinentes d'un point de vue méthodologique. Elles correspondent en effet aux exigences les plus strictes en matière de validation de modèles. Pour l'étude de modèles complexes en revanche, pour lesquels les modélisateurs ne peuvent se permettre de renoncer à une partie des données pour la détermination de leurs paramètres, les stratégies de *Split-Sample Test* ne sont pas applicables.

Bien que les modèles utilisés dans cette thèse soient tous conceptuels et adaptés aux stratégies

de test en évaluation, nous avons cherché à proposer une méthode plus générale et applicable à tout type de modèle. Une telle stratégie ne peut pas reposer sur le recours au calage-contrôle : sa mise en œuvre ne peut donc que résulter d'un fléchissement du niveau d'exigence méthodologique. Puisque l'évaluation en extrapolation n'est pas envisageable, cette nouvelle méthode devrait incarner une stratégie de test en interpolation.
La courbe des biais glissants présentée au Chapitre 2 est un exemple d'outil d'analyse en interpolation, dans la mesure où elle permet de rendre compte de la capacité d'un modèle à simuler les variations interannuelles du débit à partir d'une seule réalisation du modèle. La courbe des biais glissants est toutefois un outil graphique, qui suppose d'une part une certaine subjectivité pour son interprétation, et qui d'autre part ne permet pas d'estimer rapidement la robustesse d'un modèle sur un large échantillon de bassins versants sans une analyse détaillée de son comportement sur chaque bassin.

Nous proposons de valoriser la courbe des biais glissants par le calcul d'un critère de performance synthétique, offrant un aperçu des capacités d'un modèle à représenter l'évolution temporelle du régime hydrologique. Ce travail est présenté dans la suite sous la forme d'une note technique, soumise pour publication au journal *Hydrology and Earth System Sciences* (HESS, Royer-Gaspard *et al.*, 2021). La note technique n'implique que le modèle GR4J sur le sous-échantillon de bassins français, mais les résultats obtenus pour les autres modèles et sur le reste de l'échantillon sont exposés et discutés à la suite de la note technique. La note technique est intégrée telle quelle au mémoire de thèse, bien que sa mise en forme ait été légèrement adaptée. Les références citées dans la note sont incluses dans celles du mémoire. L'Annexe A regroupes les annexes à la note technique.

4.2. Note technique : *A proxy metric to assess hydrological model robustness in a changing climate*

Abstract

The ability of hydrological models to perform in climatic conditions different from those encountered in calibration is crucial to ensure a reliable assessment of the impact of climate change in water management sectors. However, most evaluation studies based on the Differential Split-Sample Test (DSST) endorsed the consensus that rainfall-runoff models lack climatic robustness. Models typically exhibit substantial errors on streamflow volumes applied under climatologically different conditions. In this technical note, we propose a new performance metric to evaluate model robustness without applying the DSST and which performs with a single hydrological model calibration. The Proxy for Model Robustness (PMR) is based on the systematic computation of model error on sliding sub-periods of the whole streamflow time series. We demonstrate that the metric shows patterns similar to those obtained with the DSST for a conceptual model on a set of 377 French catchments. An analysis of sensitivity to the length of the sub-periods shows that this length influences the values of the PMR and its adequation with DSST biases. We recommend a range of a few years for the choice of sub-period lengths, although this should be context-dependent. Our work makes it possible to evaluate the temporal transferability of any hydrological model, including uncalibrated models, at a very low computational cost.

4.2.1. Introduction

In the context of climate change, quantifying the performance of the models used for assessing the impact of a changing climate is essential for informing model selection and estimating uncertainty. Assessing the impact of a changing climate typically involves a modeling chain ranging from general circulation models to impact models such as catchment hydrological models (Clark *et al.*, 2016). It is now acknowledged that the contribution of hydrological models to the total uncertainty of projections may be significant and should be addressed along with other sources of uncertainty (e.g. Hagemann *et al.*, 2013; Schewe *et al.*, 2014; Vidal *et al.*, 2016; Melsen *et al.*, 2018). A key issue in the reduction of hydrological model uncertainty is the assessment of robustness to climatic changes, i.e., their ability to perform in climatic conditions that differ from those encountered in calibration.

Advocating that hydrological models needed to be tested under conditions that would "represent a situation similar to which the data are to be generated," Klemeš (1986) suggested a series of tests to evaluate the robustness of hydrological models. Among these testing procedures, the most popular scheme to assess model robustness to varying climatic conditions is the Differential Split-Sample Test (DSST). The DSST consists in a calibration-evaluation exercise in two periods of the available time series chosen to be as climatically different as possible. Variants of the DSST have also been proposed for specific purposes, such as the Generalized Split-Sample Test (Coron *et al.*, 2012), which consists in a systematic calibration-evaluation experiment on every couple of independent periods that one can possibly define. However, these variants all rely on the same principles as the DSST (e.g. Dakhlaoui *et al.*, 2019).

Many studies report poor model simulations resulting from the application of the DSST in various modeling contexts (e.g. Thirel *et al.*, 2015b). Among the deficiencies observed in the tested models, a common feature is their tendency to produce biased streamflow simulations in evaluation conditions (e.g. Vaze *et al.*, 2010; Merz *et al.*, 2011; Broderick *et al.*, 2016; Dakhlaoui *et al.*, 2017; Mathevet *et al.*, 2020). Although changes in catchment temperature and/or precipitation are usually associated with volume errors, these errors vary across the tested models and catchments (e.g. Vaze *et al.*, 2010; Broderick *et al.*, 2016; Dakhlaoui *et al.*, 2017). The dire need to improve hydrological models is widely recognized and is considered as one of the 23 unsolved problems in modern hydrology (Blöschl *et al.*, 2019, UPH n°19). However, to improve models we first need a good diagnostic method, and the design of alternatives to the DSST for the evaluation of model robustness could contribute to these advancements.

The first shortcoming of the DSST is its limited application regarding a particular category of hydrological models. Indeed, Refsgaard *et al.* (2014) pointed out that split-sample procedures were not applied to models that are not calibrated. The evaluation of such models is usually performed by testing their spatial transferability with data from proxy sites. It is therefore difficult to compare the robustness of highly complex hydrological models to simpler models such as the ones typically tested in the aforementioned DSST studies. A further limitation is the necessity to determine a set of climatic variables to define different calibration and evaluation periods. This is of course highly relevant in contexts where the direction of future changes is unambiguously predicted. In other situations, however, robustness assessment would benefit from evaluating the model on a wider spectrum of hydro-climatic changes. Variants of the DSST, such as the Generalized Split-Sample test, may circumvent this problem, but at a high computational cost that not all modelers can afford (Coron *et al.*, 2012).

This technical note presents and assesses a way to quantify model robustness as a mathematical performance criterion computed without splitting time series into calibration and evaluation periods.

This criterion is conceived to be a proxy for model robustness, i.e. to reproduce the hydrological model average error as obtained by applying the DSST. It is based on the computation of interannual model bias derived from graphical considerations in the work of Coron *et al.* (2014). In order to be reliable, the Proxy for Model Robustness (PMR) must allow for the indication of typical model biases in independent evaluation periods. It should also help to identify catchments where a model lacks robustness. We summarize the important aspects that we discuss in the following with two research questions:

- Does the PMR faithfully relate to model robustness as assessed in DSST experiments?
- How do computation choices affect the results obtained when applying the PMR?

The first question will be addressed by comparing the metric with model bias obtained in the DSST for a conceptual model across a large set of French catchments. The underpinning mathematical choices will be discussed in a sensitivity analysis comparing the metric and the results obtained by applying the DSST. The description of the PMR is given in Section 4.2.2. The hydrological model and the data are presented in Section 4.2.3. The reliability of the metric is assessed in Section 4.2.4, and the results of the sensitivity analysis are shown and discussed in Section 4.2.5.

4.2.2. Description of the Proxy for Model Robustness

4.2.2.a. Building the "moving bias curve"

Model robustness to climate change is the ability of a hydrological model to perform well under different climatic conditions without parameters being recalibrated to match the changes in the precipitation–streamflow relationship. A robust model should thus adequately simulate streamflow volumes for any type of climatic conditions experienced by a catchment. Coron *et al.* (2014) suggested a simple way to visualize model robustness by computing the bias of a model simulation on sliding sub-periods of the available time series (Figure 4.1). The curve of model bias on the moving sub-periods, named here the "moving bias curve," indicates the temporal evolution of model errors. Since a robust model should perform similarly well whatever the considered sub-period, the flatter the moving bias curve, the more robust a model.

Coron *et al.* (2014) showed that hydrological models would typically not have the ability to flatten their associated moving bias curve. The authors indeed calibrated model parameters on each sub-period of the data and plotted all the produced moving bias curves on the same graph. One of the main conclusions of their study was that the obtained moving bias curves were all almost parallel and that calibration conditions influenced more the vertical positioning of the curves rather than their shape. This observation was true for models of different complexities across a small set of catchments. The phenomenon described by Coron *et al.* (2014) is illustrated in Figure 4.2.

The moving bias curve obtained with the model calibrated on the blue sub-period (1984–1988, coldest sub-period of the time series) is almost parallel to the moving bias curve derived from the calibration on the total period. The y-axis shift corresponds to a model bias on the calibration sub-period almost equal to zero. Calibrating the model on another sub-period (1999–2003, warmest sub-period of the time series, in red) yields a different shift of the moving bias curve, which corresponds to a null model bias on the red calibration sub-period. The shape of the curve being almost identical whatever the calibration period in the illustrated case, it offers an interesting perspective on model robustness. The flatness of the curve is indeed almost independent of the period used for model calibration.
Whether they are parallel or not, depending on the modeling context (model, catchment, data, etc.), the moving bias curves appear to be a relevant tool for analyzing model robustness. Before per-

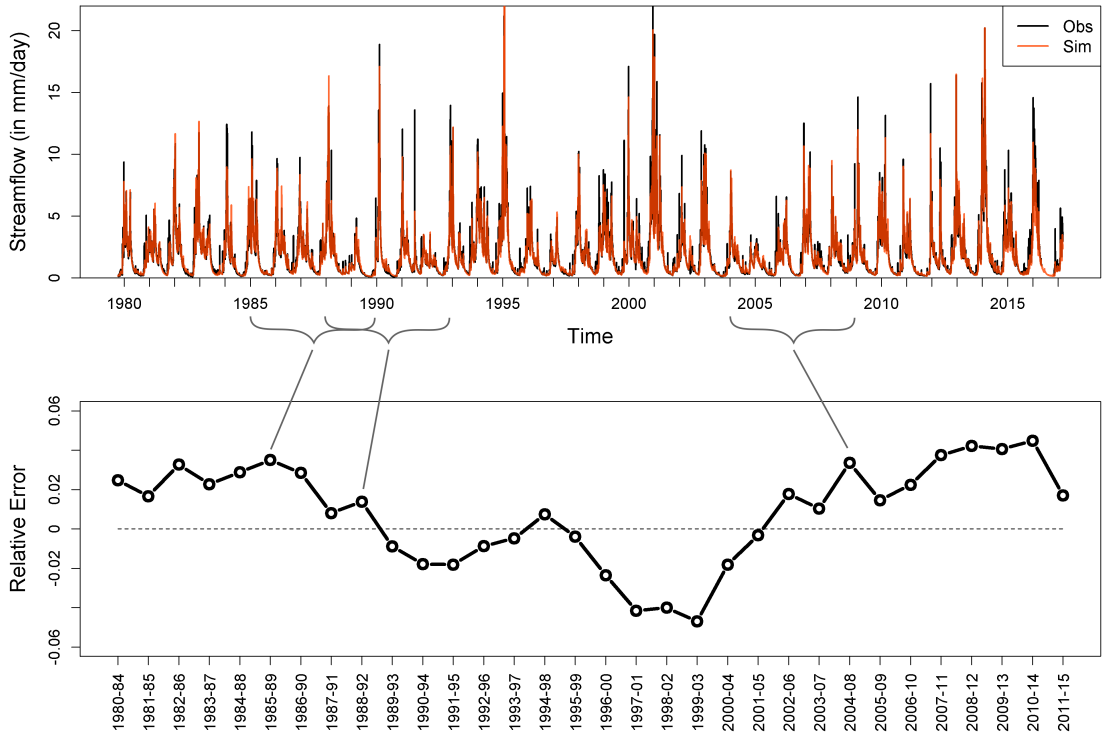


Figure 4.1: Construction of the "moving bias curve." Top panel: daily simulated (orange line) and observed (black line) streamflow. Bottom panel: moving bias curve expressed as the relative error (in %); each point of the curve corresponds to a 5-year sub-period of the time series (as exemplified by the brackets).

forming calibration-evaluation tests, assessing the flatness of the moving bias curve obtained by calibrating a model as well as possible (i.e., with all available data) could be seen as a first estimate for model robustness. We thus propose a simple mathematical expression to calculate this flatness, i.e., a performance metric designed to be a PMR.

4.2.2.b. Computation of the Proxy for Model Robustness

The PMR is based on the computation of the average absolute difference between the actual moving bias curve computed on 5-year sub-periods and a hypothetical flat curve. This hypothetical flat curve is defined as the curve obtained for a hypothetical model, perfectly robust so that its bias on different time sub-periods would remain constant, but imperfect so that this bias would be equal to the mean bias of the evaluated model on the total period. It should be noted that, if the evaluated model is unbiased, as is the case on Figure 4.2 for the moving bias curve obtained by calibrating the model on all the available data, then this reference hypothetical curve is simply defined by "y=0." The PMR is computed as the mean of absolute differences between the model average error on the 5-year sub-periods and the model average error on the total period, normalized by the average observed streamflow (Equation (4.1)). It thus corresponds to the normalized area between the moving bias curve and the hypothetical flat curve.

$$PMR = 2 \times \frac{1}{N} \times \sum_{i=1}^{N} \left| \left(\overline{Q}_{s,i} - \overline{Q}_{obs,i} \right) - \left(\overline{Q_s} - \overline{Q_o} \right) \right| \times \frac{1}{\overline{Q_o}} \tag{4.1}$$

$\overline{Q_o}$ and $\overline{Q_s}$ are the respective averages of the observed and of the simulated streamflows on the total period. $\overline{Q}_{o,i}$ and $\overline{Q}_{s,i}$ are the respective averages of the observed and of the simulated streamflows on the sub-period whose index is i. N is the number of sub-periods that can be defined with a 5-year moving window. The reason for including a factor 2 is to reproduce the bias that would be obtained

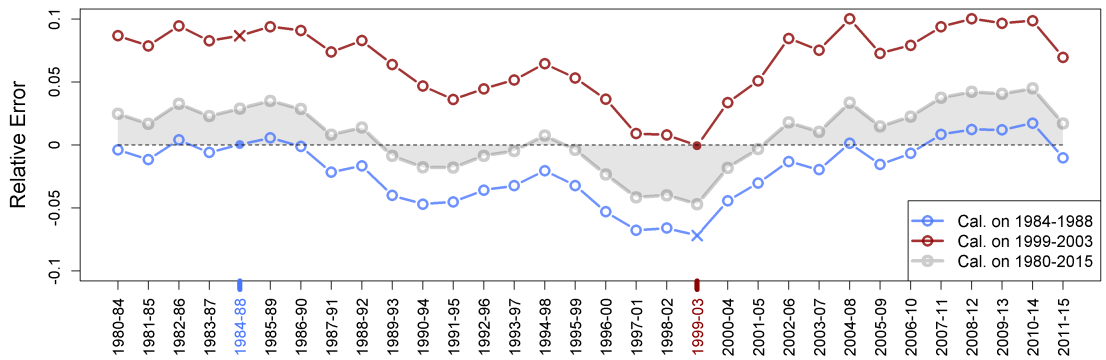


Figure 4.2: Moving bias curves of a hydrological model on the Steir at Guengat catchment (J4313010, in France) for three different calibration periods. The filled dot of the blue (respectively red) curves indicates the 1984–1988 (respectively 1999–2003) calibration period of the model. The crosses indicate the evaluation bias on 1999–2003 of the 1984–1988 calibration and the evaluation bias on 1984–1988 of the 1999–2003 calibration. The grey moving bias curve is obtained by calibrating the model on the total period. The grey area corresponds to a proxy for model robustness, as suggested in Section 4.2.2.b.

in a DSST on sub-periods that are on the opposite side of the moving bias curve (see Figure 4.2). Although the errors for each sub-period are calculated in absolute terms, the normalization by the average observed streamflow allows the resulting value to represent the average relative error produced by the model on the sub-periods as compared with mean observed streamflow.
One reason for computing errors relative to the average streamflow on the whole time series instead of the average streamflow of each sub-period is that this reduces the weight of very dry years. It also avoids dealing with zeros in the denominator in intermittent catchments. This choice is further discussed in Appendix A.2.

4.2.3. Material and methods

4.2.3.a. Dataset

A set of 377 French catchments was used (Figure 4.3) (Delaigue *et al.*, 2020). The selected French catchments cover a variety of physical and hydroclimatic characteristics and were selected as they are little impacted by human activities and have limited solid precipitation ($< 10\%$ of the total precipitation on average). Western France is characterized by an oceanic climate with no marked wet and dry seasons. The climate of the eastern part of the country is more continental, with a larger annual temperature range. Southeastern France has a Mediterranean climate, with humid springs and autumns and dry summers. The yearly average precipitation of the catchments ranges from 662 mm to almost 1926 mm, while the average temperatures vary from 8 to 14.4 °C.

Daily streamflow measurements at the outlet of the catchments were retrieved from the Banque HYDRO (`http://www.hydro.eaufrance.fr/` (last accessed: 21 January 2019), Leleu *et al.*, 2014). Daily meteorological data were supplied by the SAFRAN atmospheric reanalysis (Vidal *et al.*, 2010) aggregated at catchment scale. We used the temperature- and radiation-based formula proposed by Oudin *et al.* (2005a) to compute potential evaporation. In every catchment, streamflow observations cover at least 20 years (40 years on average).

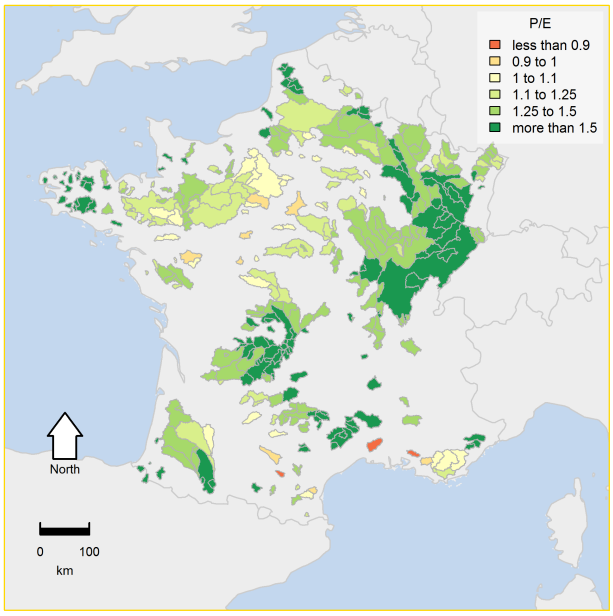


Figure 4.3: Map of the French catchments used in this study. The humidity index is defined as the ratio between average precipitation and average potential evaporation.

4.2.3.b. Hydrological model

The tests were performed with GR4J (Perrin *et al.*, 2003), a daily lumped hydrological model. The model is parsimonious (four parameters to calibrate, two reservoirs, two unit hydrographs) and has been widely used in research studies focusing on hydrological model robustness (e.g. Coron *et al.*, 2014; Broderick *et al.*, 2016; Fowler *et al.*, 2016). The two-parameter CemaNeige degree-day snow module (Valéry *et al.*, 2014) was used to account for solid precipitation. The parameters of the snow module were fixed to median values as recommended by Valéry *et al.* (2014) for catchments with limited impact of snow. The GR4J and CemaNeige models are used with the airGR R package (Coron *et al.*, 2017, 2018).
The parameters of the hydrological models were calibrated by optimizing the Kling-Gupta Efficiency (KGE, Gupta *et al.*, 2009) computed on the square-root of streamflow in order to limit error heteroscedasticity. The optimization algorithm is a simple procedure consisting in a prior global screening on a gross predefined grid, followed by a descent local search from the best parameter set of the grid. The procedure has been successfully used in multiple studies involving GR4J (e.g. Mathevet, 2005; Coron *et al.*, 2014).

4.2.3.c. DSST experiments

DSST experiments consist in selecting contrasted periods (according to some hydrologically relevant indicator) and performing a calibration-evaluation experiment. Our DSST experiments are based on three hydroclimatic variables. The procedure consists in dividing the time series in sub-periods of L consecutive years, and selecting six sub-periods from these. The sub-periods of the DSST are chosen to be:

- The driest and the wettest in terms of precipitation
- The warmest and the coldest in terms of temperature
- The least and the most productive in terms of runoff ratio (computed as the ratio of mean observed streamflow to mean precipitation)

The model parameters are then calibrated on each sub-period and transferred to the sub-period of opposite climate. The process is summarized in Table 4.1.

Table 4.1: Summary of the different setups of the Differential Split-Sample Test. $\bar{Q}$, $\bar{P}$ and $\bar{T}$ respectively stand for average observed streamflow, precipitation and temperature computed on the sub-periods.

	Name of the DSST setup					
	"dry"	"humid"	"warm"	"cold"	"unproductive"	"productive"
Calibration	$\min \bar{P}$	$\max \bar{P}$	$\max \bar{T}$	$\min \bar{T}$	$\min \bar{Q}/\bar{P}$	$\max \bar{Q}/\bar{P}$
Evaluation	$\max \bar{P}$	$\min \bar{P}$	$\min \bar{T}$	$\max \bar{T}$	$\max \bar{Q}/\bar{P}$	$\min \bar{Q}/\bar{P}$

The runoff ratio was preferred to the humidity index since the latter is highly correlated to average precipitation in France and would therefore be redundant with DSST experiments based on precipitation. Since runoff ratio is computed from average streamflow, it cannot be used for predictive purposes of model biases in future climate conditions. However, it estimates how catchments respond to precipitation forcings. Its use in the DSST may thus indicate how well a model is able to represent variations in catchment response to climatic conditions.
The sub-period length for the DSST experiment has been fixed at $L = 5$ years, so as to match the length of the sub-period involved in the computation of the PMR. The length of sub-periods used in the computation of the PMR is discussed in Section 4.2.5. The length of the sub-periods used for the DSST are discussed in Appendix A.3. We remind the reader that the PMR is computed from model simulations obtained by calibrating the model on the whole time series, while the DSST results are obtained through calibration evaluation on sub-periods of the time series. It should also be mentioned that model biases obtained in the DSST were calculated as the absolute differences to 1 so that they could be compared to PMR values, which are positive by definition, as follows:

$$\text{Model bias on subperiod } i \text{ (in \%)} = \left|1 - \frac{\overline{Q}_{s,i}}{\overline{Q}_{o,i}}\right| \tag{4.2}$$

A drawback of this way of computation is that it removes the sign of model errors. Therefore, the sign of model errors in the different DSST setups has been analyzed in Appendix A.1. In the following, model bias obtained in the DSST will systematically be calculated in absolute terms without further notice.

4.2.4. Results: Reliability of the Proxy for Model Robustness to indicate model robustness

4.2.4.a. Comparison of the distributions of PMR values and DSST bias

The PMR is theoretically designed to quantify the average bias that would be obtained from DSSTs of the model if they were calculated in an absolute way. The bias obtained for GR4J for each type of DSST setup is plotted in Figure 4.4.

Compared to absolute biases obtained in the different DSST setups, PMR values have the same order of magnitude as biases in precipitation- or temperature-based experiments. However, the distribution of PMR values exhibits less spread than DSST biases. In the case of DSSTs designed on changes in runoff ratio, model biases are larger than PMR values. The PMR thus seems to relate rather well to model biases observed in typical differential calibration-validation experiments, but also appears to give an underestimated estimation of model biases in highly adverse transfer conditions (see Appendix A.1 for more details about DSST results). In summary, one can say that the results presented in Figure 4.4 simply indicate that, on average, PMR is the same order of magnitude as model bias in DSST.

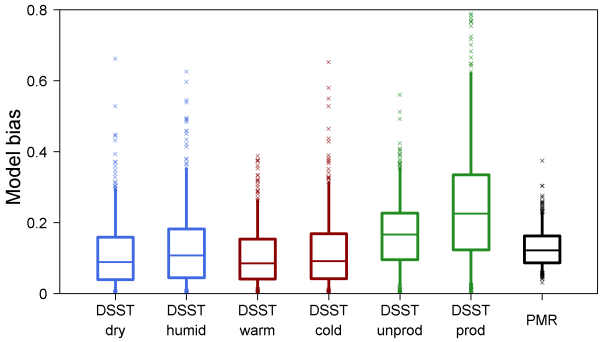


Figure 4.4: Distribution over the catchment set of the absolute value of model bias in DSST for each type of setup, compared to PMR values. The boxplots represent the 5, 25, 50, 75 and 95 quantiles and the crosses denote the outliers. Blue, red and green boxplots are respectively associated to DSST setups based on precipitation, temperature and runoff ratio.

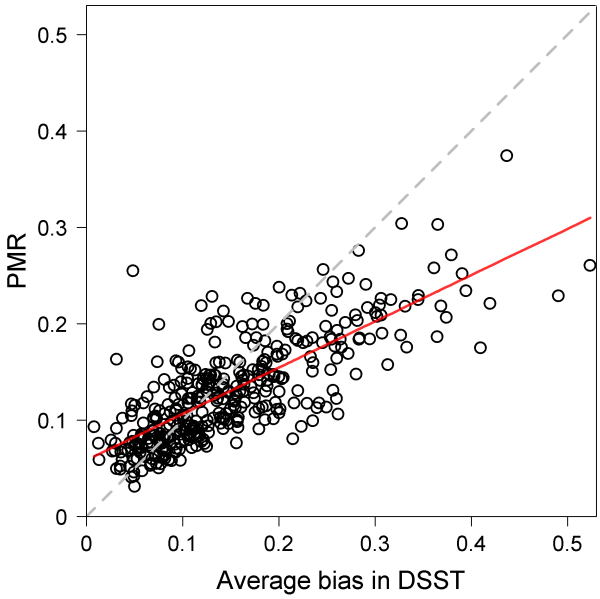


Figure 4.5: Catchment-wise PMR values against catchment-wise average of DSST bias (one point per catchment). The red line corresponds to the fitted linear model relating average bias to PMR.

4.2.4.b. Assessment of the predictive ability of the PMR for model robustness

To further investigate the link between the PMR and model robustness as measured by the DSST, we plotted the average model bias across the DSST setups for each catchment against PMR values (Figure 4.5). The reader is reminded that model bias is calculated in absolute terms, and thus there are no compensations between the averaged six model biases for each catchment. This comparison was made in order to assess the ability of the PMR to assess the variations of model robustness across the catchment set.

Figure 4.5 shows an acceptable correlation between the two indicators. Overall, the PMR seems to be a satisfactory approximation for model robustness, even if the PMR values underestimate model bias in the worst catchments (and thus overestimate model robustness). The predictive power of the PMR for model bias is further confirmed in Table 4.2 by the fair correlation values especially when model bias obtained in the DSST is averaged. DSST setups based on runoff ratio also relate better with the PMR. This indicates that the PMR has a high potential to help identify catchments

where the model struggles in particularly adverse transferability conditions. The fact that the higher correlation is obtained by averaging model biases in the DSST can be explained by the fact that the PMR is designed to be an average of model biases, and is thus more representative of an average of model errors on a few sub-periods rather than model bias on a single sub-period.
Moreover, although model biases in the DSST based on runoff ratio and PMR values have quite different spreads, the good correlation between model biases in the DSST based on runoff ratio and PMR values indicates that the PMR could be particularly useful to compare model robustness in relative terms rather than in absolute terms. Since it seems meaningful to compare PMR values reached by the model in these conditions where model errors vary drastically from one catchment to another, it is likely that the PMR could be used to relatively compare different models or different model configurations showing various behaviors.

Table 4.2: Pearson's correlation between the PMR and model biases obtained in DSST across the catchment set.

DSST setup	'dry'	'humid'	'warm'	'cold'	'unprod'	'prod'	average bias
Correlation	0.48	0.44	0.53	0.57	0.71	0.70	0.76

Therefore, the PMR appears to be a reliable estimator of model robustness, even if it exhibits tighter variations across the catchment set and thus underestimates typical biases encountered in catchments where GR4J is poorly robust. It should be noted that we did not find any particular differences in topographic or climatic properties between catchments where PMR values and DSST biases closely match and catchments where they do not.
Even if the PMR as defined in Equation (4.1) provides satisfactory results and allows a reliable assessment of model robustness without performing a DSST, the metric could have been computed in different ways. The chosen length of sub-periods or the mathematical expression of model bias are possibly important factors influencing the behavior of the metric. In the following, we conducted a sensitivity analysis with the objective of identifying the way to compute the PMR that best matches the bias that would be obtained by applying DSST procedures. Therefore, we strived to define the metric so that it corresponds as closely as possible to errors on streamflow volumes typically made by the model in adverse simulation conditions.

4.2.5. Discussion

An important element to discuss in the definition of the PMR is the length of sub-periods on which model errors are computed. Shorter sub-periods make it possible to reduce compensations between model errors, while longer sub-periods make it possible to reduce the weight of years when the model is drastically worse than others or when there are large measurement errors in the data. Sub-period length may also influence the adequation between model biases in DSST and PMR values.

4.2.5.a. Effect of sub-period length on the sensitivity of the PMR

We tested the sensitivity of the metric values to the length of sub-periods used for its computation in Figure 4.6.

The PMR metric values decrease when the sub-period length used in the computation increases. This result indicates that model interannual errors on streamflow volumes tend to compensate when the PMR is computed on longer sub-periods. Therefore, sub-period length should preferably not be too long in order to avoid loss of information about model bias across the years. This statement

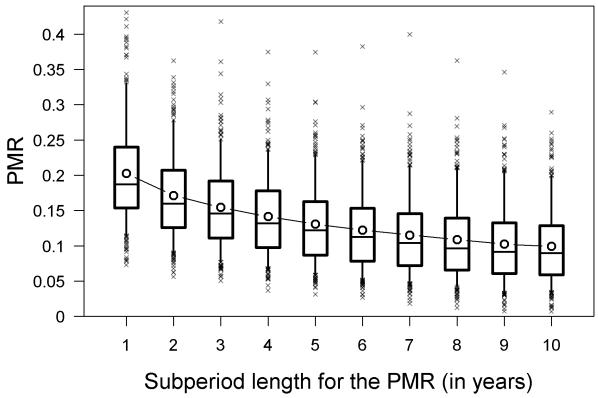


Figure 4.6: Distribution of PMR values on our catchment set for different sub-periods lengths, from 1 to 10 years. The model is calibrated on the whole time series. Black circles indicate the average PMR value.

is corroborated by the slight decrease in metric variability when sub-period length increases (standard deviation of the metric on the catchment set decreases from 7% to 5%), which suggests that differences in model robustness across the catchment set are less clear when sub-periods are too long.

4.2.5.b. Effect of sub-period length on the reliability of the PMR

Previous results show that the length of the sub-periods influences the values of the PMR. Given that the metric should relate to model biases in the DSST to be useful and reliable, it is thus necessary to verify that its correlation with DSST biases remains high when the length of the sub-periods on which the PMR is computed varies. Figure 4.7 displays the evolution of the correlation between PMR values and the DSST biases averaged for each catchment.

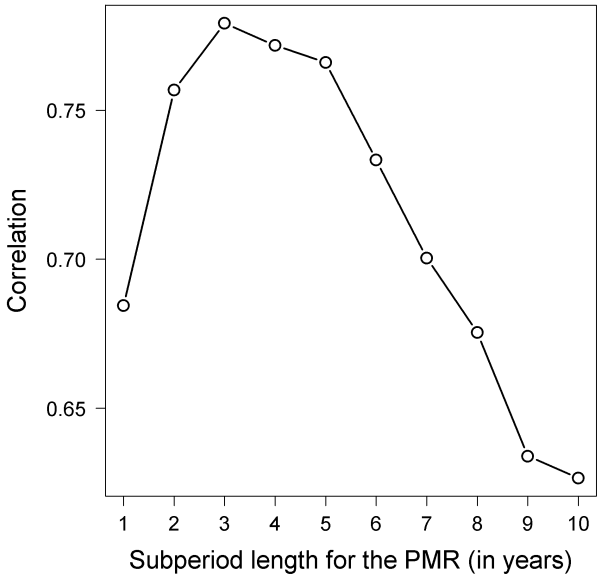


Figure 4.7: Pearson's correlations between average DSST biases and PMR values computed on varying sub-period lengths.

The correlation of the PMR with DSST biases depends in effect on the length of the sub-periods. Overall, the best scores are obtained for lengths between 2 and 5 years. Therefore, the results

described in Table 4.2 show a near-optimal situation for the reliability of the metric, although the worst correlation score, associated with 10-year sub-periods, still demonstrates a fair adequation between PMR values and DSST biases. Interestingly, computing PMR on very short sub-periods does not benefit the reliability of the metric. We suggest that adequate sub-period length for the PMR should be close to the length of the evaluation periods of the DSST, so that compared model errors relate to similar temporal scales and patterns. This issue is further discussed in Appendix A.3.

4.2.5.c. The choice of an adequate sub-periods length

Overall, the choice of the best sub-period length for the computation of the PMR must satisfy two conditions: i) be small enough to limit loss of information about model robustness, and ii) maximize correlation scores with DSST biases in order to ensure its robustness. The first condition relates to the sensitivity of the PMR to the model's actual robustness, while the second condition relates to the reliability of the PMR in different DSST setups for the evaluation of model robustness. We would suggest that the interpretability of the moving bias curve associated with the PMR accounts for a third condition for the choice of the sub-period length. In order to be interpreted easily, the curve should be smooth enough to clearly distinguish periods during which a model overestimates or underestimates observed streamflow, and thus involve long enough sub-periods. Of course, for cases where only PMR values were to be used without any analyses of the moving bias curves, this issue is incidental.
Under the conditions of our experiment, we found that lengths between 2 and 5 years were relevant to fulfil the second requirement. The sensitivity requirement would lead to computing the PMR on 2-year sub-periods; however, we acknowledge from our experience with moving bias curves that such sub-periods are too short for quick visual analyses. Therefore, we consider 3–5 years to be adequate lengths for the computation of the PMR.
However, it should be pointed out that these results are likely to be context-dependent and may have been different for other models or for another catchment set. For these reasons, the aim of the study was more the demonstration that it is possible to assess hydrological model robustness to climatic changes without performing a DSST, rather than demonstrating that the PMR is perfectly reliable and that it should substitute Split-Sample Tests. Moreover, the length of the sub-periods involved in the computation of the PMR should also reflect the particular needs of each model evaluation study.

4.2.6. Conclusions

Traditional methods to assess the robustness of hydrological models to changes in climatic conditions rely on calibration-evaluation exercises, preferably performed on climatically different periods of a time series. Although the DSST or its variants represent the most appropriate procedure one can imagine in terms of model-robustness evaluation, it is never used on models that are so complex that they need to be calibrated on all the available data. Furthermore, the DSST is based on the selection of hydro-climatic variables whose change is supposed to place the model in unfavorable conditions to perform, but whose actual link with robustness is strongly context-dependent.
In this technical note, we propose a performance metric able to evaluate model robustness on a single model realization. The so-called PMR thus does not need multiple calibrations of the model on sub-periods of the time series and can be used for any kind of hydrological model. The PMR is constructed as an indicator of the flatness of the"moving bias curve," which is a graphical representation of the temporal evolution of model bias across sliding sub-periods of the data.

The reliability of the PMR was compared with the results obtained by applying different DSST setups on GR4J, a typical conceptual model, on a dataset of 377 French catchments. We tested the predictive ability of the metric to estimate model bias obtained by transferring model parameters

from calibration periods to climatically opposite evaluation periods, for six types of hydro-climatic changes (changes in both directions of average precipitation, average temperature and average runoff ratio).
Our results show that PMR relates well to absolute model biases in the DSST, especially when these biases derived from the six DSST setups are averaged. Although the metric values do not vary much across the catchment set, this sensitivity can be enhanced by reducing the length of the sub-periods on which PMR is computed. An analysis of the correlation between the PMR and model biases in the DSST for different sub-period lengths pinpointed that the reliability of PMR was better when the metric was computed on sub-periods with lengths between 2 and 5 years. Ultimately, the need to find a balance between metric sensitivity and reliability lead us to recommend computing the PMR on 3- to 5-year sub-periods for GR4J.

Our results should encourage hydrological modelers to include the PMR as part of their panoply of evaluation metrics to judge their models or to inform model selection in climate change impact studies as it can be applied to any kind of model. Further work should examine the potential of PMR to be incorporated as a hydrological signature in multi-objective calibration procedures, and as an additional constraint on model parameters governing temporal changes in catchment response to climatic conditions.

4.3. Compléments sur le *Proxy for Model Robustness*

Dans le reste de la thèse, le PMR a été calculé à partir des sous-périodes de 5 ans et a été inclus comme un critère de performance à part entière pour les comparaisons des fonctions objectif et des formulations d'évaporation potentielle réalisées aux Chapitres 5 et 6.

Dans les paragraphes suivants, nous proposons quelques compléments aux résultats présentés dans la note technique et dans l'Annexe A. Nous avons vérifié dans un premier temps l'adéquation du PMR avec les biais de GR4J en DSST sur le sous-échantillon de bassins versants australien. Dans un second temps, nous avons mené une comparaison succincte de la robustesse des modèles GR4J, TOPMO et SimHyd à partir du PMR. Enfin, le PMR a été utilisé pour l'estimation de la robustesse d'un modèle à base physique.

4.3.1. Vérification des résultats sur l'échantillon de bassins versants australien

La Figure 4.8 montre une comparaison des biais moyens de GR4J en DSST, mis en œuvre selon les mêmes principes que dans la note technique, avec les valeurs de PMR atteintes par le modèle calé sur la période totale de chaque bassin versant australien. Cette figure est le pendant australien de la Figure 4.5.

Malgré une cohérence satisfaisante entre les valeurs de PMR et les biais du modèle en contrôle, les valeurs de PMR sont plus nettement plus faibles que les biais en DSST pour les bassins où le modèle est le moins robuste. La sous-estimation est plus marquée sur le sous-échantillon de bassins australien que sur le sous-échantillon de bassins français (pente de la régression de 0,23 en Australie contre 0,48 en France). Ce constat est toutefois explicable par la plus grande aridité des bassins australiens, ainsi que la présence d'années très sèches dans la plupart des chroniques

de débit. Lorsque le modèle est évalué sur ces années, son biais en DSST peut être très important. Le PMR au contraire est construit de manière à réduire l'influence des erreurs des années très sèches sur l'évaluation de la robustesse. Cette hypothèse est corroborée par le score de corrélation relativement faible du PMR avec les biais de GR4J en DSST « humide », c'est-à-dire calé sur une période pluvieuse et évalué sur une période sèche (Tableau 4.3). Dans l'ensemble, la corrélation du PMR avec les biais en DSST est assez forte.

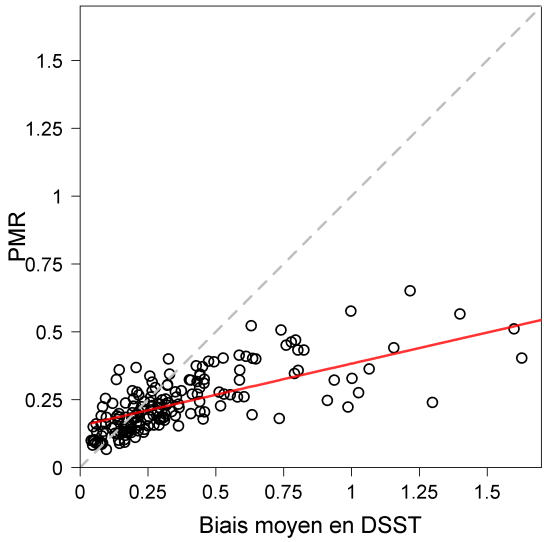


FIGURE 4.8 : Comparaison des valeurs de PMR sur les bassins versants de l'échantillon australien par rapport aux biais moyens obtenus en DSST (un point par bassin). La ligne rouge correspond à la régression linéaire du PMR sur les biais moyens.

TABLEAU 4.3 : Corrélations de Pearson entre le PMR et les biais obtenus en DSST pour le sous-échantillon de bassins versants australien.

DSST	« sec »	« humide »	« chaud »	« froid »	« improd »	« prod »	biais moyen
Corrélation	0,51	0,39	0,59	0,50	0,74	0,64	0,67

4.3.2. Comparaison de la robustesse des modèles hydrologiques utilisés dans la thèse

La Figure 4.9 présente la comparaison des valeurs de PMR calculées pour les trois modèles hydrologiques conceptuels présentés dans le Chapitre 3. Tous les modèles ont été forcés en calculant l'évaporation potentielle avec la formulation d'Oudin, et calés sur la période totale des données disponibles pour chaque bassin. La fonction objectif utilisée est le KGE calculé sur la racine carrée des débits. Les résultats sur les sous-échantillons français et australien ont été mélangés afin d'exposer une comparaison générale des modèles.

Les trois modèles hydrologiques testés démontrent des robustesses assez cohérentes sur les bassins de l'échantillon. Aucun modèle ne se distingue par des scores significativement meilleurs sur l'ensemble de l'échantillon. Les valeurs moyennes du PMR de GR4J, TOPMO et SimHyd sont respectivement 0,084, 0,086 et 0,088. De plus, les bassins où les modèles sont les plus robustes sont généralement les mêmes pour les trois modèles, et inversement pour les bassins où ils sont les

moins robustes.
Ce constat montre d'une part que des structures différentes de modèles peuvent donner lieu à des performances similaires, et d'autre part qu'aucune de ces structures ne convient pour la simulation des variations interannuelles du débit de certains bassins particulièrement difficiles à modéliser. À ce niveau d'analyse, il n'est pas possible d'identifier les raisons de ces difficultés : calage des paramètres, évaporation potentielle, précipitations ou débits observés, ou encore problème commun aux structures des trois modèles.

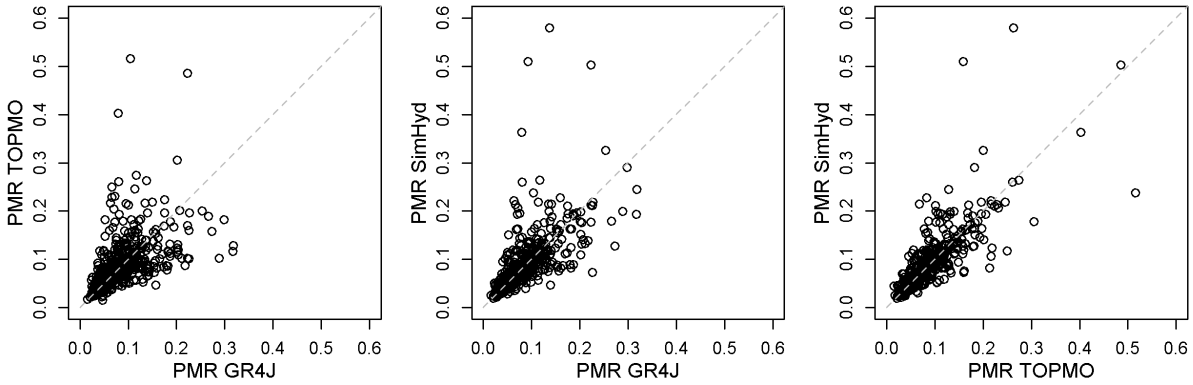


FIGURE 4.9 : Comparaison des valeurs de PMR obtenues par les modèles GR4J, TOPMO et SimHyd sur les bassins versants de l'échantillon (français et australien).

4.3.3. Estimation de la robustesse d'un modèle à base physique

Afin d'illustrer les possibilités d'application du PMR à des modèles plus complexes que les modèles conceptuels étudiés dans le cadre de la thèse, nous avons entrepris d'estimer la robustesse d'un modèle à base physique. Les simulations du modèle hydro-météorologique SIM2 (SAFRAN-ISBA-MODCOU, Habets *et al.*, 2008; Le Moigne *et al.*, 2020) nous ont été fournies par Météo-France sur un sous-ensemble de 144 bassins versants du sous-échantillon français (Figure 4.10). Ce sous-ensemble est constitué en majeure partie des bassins versants les plus grands disponibles (1750 km^2 en moyenne, contre 210 km^2 pour le reste du sous-échantillon français).

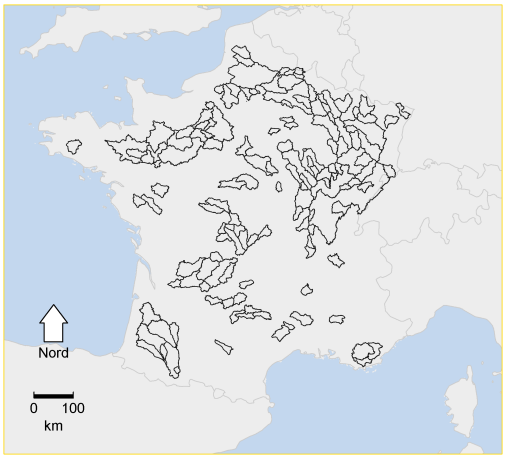


FIGURE 4.10 : Carte des 144 bassins versants de l'échantillon français pour lesquels nous disposons à la fois des simulations de GR4J et de SIM2.

Le modèle SAFRAN a déjà été décrit au Chapitre 3. Le modèle de surface SURFEX, contenant notamment le coeur de code ISBA (Noilhan et Planton, 1989; Noilhan et Mahfouf, 1996), utilise les analyses SAFRAN en entrée et calcule les bilans d'eau et d'énergie sur les mailles SAFRAN. Le bilan en eau est calculé comme l'équilibre entre les apports des précipitations et les pertes liées à l'évaporation, le ruissellement en surface, et l'infiltration dans le sol. Ces deux termes alimentent le modèle hydrogéologique MODCOU (Ledoux *et al.*, 1989; Habets, 1998), qui simule l'évolution du débit à l'exutoire des bassins versants et du niveau piézométrique des aquifères.
Les simulations de SIM2 ont été réalisées en « offline », c'est-à-dire en découplant le module météorologique du module hydrogéologique. Les données météorologiques provenant du modèle SAFRAN, les forçages utilisés pour les simulations hydrologiques de SIM2 sont exactement les mêmes que ceux utilisés pour les modèles conceptuels de la thèse sur les bassins versants français.

Les simulations de débit de SIM2 étant disponibles sur l'ensemble des chroniques SAFRAN (1958–2018), nous avons calculé les valeurs de PMR pour chaque bassin sur la partie des chroniques pour lesquelles nous avions réalisé les simulations des modèles hydrologiques conceptuels de la thèse. Les valeurs de PMR du modèle SIM2 sont comparées à celles de GR4J en Figure 4.11.

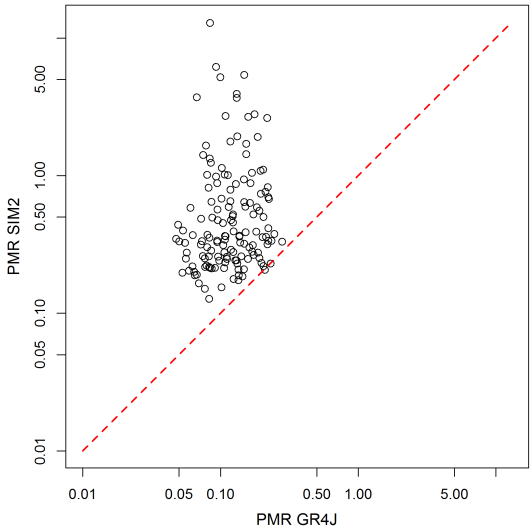


FIGURE 4.11 : Comparaison des valeurs de PMR obtenues par les modèles GR4J et SIM2 sur le sous-ensemble de 144 bassins versants du sous-échantillon de bassins français.

Il apparaît clairement que les performances de SIM2 sont moins bonnes que celles de GR4J au regard du PMR. Dans une portion importante de l'échantillon, les variations de l'erreur de SIM2 sur les sous-périodes de 5 ans dépassent la moyenne du débit observé de plus de 50%. Les débits simulés par SIM2 et GR4J étant issus de simulations en « offline » forcées par les mêmes données météorologiques, la comparaison des PMR permet *a priori* d'évaluer la robustesse des représentations des processus hydrologiques dans les deux modèles. Ce résultat montre que les modèles à base physique sont pareillement concernés par des problèmes de robustesse au même titre que les modèles conceptuels.

La Figure 4.12 montre l'exemple de quatre bassins versants. Ces bassins ont été choisis pour illustrer quatre situations distinctes :

- Les modèles GR4J et SIM2 obtiennent des valeurs de PMR comparables et faibles (bassin de la Seille à Saint-Usuge)
- Les modèles GR4J et SIM2 obtiennent des valeurs de PMR comparables et élevées (bassin

de la Sauldre à Selles-sur-Cher)

- Le modèle SIM2 obtient une valeur de PMR très supérieure à celle de GR4J, qui obtient une valeur de PMR faible (bassin de la Sarthe à Saint-Denis-d'Anjou)
- Le modèle SIM2 obtient une valeur de PMR très supérieure à celle de GR4J, qui obtient une valeur de PMR élevée (bassin de l'Aronde à Clairoix)

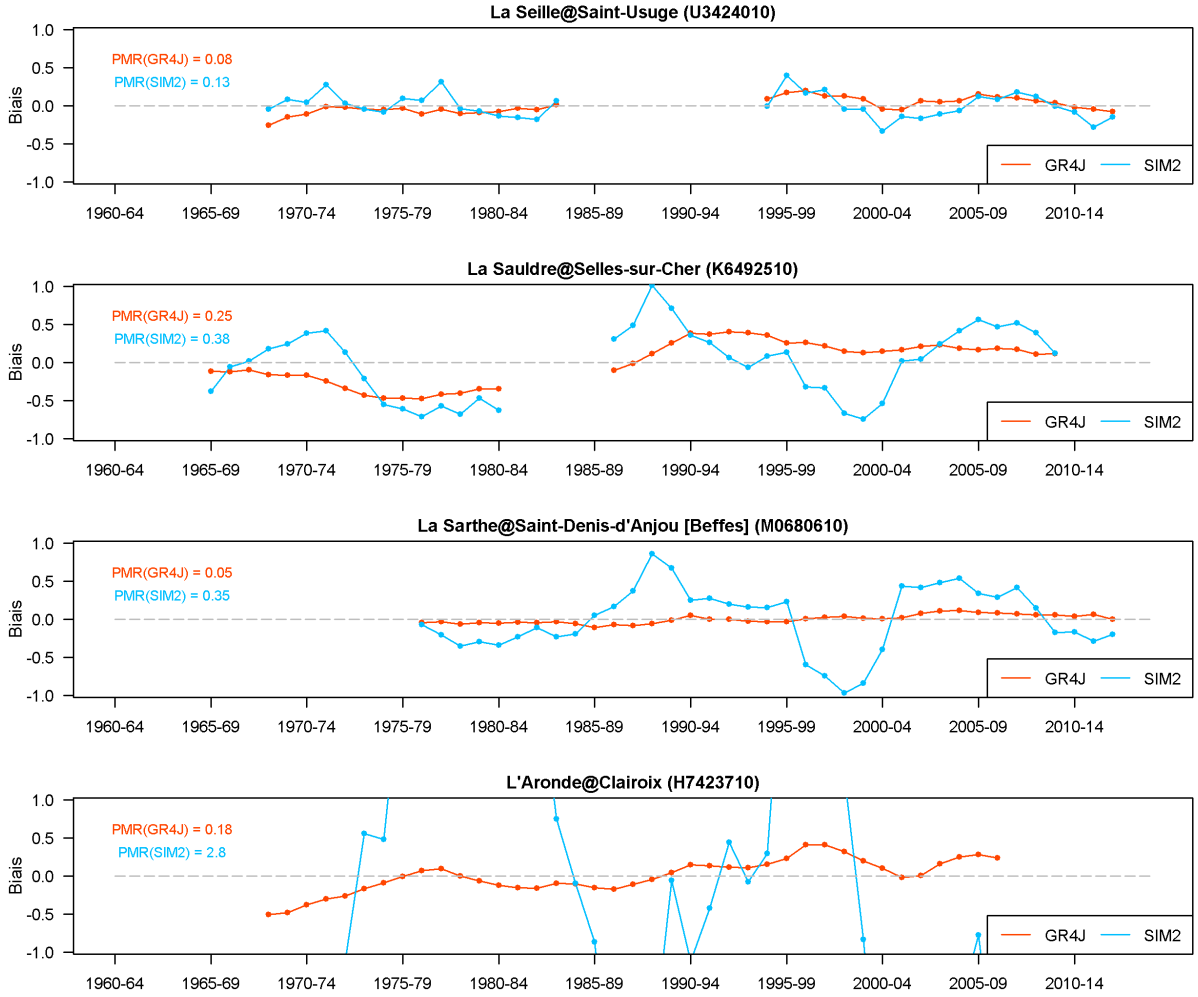


FIGURE 4.12 : Courbes des biais glissants des modèles GR4J (orange) et SIM2 (bleu) sur les bassins de la Seille à Saint-Usuge (U3424010), de la Sauldre à Selles-sur-Cher (K6492510), de la Sarthe à Saint-Denis-d'Anjou (M0680610) et de l'Aronde à Clairoix (H7423710). Les sous-périodes de 5 ans sont définies sur les années hydrologiques.

Il convient néanmoins de nuancer l'interprétation des très mauvais scores de SIM2, qui ne signifient pas forcément que le modèle n'est pas du tout robuste.

Il se trouve que sur certains bassins, le modèle SIM2 simule des écoulements très surestimés, et que le PMR du modèle est très élevé précisément sur ces bassins (Figure 4.13.a). Cela s'explique par la manière dont est calculé le PMR, à savoir le fait que la moyenne des écarts est rapportée à la moyenne du débit observé. Dans les cas de simulations très surestimées, les erreurs moyennes des simulations sur chaque sous-période peuvent varier très fortement d'une période à l'autre, d'un ordre de grandeur supérieur à celui du débit moyen observé. Puisque le PMR est justement un indicateur de la variation des erreurs d'un modèle au cours du temps, cette normalisation peut l'amener à des valeurs très hautes.

Une tentative de normalisation par la moyenne du débit simulé plutôt que par celle du débit obervé a été testée pour contourner ce problème, mais a toutefois conduit à l'effet inverse en pénalisant

lourdement les simulations les plus sous-estimées (résultats non montrés). Nous avons donc gardé la normalisation par le débit moyen observé, qui donne des résultats plus facilement interprétables. Il convient également de noter par avance que de tels biais (> 5) n'ont pas été rencontrés dans le reste de la thèse avec les modèles hydrologiques conceptuels, et donc que ce problème n'a presque pas influencé nos résultats.

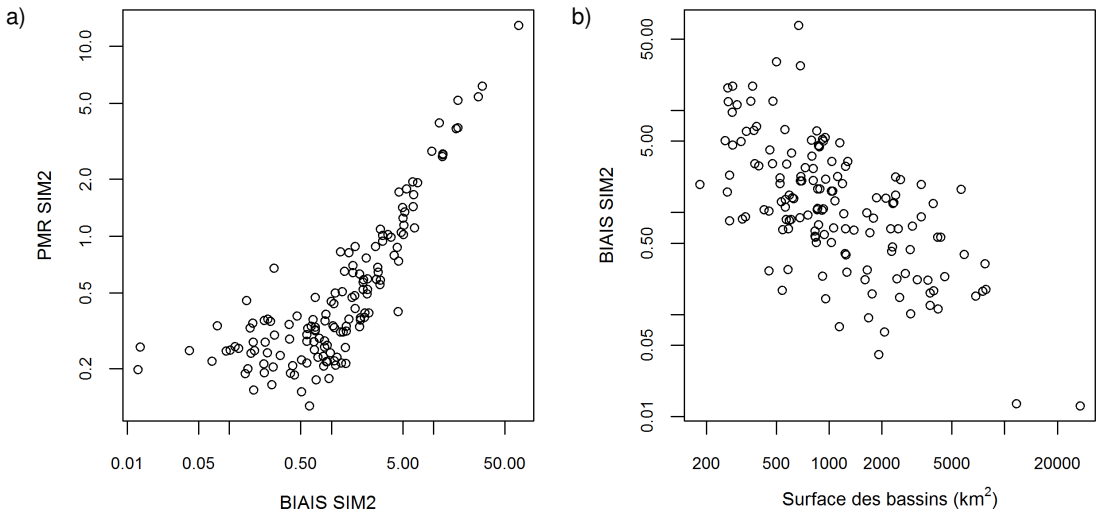


FIGURE 4.13 : a) Comparaison des biais et des valeurs de PMR obtenus par le modèle SIM2 sur le sous-ensemble de 144 bassins versants du sous-échantillon de bassins français. **b)** Comparaison des biais du modèle SIM2 à la taille des bassins étudiés.

Enfin, il est nécessaire d'ajouter que les objectifs de modélisation de SIM2 sont plus nombreux que ceux d'un modèle comme GR4J (par exemple : résolution du bilan d'énergie, simulation du niveau des nappes, cohérence spatiale sur le territoire). Et même si la multiplicité des objectifs aurait peut-être dû contribuer à rendre SIM2 plus robuste, il faut se rendre à l'évidence que ce n'est pas le cas dans les bassins de notre étude. Cependant, reconnaissons qu'une évaluation simplement indexée sur les débits à l'exutoire d'un bassin ne peut pas être considérée comme suffisante pour donner un aperçu des qualités du modèle.
Par ailleurs, au contraire des paramètres de GR4J, calés sur chaque bassin, les paramètres de SIM2 ne sont pas calés à proprement parler à partir de l'optimisation d'une fonction objectif, mais plutôt estimés à partir de calages manuels multi-objectifs sur toute la période d'observation et d'estimations *ad hoc* à partir des caractéristiques physiques des zones modélisées, sur l'ensemble de la France. Dans la mesure où SIM2 est avant tout utilisé pour la modélisation des variables de surface, comme l'humidité des sols et le manteau neigeux, les objectifs de simulation des variables hydrologiques concernent seulement l'estimation correcte des ordres de grandeur du débit à large échelle. Ces considérations contribuent à expliquer la raison pour laquelle le modèle SIM2 est fortement biaisé sur les bassins de petite taille en particulier (Figure 4.13.b), alors que ceux-ci sont déjà parmi les plus grands disponibles dans notre échantillon d'étude.

Ces résultats soulignent la difficulté de définir des méthodes d'évaluation communes aux modèles hydrologiques conceptuels et aux modèles hydrologiques à base physique, tant parce qu'ils visent des objectifs de modélisation différents que parce que les échelles spatiales auxquelles sont utilisés les premiers pour la simulation des écoulements (bassins et sous-bassins versants) contrastent avec celles étudiées avec les seconds. Cependant, l'étude proposée ici reste très succincte et une recherche approfondie pourrait conduire à identifier des références et des techniques que pourraient partager ces deux approches de modélisation. L'élaboration PMR pourrait à cet égard compter parmi les diagnostics utiles à l'atteinte de cet objectif.

4.3.4. Remarques pour la suite de la thèse

Dans la suite de la thèse, le PMR a été calculé sans le facteur 2 introduit dans la note technique à l'équation 4.1. Pour rappel, ce facteur 2 était motivé par la comparaison des ordres de grandeur des valeurs du PMR avec les biais de GR4J en DSST. En effet, le calcul de l'aire sous la courbe des biais glissants tend à diviser par deux l'écart entre les erreurs d'une période avec une autre période très contrastée, de « l'autre côté » de la courbe des biais glissants.
Le PMR a donc été calculé dans la suite selon la formule suivante :

$$PMR = \frac{1}{N} \times \sum_{i=1}^{N} \left| \left(\overline{Q}_{s,i} - \overline{Q}_{o,i} \right) - \left(\overline{Q_s} - \overline{Q_o} \right) \right| \times \frac{1}{\overline{Q_o}} \tag{4.3}$$

Dans la mesure où le PMR s'apprécie davantage de manière relative que de manière absolue, la suppression du facteur 2 ne change rien aux comparaisons menées dans la suite.

4.4. Synthèse

Stratégie d'évaluation de la robustesse en extrapolation

Ce chapitre avait pour objectif de définir les stratégies mises en œuvre pour l'évaluation de la robustesse des modèles hydrologiques.
Nous sommes d'abord partis du constat que *Differential Split-Sample Test* (DSST) présentait des défauts pour la réalisation de diagnostics de modèles, comme par exemple la nécessité de formuler des hypothèses *a priori* sur les variations climatiques défavorables aux performances des modèles.
Nous avons proposé d'adopter la procédure du *Generalized Split-Sample Test* (GSST), en l'adaptant de sorte à ce qu'une unique période de contrôle serve à évaluer les modèles en conditions de transfert de leurs paramètres. Ce choix est justifié par le besoin de comparer entre elles plusieurs réalisations d'un même modèle.
Dans toute la suite de la thèse, toute référence au GSST renvoie à la procédure décrite en Section 4.1.1.

Stratégie d'évaluation de la robustesse en interpolation

En constatant que l'emploi de procédures de calage-contrôle n'était pas possible pour les modèles complexes, dont les paramètres sont estimés avec toutes les données disponibles, nous avons ensuite proposé un critère de performance pour l'estimation de la robustesse de tout modèle hydrologique, quelle que soit sa complexité. Ce critère de performance, baptisé *Proxy for Model Robustness* (PMR), est construit à partir des courbes de biais glissants présentées en Section 2.1.1.c. Le PMR permet d'estimer en première approche la robustesse d'un modèle à partir d'une unique réalisation, son comportement du critère étant assez similaire avec les biais d'un modèle hydrologique testé en DSST.
Le PMR a été intégré aux métriques de performance utilisées dans la thèse.

Chapitre 5.

Étude de l'influence du calage sur la robustesse des modèles hydrologiques

Introduction

Ce chapitre étudie l'influence du calage sur la robustesse de trois modèles hydrologiques GR4J, TOPMO et SimHyd grâce à la mise en œuvre du GSST. Les développements menés s'articulent autour de la question de la pertinence des choix liés au calage, en excluant la structure des modèles du champ de la réflexion. Dans un premier temps, nous avons mené une comparaison de plusieurs fonctions objectif de calage au regard de la robustesse des modèles hydrologiques, dans le but d'écarter les fonctions objectif les moins adaptées pour la simulation des débits en conditions climatiques changeantes. Nous avons ensuite exploré la sensibilité de la robustesse aux choix effectués en calage, tels que la période de calage et l'algorithme utilisé. Enfin, une analyse des changements hydro-climatiques les plus difficiles à modéliser a été menée en complément des résultats du Chapitre 4 précédent, en s'inscrivant dans un cadre d'étude plus général.

5.1. Recherche d'une fonction objectif pour un calage robuste

Il convient dans un premier temps de déterminer la ou les fonctions objectif permettant de maximiser la robustesse des paramètres des modèles hydrologiques. Pour cela, nous avons fait le choix de forcer les modèles avec l'évaporation potentielle d'Oudin, dans la mesure où elle constitue la formule de référence en modélisation hydrologique conceptuelle en France. Pour rappel, les fonctions objectif comparées sont le KGE, le KGE racine (i.e. calculé sur la racine carrée des débits simulés et observés), le SKGE et le SKGE racine. Ces critères de performance ont été présentés en détail dans le Chapitre 3.
Les résultats sont d'abord présentés en agrégeant les performances des trois modèles et sur l'ensemble de l'échantillon d'étude, puis discutés en distinguant les différents modèles et les sous-échantillons de bassins versants.

5.1.1. Comportement des fonctions objectif

5.1.1.a. Comportement en calage

Une étape préliminaire à la comparaison des performances des modèles en contrôle dans le cadre du GSST est la comparaison des valeurs des fonctions objectif en calage sur la période totale, afin d'identifier des ressemblances éventuelles entre les fonctions objectif. La distribution des valeurs atteintes pour les fonctions objectif sur l'échantillon de bassins versants en calage avec chaque fonction objectif est représentée sur la Figure 5.1.
Comme attendu, les meilleures valeurs obtenues pour chaque fonction objectif sont atteintes lorsque celles-ci ont été optimisées en calage. Les différences entre les performances des modèles calés avec les quatre fonctions objectif ne sont pas les mêmes selon le critère considéré. Ainsi, elles se démarquent assez nettement au regard des valeurs de KGE, un calage avec le KGE permettant d'atteindre des performances nettement meilleures qu'avec le SKGE ou le SKGE racine. Les différences entre les distributions sont toutefois plutôt ténues au regard des valeurs de SKGE racine. On note par ailleurs qu'aucun sous-groupe de fonctions objectif ayant des comportements similaires ne semble se dessiner. Les deux critères SKGE et SKGE racine utilisés comme fonctions objectif

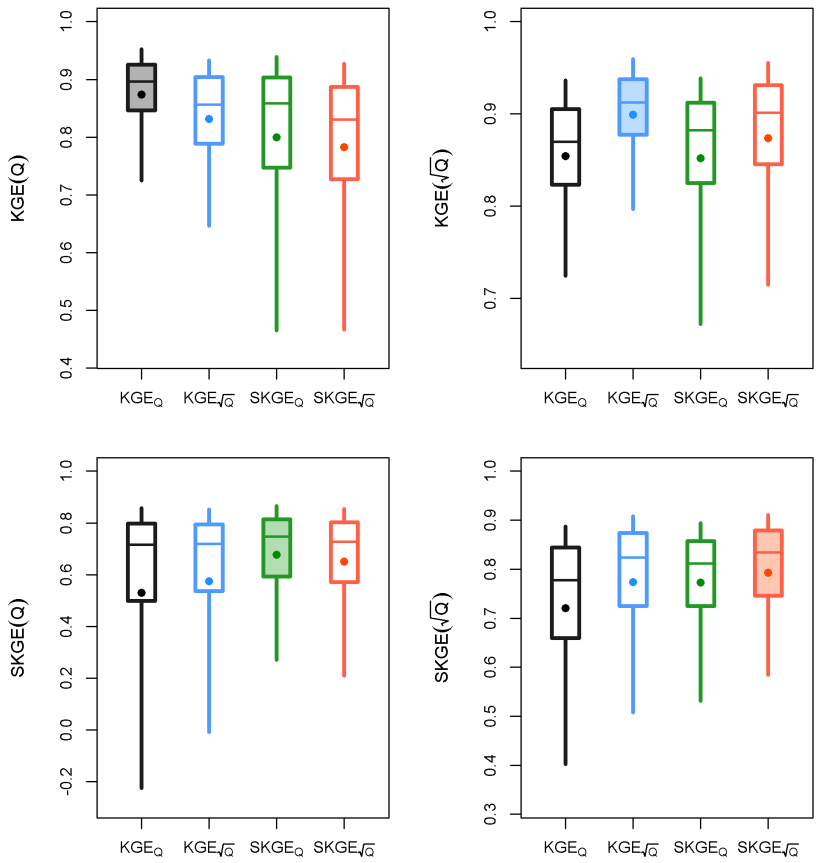


FIGURE 5.1 : Distributions des valeurs des quatre critères utilisés comme fonctions objectif (un par graphe), sans distinction des modèles, atteintes en calage avec chacune des quatre fonctions objectif sur les bassins de l'échantillon. Les distributions obtenues en calage avec le KGE, le KGE racine, le SKGE et le SKGE racine sont respectivement en noir, en bleu, en vert et en orange. Les boîtes à moustache grisées indiquent les résultats atteints en calant avec la fonction objectif représentée dans chaque graphe.

permettent par exemple d'obtenir des distributions assez proches en terme de KGE, mais assez différentes en terme de KGE racine, à l'avantage du SKGE racine. Enfin, malgré leurs similarités les fonctions objectif varient dans des domaines variés sur l'échantillon, le plus large étant le domaine de variation du SKGE et le plus restreint celui du KGE racine.

5.1.1.b. Comportement en contrôle

Les différences entre les performances moyennes par bassin obtenues en contrôle pour les quatre fonctions objectifs et les performances en calage sont comparées entre elles sur la Figure 5.2. Il convient de noter que les performances sont agrégées par bassin et donc que chaque point des courbes intègre les résultats de plusieurs modèles et plusieurs expériences de calage-contrôle. La décision de moyenner par bassin est justifiée par le besoin de donner un poids équivalent à tous les bassins, en évitant que les bassins possédant les plus longues chroniques d'observation, où plus de calages-contrôles ont été réalisés, cachent les autres. L'agrégation des performances des trois modèles répond à la nécessité d'évaluer la capacité générale des fonctions objectif à cibler des jeux de paramètres robustes. Toutefois certains phénomènes sont lissés par cette approche et il en résulte une perte d'information sur les différences entre les fonctions objectif. Ces phénomènes seront détaillés dans la suite du chapitre.

Sur le graphe, plus les distributions tendent vers des valeurs négatives, plus la dégradation de

performances est forte entre calage et contrôle sur cette période. Il apparaît globalement que le meilleur moyen de minimiser cette dégradation au regard d'un critère est de caler le modèle en utilisant ce critère comme fonction objectif. Dans l'ensemble, les relations d'ordre entre les distributions des performances ressemblent à celles constatées en calage. C'est particulièrement visible avec le KGE utilisé comme critère d'évaluation. Quelques différences se remarquent cependant entre les fonctions objectif dans les écarts aux performances en calage. Des écarts très amplifiés entre les fonctions objectif par rapport à ceux observés en calage, sont obtenus pour les valeurs de SKGE racine et pour celles de SKGE dans une moindre mesure. Les performances au regard de ces deux critères obtenues en calage avec le KGE sont par exemple nettement moins bonnes que celles obtenues en calant avec les autres fonctions objectif.
Enfin, en comparaison des autres fonctions objectif, les calages réalisés avec le KGE racine conduisent à des performances en contrôle relativement proches des valeurs en calage, quel que soit le critère d'évaluation considéré.

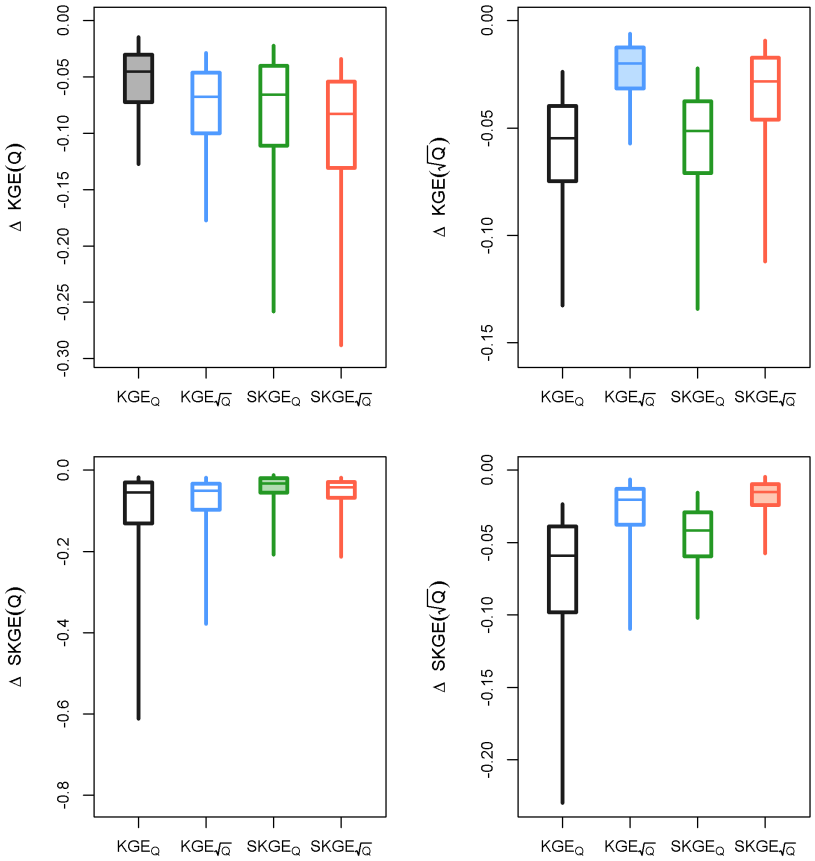


FIGURE 5.2 : Distributions des différences entre calage et contrôle des quatre critères utilisés comme fonctions objectif (un par graphe), moyennées par bassins sans distinction des modèles, dans le cadre du GSST avec chacune des quatres fonctions objectif sur les bassins de l'échantillon. Les distributions obtenues en calage avec le KGE, le KGE racine, le SKGE et le SKGE racine sont respectivement en noir, en bleu, en vert et en orange. Les boîtes à moustache grisées indiquent les résultats atteints en calant avec la fonction objectif représentée dans chaque graphe.

La comparaison des fonctions objectif est cependant limitée par la faible interprétabilité des critères utilisés comme fonctions objectif, notamment parce qu'ils sont construits comme une agrégation de sous-critères et utilisent des transformations. Dans la suite, les robustesses des modèles en calage avec les quatre fonctions objectif sont comparées au moyen des cinq critères d'évaluation présentés aux Chapitre 3 et du critère PMR introduit au Chapitre 4.

5.1.2. Comparaison des fonctions objectif à travers la robustesse des modèles hydrologiques

5.1.2.a. Performances générales

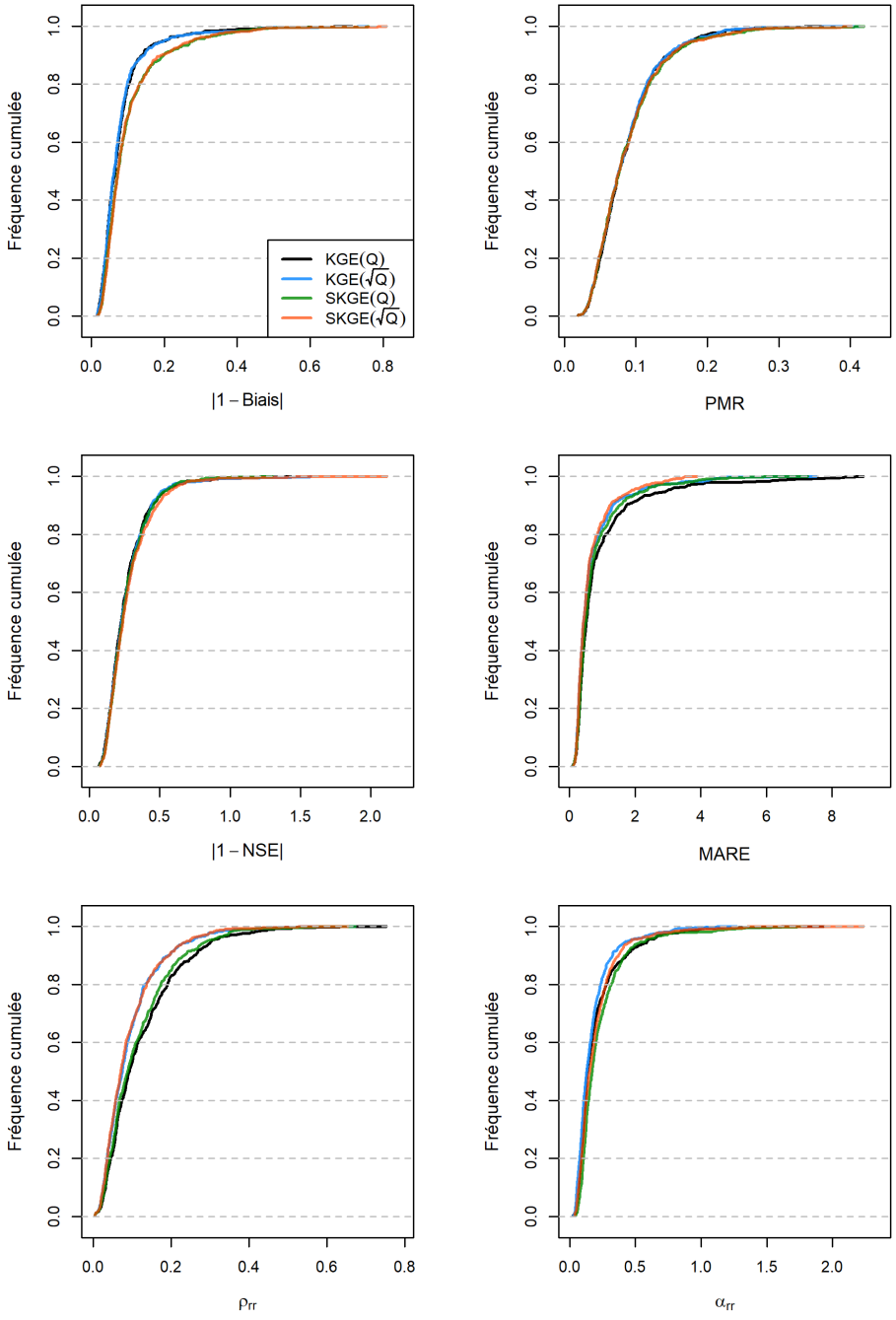


FIGURE 5.3 : Distributions des valeurs des six critères d'évaluation (un par graphe), moyennées par bassins sans distinction des modèles, atteintes en contrôle dans le cadre du GSST avec chacune des quatre fonctions objectif sur les bassins de l'échantillon. Les distributions obtenues en calage avec le KGE, le KGE racine, le SKGE et le SKGE racine sont respectivement en noir, en bleu, en vert, et en orange. Les valeurs des critères ont été calculées à partir de l'écart à la valeur optimale, de telle sorte qu'une simulation parfaite serait située tout à gauche des graphes.

Les distributions des performances des modèles, moyennées par bassin, obtenues en calant avec les différentes fonctions objectif sur les sous-périodes du GSST sont montrées en Figure 5.3. Comme précédemment, les performances des trois modèles issues des calages sur toutes les sous-périodes sont agrégées ensemble par bassin, afin d'éviter de sur-représenter les bassins ayant les plus longues chroniques.
Dans l'ensemble, il semble que ces performances soient assez peu sensibles au choix de la fonction objectif utilisée en calage, en particulier dans les 60% des bassins où les performances sont les meilleures. Les différences apparaissent en général pour les 20% des bassins où les performances sont les pires. Bien que le moyennage des performances puisse participer à l'impression d'insensibilité, celle-ci est rassurante dans la mesure où elle permet de repousser l'étape du choix de la fonction objectif au second plan dans la majorité des bassins étudiés. Néanmoins, ce choix importe dans une part non négligeable de l'échantillon et peut permettre d'améliorer la transférabilité des paramètres.

Le KGE et le KGE racine sont les fonctions objectif permettant d'obtenir les simulations les moins biaisées en comparaison du SKGE et du SKGE racine. Il convient de nuancer en remarquant que le fait que la période de contrôle ne soit pas strictement indépendante des périodes de calage favorise le KGE et le KGE racine. Ces deux critères intègrent en effet le biais sur la sous-période comme composante, au contraire du SKGE et du SKGE racine qui sont construits comme la moyenne de critères annuels, et ne ciblent donc pas autant les erreurs volumiques moyennes sur les sous-périodes de calage. Ceci se traduit par des biais des modèles plus conséquents en calage avec le SKGE et le SKGE racine (voir Figure 5.4), et donc des biais en contrôle puisque la période de contrôle contient les sous-périodes de calage.
Par ailleurs, il semble que la forme des courbes des biais glissants, ou tout au moins l'aire sous ces courbes, soit assez insensible à la fonction objectif utilisée en calage, puisque les distributions obtenues sont superposées. Ce constat est cohérent avec les résultats de Coron *et al.* (2014), qui trouvaient déjà que la forme des courbes des biais sur les sous-périodes glissantes étaient insensibles à la sous-période de calage.
Les distributions du critère de Nash montrent un désavantage très léger du SKGE racine pour la simulation des hautes eaux en contrôle. Toutefois, l'écart entre les distributions de NSE obtenues avec le KGE et le SKGE racine n'est pas statistiquement significatif, selon le test de Wilcoxon. Il l'est en revanche pour les distributions du critère MARE, où les simulations dérivées d'un calage avec le KGE sont nettement moins bonnes pour la simulation des étiages que les simulations dérivées du SKGE racine. Le test de Wilcoxon ne montre pas de différence significative entre les distributions obtenues avec le SKGE racine et le KGE racine, bien que visuellement cette dernière apparaisse légèrement moins bonne.
Enfin, le KGE racine et le SKGE racine produisent des simulations de rendement mensuel des précipitations mieux corrélées avec les observations que celles produites par le KGE ou le SKGE. Le KGE racine est également la meilleure fonction objectif au regard du critère de variabilité du rendement mensuel, les écarts avec les autres distributions étant significatifs.
En résumé, il semble donc que le KGE soit dépassé par d'autres fonctions objectif comme le KGE racine ou le SKGE racine pour la simulation des étiages et des chroniques de rendement mensuel des précipitations. De ces deux critères, le KGE racine a l'avantage de produire des débits simulés moins biaisés, bien que ce constat soit en partie dû au chevauchement des périodes de calage et de contrôle. La robustesse des modèles telle qu'estimée avec le PMR ne semble pas affectée par le choix de la fonction objectif.
Dans la suite, nous vérifions la cohérence de ces conclusions en détaillant par sous-échantillon et par modèle hydrologique.

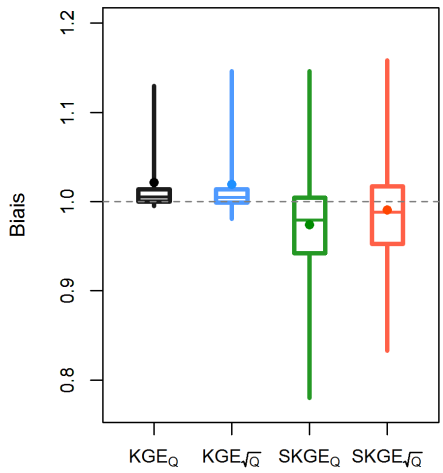


FIGURE 5.4 : Distributions des biais des modèles en calage, sans distinction des modèles, obtenus en calant avec chacune des quatre fonctions objectif. Les distributions obtenues en calage avec le KGE, le KGE racine, le SKGE et le SKGE racine sont respectivement en noir, en bleu, en vert, et en orange.

5.1.2.b. Distinction par région

Les Figures 5.5 et 5.6 illustrent les distributions des valeurs des six critères d'évaluation respectivement sur le sous-échantillon français et sur le sous-échantillon australien.
Sur le sous-échantillon français, la supériorité du KGE et du KGE racine pour le biais est visuellement moins claire bien que statistiquement significative. Le KGE racine est également mieux adapté pour la simulation des basses eaux (MARE) et pour la cohérence des chroniques de rendement avec les observations en termes de corrélation comme de variabilité.
Sur le sous-échantillon australien, on constate que les écarts de biais sont importants entre les critères KGE et les critères SKGE. Cela s'explique par le fait que les débits australiens sont en moyenne plus faibles que les débits français, ainsi que par le fait que les chroniques sont plus courtes et donc le chevauchement plus important entre périodes de calage et période d'évaluation. On note également une sensibilité accrue des distributions de PMR à la fonction objectif, bien que les écarts ne soient pas statistiquement significatifs. Les écarts en NSE ne sont également pas significatifs, à l'instar des résultats précédents sur l'échantillon entier. Les meilleures performances en étiage obtenues avec le SKGE racine sont bien visibles, bien que l'écart ne soit pas significatif avec la distribution obtenue avec le KGE racine. Enfin, de même que dans l'échantillon français, le KGE racine et le SKGE racine améliorent la simulation des chroniques de rendement mensuel. La distribution du score de variabilité du rendement obtenue avec le KGE racine se démarque également significativement des autres distributions.

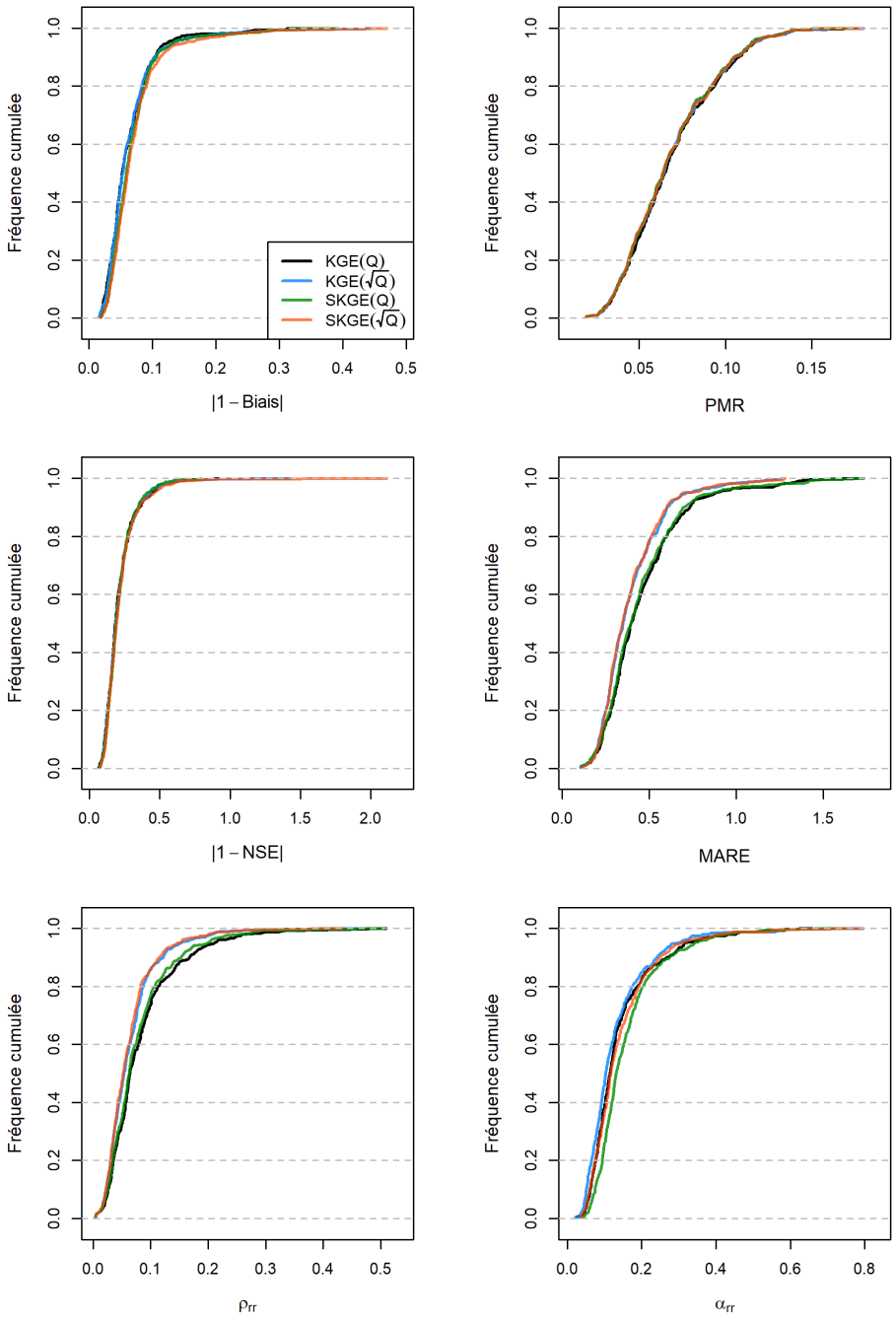


FIGURE 5.5 : Distributions des valeurs des six critères d'évaluation (un par graphe), moyennées par bassins sans distinction des modèles, atteintes en contrôle dans le cadre du GSST avec chacune des quatre fonctions objectif sur les bassins du sous-échantillon français. Les distributions obtenues en calage avec le KGE, le KGE racine, le SKGE et le SKGE racine sont respectivement en noir, en bleu, en vert, et en orange. Les valeurs des critères ont été calculées à partir de l'écart à la valeur optimale, de telle sorte qu'une simulation parfaite serait située tout à gauche des graphes.

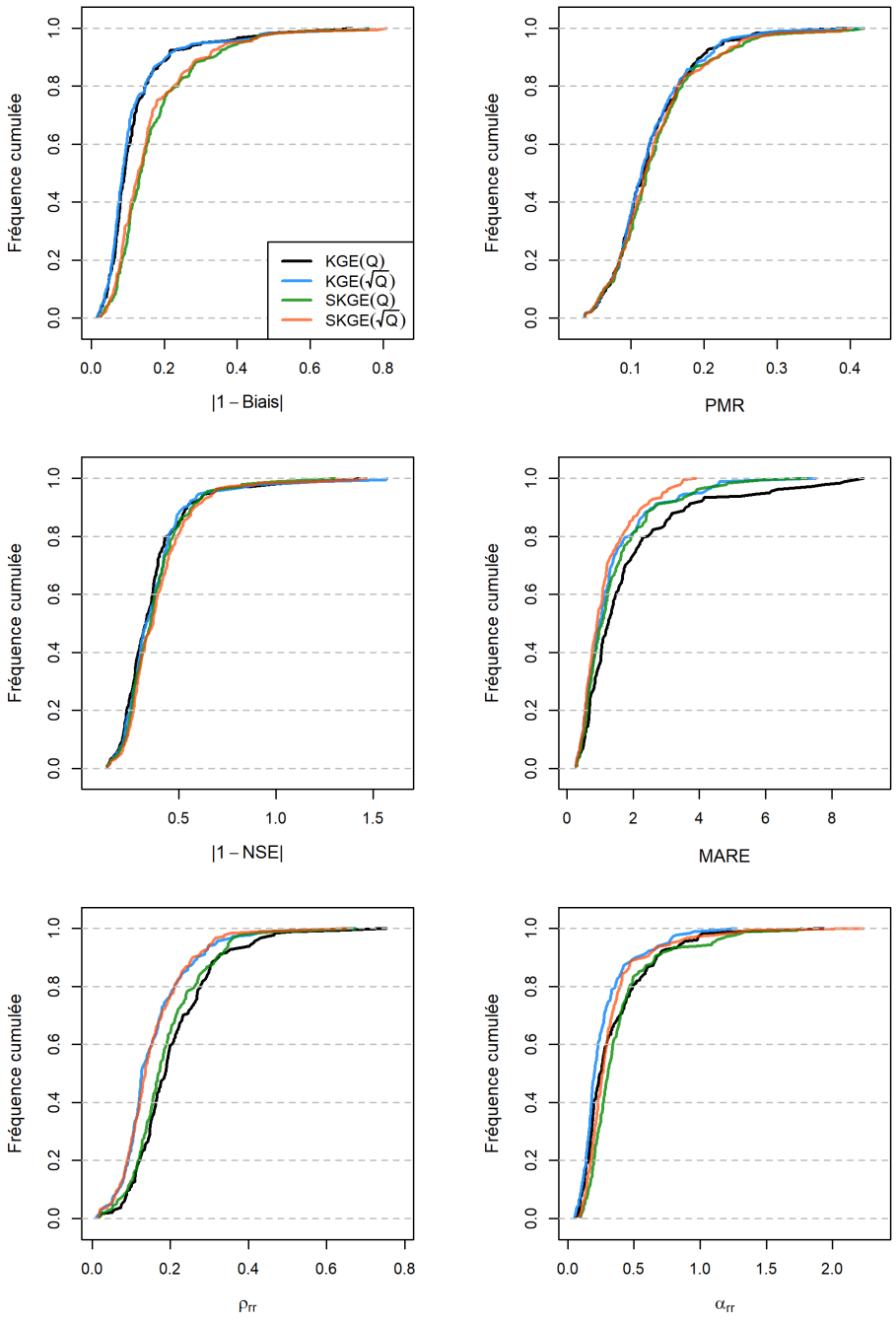


FIGURE 5.6 : Distributions des valeurs des six critères d'évaluation (un par graphe), moyennées par bassins sans distinction des modèles, atteintes en contrôle dans le cadre du GSST avec chacune des quatre fonctions objectif sur les bassins du sous-échantillon australien. Les distributions obtenues en calage avec le KGE, le KGE racine, le SKGE et le SKGE racine sont respectivement en noir, en bleu, en vert, et en orange. Les valeurs des critères ont été calculées à partir de l'écart à la valeur optimale, de telle sorte qu'une simulation parfaite serait située tout à gauche des graphes.

5.1.2.c. Distinction par modèle hydrologique

La distinction des résultats par modèle hydrologique donne des résultats globalement cohérents entre eux et similaires aux résultats agrégés, à quelques détails près. Pour ne pas alourdir l'exposé, nous avons regroupé les moyennes sur l'échantillon des performances par critère pour chaque modèle dans les Tableaux 5.1, 5.2 et 5.3.
On note dans un premier temps que pour GR4J les valeurs moyennes de MARE sont nettement améliorées en calant le modèle avec le SKGE racine par rapport au KGE racine, en comparaison avec les autres modèles. Pour la plupart des critères d'évaluation, les modèles TOPMO et SimHyd sont plus sensibles que GR4J au choix de la fonction objectif, excepté pour le biais de GR4J. Celui-ci est en effet fortement réduit par l'usage du KGE et du KGE racine en calage.

TABLEAU 5.1 : Valeurs moyennées sur l'échantillon de bassins des six critères d'évaluation de GR4J calé avec les quatre fonctions objectif. Les valeurs des critères ont été calculées à partir de l'écart à la valeur optimale théorique, de sorte que les meilleures performances sont les plus proches de zéro.

Fonction objectif	Critère d'évaluation					
	Biais	PMR	NSE	MARE	ρ_{rr}	α_{rr}
$KGE\,[Q]$	0,068	0,084	0,224	0,816	0,095	0,168
$KGE\,[\sqrt{Q}]$	0,067	0,084	0,222	0,716	0,083	0,155
$SKGE\,[Q]$	0,102	0,090	0,232	0,636	0,089	0,183
$SKGE\,[\sqrt{Q}]$	0,097	0,089	0,237	0,593	0,081	0,175

TABLEAU 5.2 : Valeurs moyennées sur l'échantillon de bassins des six critères d'évaluation de TOPMO calé avec les quatre fonctions objectif. Les valeurs des critères ont été calculées à partir de l'écart à la valeur optimale théorique, de sorte que les meilleures performances sont les plus proches de zéro.

Fonction objectif	Critère d'évaluation					
	Biais	PMR	NSE	MARE	ρ_{rr}	α_{rr}
$KGE\,[Q]$	0,091	0,088	0,253	0,683	0,107	0,217
$KGE\,[\sqrt{Q}]$	0,091	0,087	0,259	0,602	0,084	0,198
$SKGE\,[Q]$	0,099	0,088	0,255	0,670	0,101	0,239
$SKGE\,[\sqrt{Q}]$	0,101	0,088	0,275	0,574	0,083	0,207

TABLEAU 5.3 : Valeurs moyennées sur l'échantillon de bassins des six critères d'évaluation de SimHyd calé avec les quatre fonctions objectif. Les valeurs des critères ont été calculées à partir de l'écart à la valeur optimale théorique, de sorte que les meilleures performances sont les plus proches de zéro.

Fonction objectif	Critère d'évaluation					
	Biais	PMR	NSE	MARE	ρ_{rr}	α_{rr}
$KGE\,[Q]$	0,087	0,092	0,319	1,193	0,168	0,236
$KGE\,[\sqrt{Q}]$	0,090	0,090	0,329	0,788	0,117	0,174
$SKGE\,[Q]$	0,104	0,091	0,321	0,976	0,154	0,275
$SKGE\,[\sqrt{Q}]$	0,107	0,090	0,332	0,757	0,115	0,227

5.1.2.d. Discussion des résultats

Les performances en contrôle des modèles calés avec les différentes fonctions objectif sont caractérisées par des motifs similaires dans les deux régions d'étude et pour chacun des modèles étudiés. Le KGE ne présente d'avantage intéressant que dans la mesure où il permet de réduire les biais, mais fait jeu égal avec le KGE racine dans ce domaine. Pour le reste, il démontre des faiblesses pour la simulation des débits d'étiages et pour la cohérence du rendement des précipitations simulé avec le rendement observé. La difficulté des modèles calés avec le KGE à simuler convenablement l'évolution de la réponse des bassins aux précipitations en comparaison des fonctions objectifs concurrentes invite à douter de la pertinence de ce critère pour la simulation de débits dans des situations où cette réponse pourrait être amenée à varier.
Le SKGE appliqué aux débits non transformés montre également peu d'avantages comparatifs, voire des performances assez médiocres au regard de l'ensemble des aspects évalués. En revanche, la version appliquée à la racine carrée des débits démontre des qualités intéressantes. Le SKGE racine est en effet la meilleure fonction objectif pour la simulation des étiages, et est au même niveau que le KGE racine pour la corrélation des rendements mensuels. Bien que les écarts observés dans les distributions de NSE ne soient pas significatifs, le SKGE racine semble toutefois sous-performer dans ce domaine, ainsi que pour l'estimation des volumes de long terme.
Le KGE racine apparaît comme le choix le plus fiable, dans la mesure où cette fonction objectif amène les meilleures capacités de modélisation de l'évolution mensuelle du rendement des précipitations, les simulations les moins biaisées, et un bon équilibre entre performances en étiage et en crue.

Il est intéressant de constater que la transformation racine carrée est le dénominateur commun des deux meilleures fonctions objectif que sont le KGE racine et le SKGE racine. Ce constat semble indiquer que la transférabilité des paramètres des modèles hydrologiques soit davantage sensible à l'hétéroscédasticité des erreurs telles que calculées par la fonction objectif qu'à la structure même du critère. Les performances obtenues avec le SKGE ne se démarquent d'ailleurs pas particulièrement des performances obtenues avec le KGE, excepté sur les bassins australiens pour lesquels les performances en étiage sont meilleures pour le SKGE. D'après Fowler *et al.* (2018b), le SKGE est censé compenser le déséquilibre du KGE défavorable à la simulation des étiages en donnant un poids équivalent aux années sèches et aux années humides. Nous supposons que dans les bassins australiens, la structure du SKGE permet de tenir compte des années les plus sèches survenues lors de la *Millenium Drought* sans que les erreurs en étiage soient masquées par les erreurs en crue durant cette même période. Les bassins français n'ayant pour la plupart pas subi de sécheresses aussi sévères et prolongées, le SKGE ne s'y démarque pas significativement du KGE pour la simulation des étiages. Cette hypothèse n'explique toutefois pas complètement ces différences dans la mesure où Fowler *et al.* (2018b) n'intègrent pas les années de la Millenium Drought dans les données de calage et obtiennent tout de même de bonnes performances. Il est possible que des différences entre les régimes hydrologiques typiques des bassins australiens et français soient à l'origine de ces écarts. La longueur accrue des étiages dans les bassins australiens comparée à celles des bassins français favorise par exemple peut-être la prise en considération de l'erreur du modèle sur les bas débits par le SKGE.
Pour améliorer la robustesse des modèles sur l'ensemble du régime hydrologique quel que soit le contexte d'étude, il semble finalement préférable de compenser l'hétéroscédasticité à des échelles infra-annuelles, ce que permet une transfomation des débits telle que la racine carrée.

Dans la suite, nous examinons la sensibilité de la robustesse des modèles à d'autres facteurs que la fonction objectif, comme la période de calage ou l'algorithme utilisé. La comparaison des sensibilités ainsi évaluées entre les différentes fonctions objectif de calage devrait permettre d'identifier les fonctions objectif les plus robustes aux variations de ces facteurs.

5.1.3. Sensibilité de la robustesse aux choix de calage

5.1.3.a. Sensibilité au choix de la période de calage

Les résultats présentés jusqu'ici ont été agrégés par bassin pour des raisons de représentation équitable des bassins de l'échantillon. Toutefois, il est essentiel de vérifier si le comportement des modèles en contrôle dépend de la période de calage. Pour cela, nous avons calculé pour chaque bassin l'écart-type des performances en contrôle obtenues en calant les modèles hydrologiques sur les sous-périodes de calage du GSST. Ces écarts-types ont été calculés de sorte qu'ils n'intègrent pas les écarts de performance entre les trois modèles. Pour cela, pour un bassin donné et une procédure de calage-contrôle donnée du GSST, les performances des trois modèles ont été moyennées : l'écart-type est ainsi calculé sur l'échantillon des performances moyennes obtenues pour chaque procédure de calage-contrôle.
En considérant qu'il est souhaitable que les performances des modèles en contrôle ne dépendent pas des conditions de calage, un écart-type plus faible des performances est la manifestation d'une fonction objectif mieux adapté au calage de paramètres robustes.

La sensibilité des performances des modèles à la période de calage selon la fonction objectif utilisée est représentée en Figure 5.7.

Les résultats obtenus font apparaître des différences nettes entre les fonctions objectif. En comparaison des autres, le SKGE rend les modèles plus sensibles à la période sur laquelle ils sont calés, excepté pour les performances en NSE où c'est le SKGE racine qui montre le plus de variabilité. Les performances des modèles calés avec le KGE sont également assez sensibles au choix de la période de calage. Les écarts les plus évidents avec le KGE racine et le SKGE racine sont observables pour la corrélation du rendement mensuel des précipitations simulé avec le rendement observé. Dans l'ensemble, le KGE racine est la fonction objectif qui montre les performances les plus stables. Associée aux performances équilibrées constatés précédemment, cette constance confirme clairement la pertinence du KGE racine pour favoriser la robustesse des paramètres des modèles hydrologiques.
Il est intéressant de noter que la sensibilité du PMR dépend significativement de la fonction objectif utilisée, tandis que la distribution de ses valeurs moyennes ne montrait aucune différence entre les fonctions objectif. Ainsi, le KGE racine permet non seulement d'obtenir des performances améliorées et plus stables en calage-contrôle, mais il accroit également la stabilité de la robustesse telle que calculée au moyen du PMR.

Nous avons également comparé la sensibilité des trois modèles hydrologiques à la période de calage, en moyennant cette fois les performances de chaque procédure de calage-contrôle obtenues avec les quatre fonctions objectif. Les distributions des écarts-types ainsi calculés sont représentées en Figure 5.8.
Excepté son biais, les performances de GR4J sont les moins sensibles à la période de calage. La corrélation du rendement des précipitations simulé par le modèle est particulièrement peu affectée en comparaison des modèles TOPMO et SimHyd. GR4J obtenant par ailleurs les meilleures performances en moyenne pour la représentation du rendement (voir Tableaux 5.1, 5.2 et 5.3), ce résultat souligne une certaine supériorité du modèle à représenter les variations de la réponse des bassins versants aux forçages climatiques. La plus grande sensibilité du biais du modèle est compensée par ses valeurs moyennes plus faibles en comparaison des autres modèles.

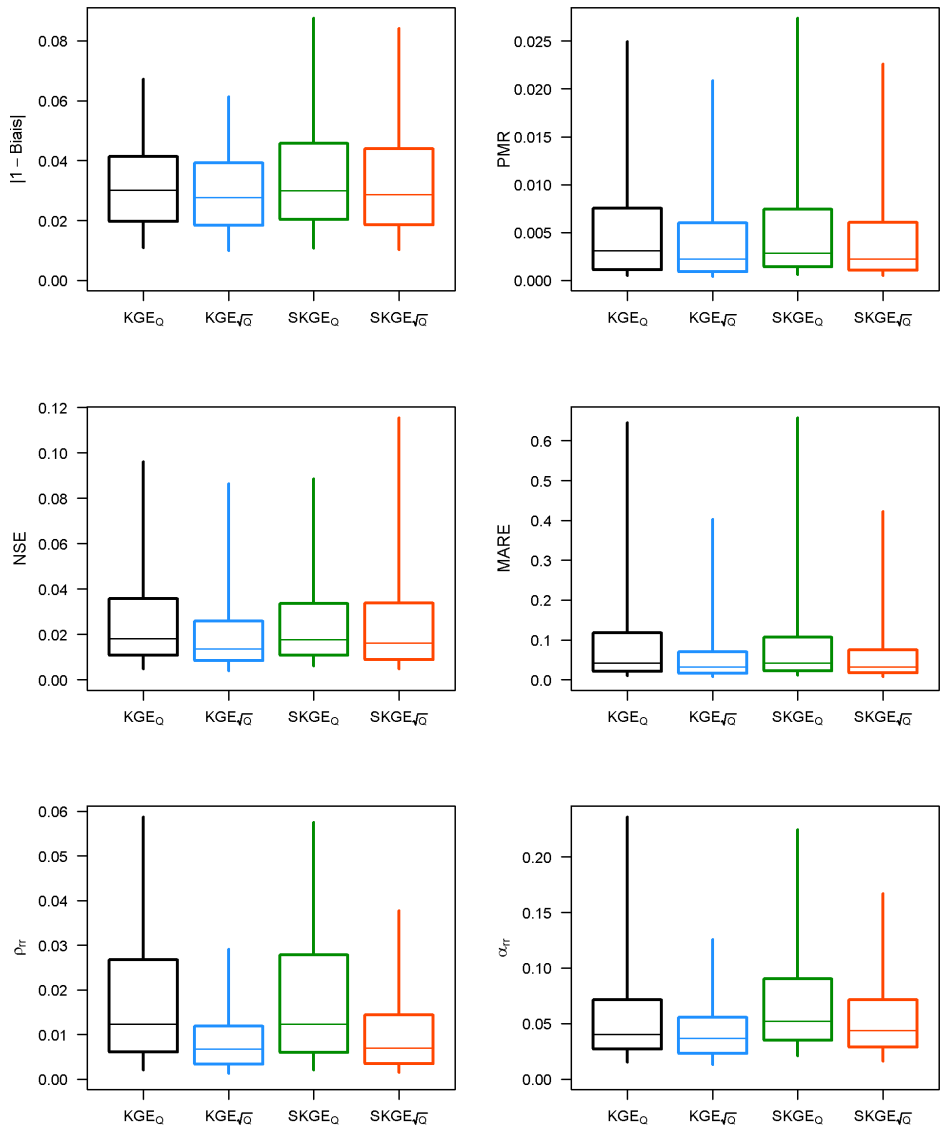


FIGURE 5.7 : Distribution des écarts-types des six critères d'évaluation (un par graphe), calculés pour chaque bassin sur l'ensemble des performances obtenues en contrôle, en calant les modèles avec chacune des quatre fonctions objectif sur les bassins de l'échantillon. Les distributions obtenues en calage avec le KGE, le KGE racine, le SKGE et le SKGE racine sont respectivement en noir, en bleu, en vert, et en orange. A chaque procédure de calage-contrôle, les performances des trois modèles ont été agrégées en une seule valeur pour ne considérer que la variabilité liée au choix de la période de calage.

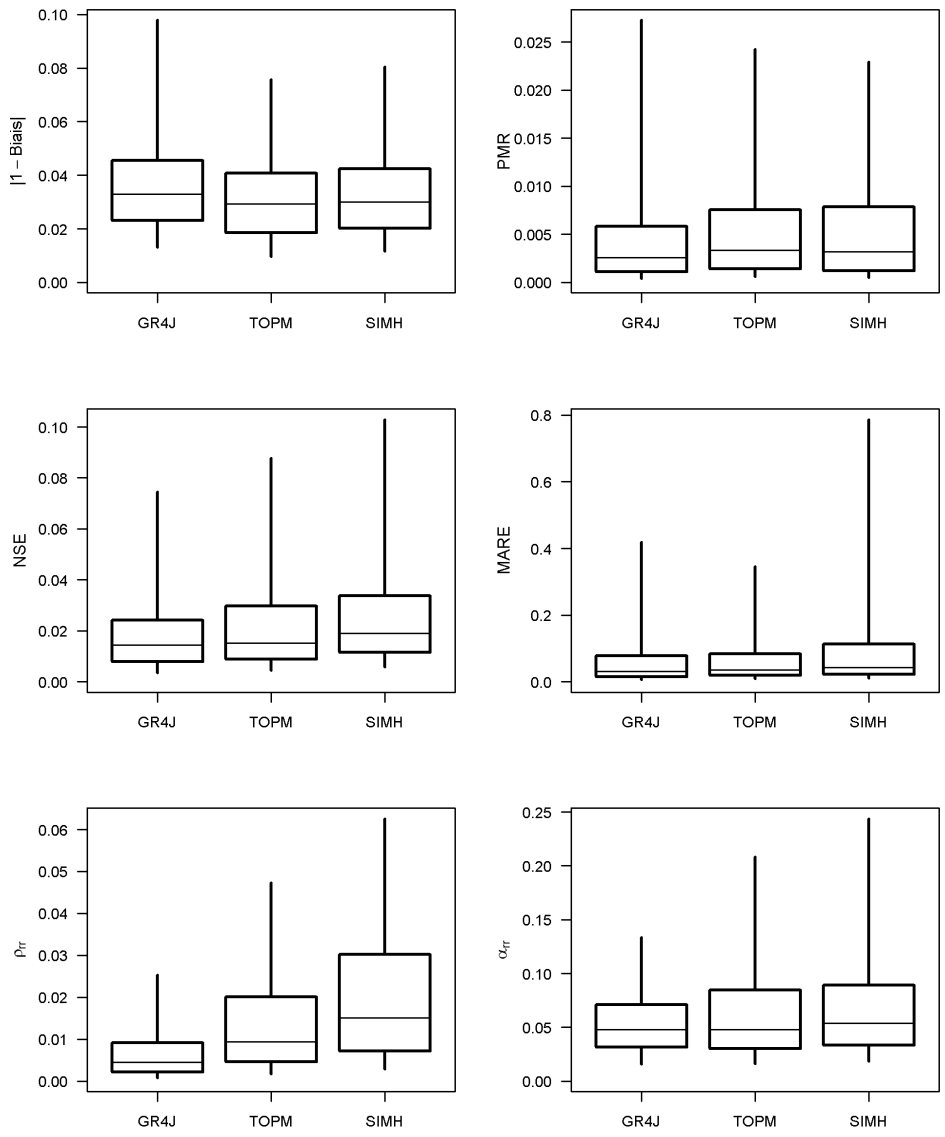


FIGURE 5.8 : Distribution des écarts-types des six critères d'évaluation (un par graphe), calculés pour chaque bassin sur l'ensemble des performances obtenues en contrôle, pour chacun des trois modèles sur les bassins de l'échantillon. A chaque procédure de calage-contrôle, les performances obtenues en calage avec les quatre fonctions objectif ont été agrégées en une seule valeur pour ne considérer que la variabilité liée au choix de la période de calage.

5.1.3.b. Sensibilité au choix de l'algorithme

Nous avons mentionné dans le Chapitre 2 que les fonctions objectif avaient rarement été comparées à l'aune de la robustesse du modèle, tandis que nous estimions que les algorithmes de calage étaient aujourd'hui suffisamment développés et validés pour considérer qu'ils ne jouaient pas un rôle important dans le constat de manque de robustesse des modèles hydrologiques. Nous avons à présent démontré que bien que la faible robustesse des modèles n'était pas vraiment imputable au choix de la fonction objectif, il était possible d'identifier des différences entre elles et d'en sélectionner une favorisant la robustesse des modèles. Il convient à présent de vérifier brièvement l'hypothèse formulée au sujet des algorithmes.
Nous avons choisi de comparer l'algorithme « Pas-à-pas » utilisé jusqu'ici à l'algorithme SCE-UA développé par Duan *et al.* (1992). SCE-UA, pour *Shuffled Complex Evolution Algorithm*, est un algorithme global assez populaire en modélisation hydrologique. Son principe repose sur l'évolution simultanée de plusieurs groupes (appelés complexes) de jeux de paramètres par la méthode des simplexes, la population de jeux de paramètres étant régulièrement redistribuée entre les groupes afin de permettre un partage d'information. L'algorithme a été paramétré avec 5 complexes comprenant chacun 9 individus. Les 45 jeux de paramètres sont initialisés sur un hypercube latin dans l'espace des 81 jeux de paramètres initiaux de la méthode pas-à-pas. Des tests de sensibilité du paramétrage de la population et de l'initialisation ont montré que ces dispositions étaient suffisantes pour garantir les meilleurs résultats pour les modèles hydrologiques utilisés dans la thèse. L'optimisation du KGE racine sur la période totale de chaque bassin est comparée entre les deux algorithmes sur la Figure 5.9 pour chacun des trois modèles.

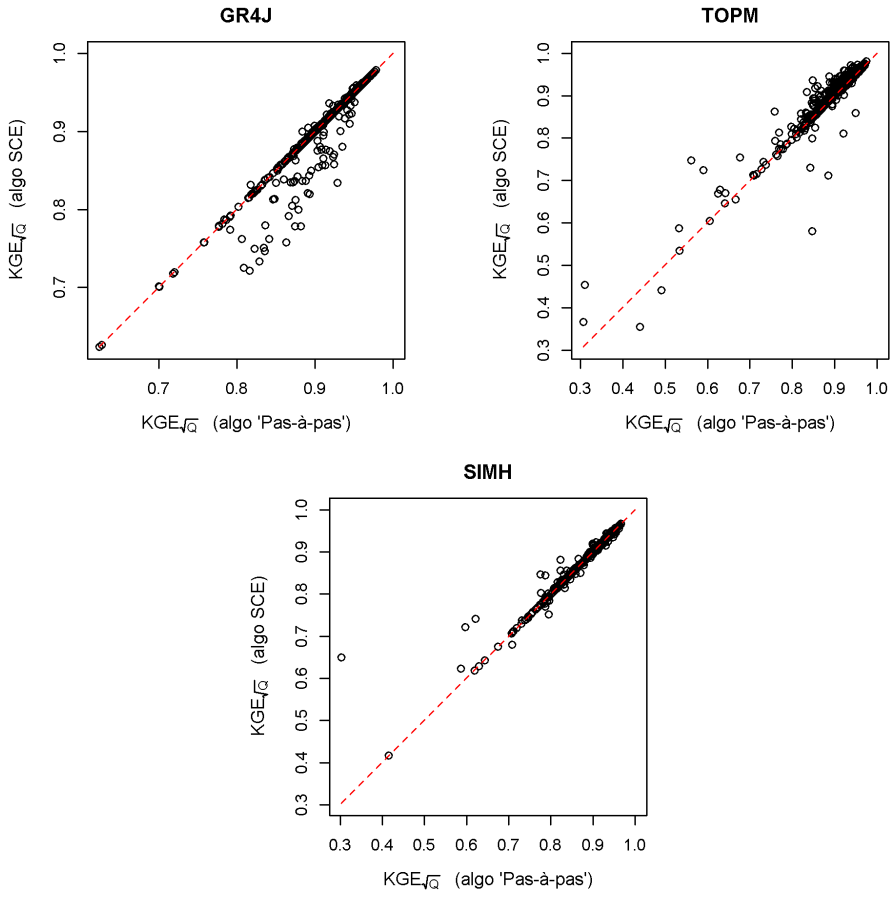


FIGURE 5.9 : Comparaison des valeurs de KGE racine atteintes en calage sur la période totale des bassins de l'échantillon avec l'algorithme « pas-à-pas » et l'algorithme SCE-UA, pour les trois modèles hydrologiques.

Il est clair que l'algorithme « pas-à-pas » est bien adapté au modèle GR4J : l'algorithme SCE-UA identifie des jeux de paramètres aux performances équivalentes ou moins élevées. Ce résultat n'est pas surprenant dans la mesure où l'algorithme « pas-à-pas » est connu pour être efficace pour le calage de GR4J (Edijatno, 1991; Nascimento, 1991; Perrin, 2000; Mathevet, 2005). Dans le cas de TOPMO en revanche, SCE-UA atteint des jeux de paramètres plus performants que la méthode « pas-à-pas » dans la plupart des bassins, excepté une poignée d'entre eux où le calage est à la faveur de la méthode « pas-à-pas ». Enfin, dans le cas de SimHyd, les deux algorithmes atteignent des jeux de paramètres aux performances équivalentes, hormis quelques rares cas à l'avantage de SCE-UA.

Pour aller plus loin dans l'estimation de l'efficacité des algorithmes, nous avons tenté d'estimer la part des bassins où le calage sur la période totale aboutissait à un jeu de paramètres sous-optimal. L'un des intérêts du GSST mis en œuvre tient dans la possibilité d'évaluer les simulations réalisées sur chaque bassin sur la même période, ce qui permet la comparaison directe de leurs performances. De ce fait, en comparant les valeurs atteintes en calage sur la période totale avec la fonction objectif de calage à l'ensemble des valeurs prises par ce critère en calant avec une autre fonction objectif ou en calant sur une sous-période, nous avons pu estimer pour chaque bassin la capacité de l'algorithme à identifier l'optimum de la fonction objectif. Ont été considérés sous-optimaux les jeux de paramètres issus du calage sur la période totale avec l'algorithme A dont la performance était dépassé par au moins un autre jeu de paramètres issu d'un calage sur une sous-période, avec une fonction objectif autre que le KGE racine, ou avec l'algorithme B. Le critère de dépassement a été fixé à un dixième de l'écart-type du KGE racine sur l'échantillon de bassins, en considérant qu'un dépassement très faible ne caractérisait pas un jeu de paramètres véritablement sous-optimal. Les résultats de cette analyse sont présentés dans le Tableau 5.4.

TABLEAU 5.4 : Pourcentage des bassins de l'échantillon où l'algorithme de calage converge vers un optimum secondaire, selon les modèles et l'algorithme utilisé.

Algorithme	Modèle		
	GR4J	TOPM	SIMH
Pas-à-pas	3,4	34,2	6,8
SCE	11,3	8,1	5,7

Il apparaît clairement que GR4J est bien calé par l'algorithme « pas-à-pas » avec seulement 3,4% de jeux de paramètres sous-optimaux. Ainsi que constaté sur la Figure 5.9, cette méthode n'est en revanche pas adaptée à TOPMO, à tel point qu'elle converge vers un optimum secondaire dans près d'un tiers des bassins de l'échantillon. L'amélioration apportée par le SCE-UA est assez forte puisqu'elle fait chuter cette part à seulement 8%. Enfin, la réussite du calage de SimHyd apparaît faiblement dépendante de l'algorithme utilisé et le taux de jeux de paramètres sous-optimaux avoisine les 6%, ce qui est convenable. Les essais menés avec les autres fonctions objectif testées précédemment avec le modèle TOPMO, dont nous ne montrons pas les résultats, sont assez similaires et contiennent les mêmes motifs. Ce résultat souligne l'importance de tenir compte de l'adéquation entre modèles et algorithmes de calage. La complexité des modèles déterminant la dimension du problème d'optimisation à résoudre en calage, il semble normal que certains algorithmes soient mieux adaptés à des modèles simples et moins à des modèles plus complexes. Le cas de la sous-optimalité de TOPMO avec l'algorithme « pas-à-pas » est problématique et invite à vérifier les conséquences sur les performances du modèle en contrôle (Figure 5.10).

Les distributions de performances obtenues avec les deux algorithmes semblent très similaires. Les différences entre les distributions dues au choix de l'algorithme sont nettement plus ténues que celles liées au choix de la fonction objectif observées précédemment. Pour quelques critères

d'évaluation néanmoins, comme le MARE, la corrélation et le rapport des variabilités du rendement mensuel, le test de significativité produit des valeurs p assez faibles. Pour le MARE et le ρ_{rr}, la probabilité que les différences soient dues au hasard est de 5%, ce que nous ne considérons pas significatif. En revanche, pour α_{rr}, cette probabilité est de 0.3% et donc significative. Pour ces trois critères, les différences sont à l'avantage de l'algorithme SCE-UA.

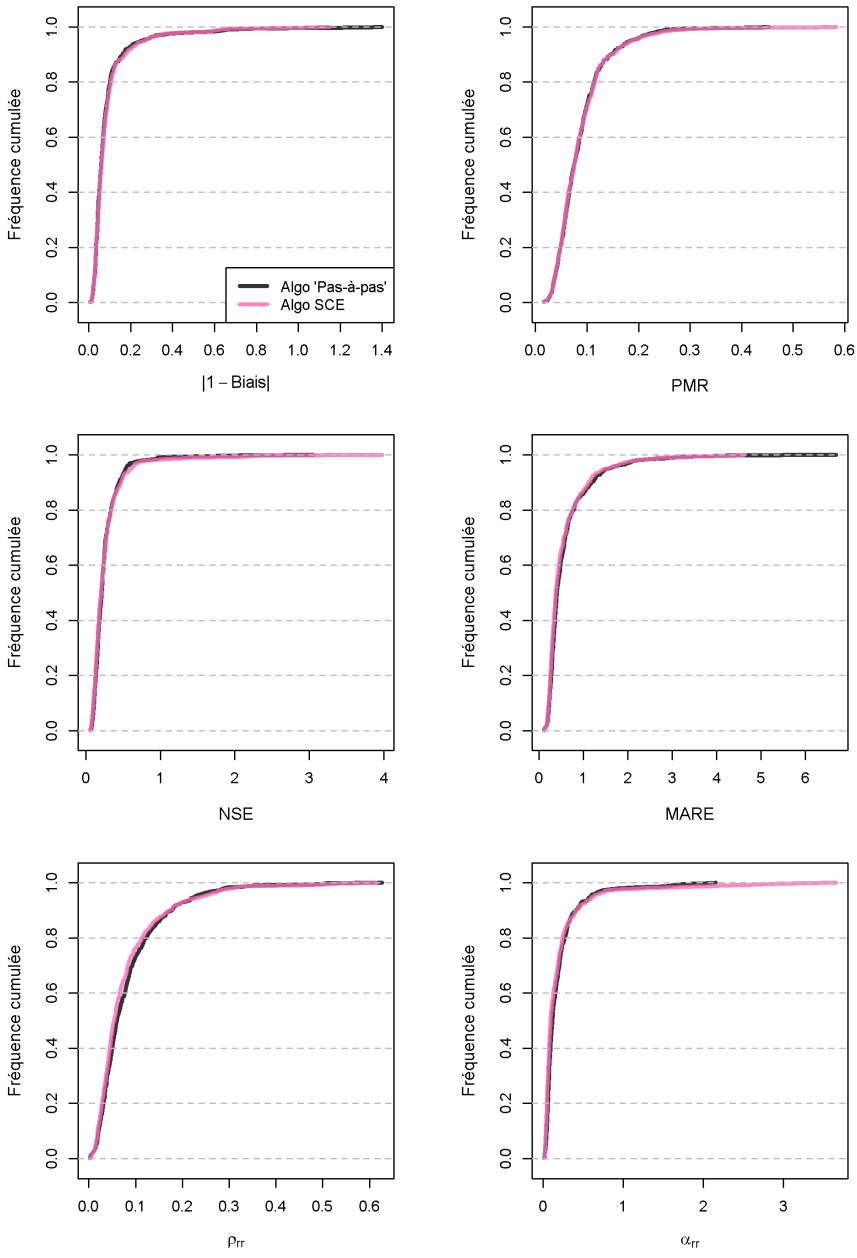


FIGURE 5.10 : Distribution des valeurs des six critères d'évaluation (un par graphe), moyennées par bassins, atteintes par le modèle TOPMO en contrôle dans le cadre du GSST en calant avec le KGE racine, sur les bassins de l'échantillon. Les distributions obtenues en calage avec l'algorithme « pas-à-pas » sont en noir, celles obtenues avec l'algorithme SCE-UA sont en rose. Les valeurs des critères ont été calculées à partir de l'écart à la valeur optimale, de telle sorte qu'une simulation parfaite serait située tout à gauche des graphes.

Dans le reste de la thèse, les calages ont tous été réalisés avec l'algorithme « pas-à-pas », bien qu'il ne soit pas optimal avec TOPMO. Les tests sur l'algorithme ont en effet été réalisés trop tardivement pendant la thèse et nous avons considéré que les écarts de performance de TOPMO étaient suffisamment légers pour ne pas justifier un recalage du modèle avec le SCE-UA, d'autant plus que

SCE-UA est moins efficient et que les calages auraient été coûteux en temps de calcul. Par ailleurs, étant donné que la plupart des développements de la thèse concerne GR4J, il est rassurant de constater que la méthode « pas-à-pas » produit des résultats très satisfaisants avec ce modèle.

5.2. Liens entre robustesse des modèles et variables hydro-climatiques

Nous avons constaté précédemment que certaines fonctions objectif étaient assez sensibles au choix de la période de calage. Le KGE racine s'est révélé à cet égard la fonction objectif la moins sensible vis-à-vis de la période de calage, et donc vis-à-vis du type de changement climatique entre calage et contrôle. Toutefois, il convient d'explorer les liens qui pourraient exister entre changements hydro-climatiques et robustesse des modèles hydrologiques, notamment pour prolonger les observations faites au Chapitre 4 avec les trois types de DSST.
Dans un premier temps, nous avons choisi de déterminer ces liens en « interpolation », c'est-à-dire en examinant les co-variations interannuelles des variables hydro-climatiques et des biais des modèles calés sur toutes les données disponibles. L'analyse est fondée sur la corrélation des erreurs annuelles avec les variations des conditions hydro-climatiques, afin de caractériser les déficiences des modèles calés avec toutes les données disponibles en formulant l'hypothèse qu'elles témoignent de leur robustesse en situation d'extrapolation climatique. Cependant, l'évolution de l'erreur annuelle n'est pas nécessairement le reflet de l'erreur moyenne sur une période de contrôle climatiquement différente des conditions de calage d'un modèle, dans la mesure où les échelles de temps considérées sont généralement plus longues dans ce cas (de quelques années à quelques décennies). Nous avons donc dans un deuxième temps analysé les résultats obtenus en « extrapolation », c'est-à-dire en confrontant les biais des modèles aux variations hydro-climatiques moyennes entre période de calage et période de contrôle, à partir des procédures de calage-contrôle du GSST.

5.2.1. Liens identifiés en interpolation

Les biais des modèles hydrologiques calés avec le KGE racine sur la période totale sont comparés à un jeu de cinq variables hydro-climatiques de référence. Les variables choisies sont le débit moyen, les précipitations moyennes, l'évaporation potentielle moyenne, le rendement moyen des pluies et l'indice d'humidité moyen. Les liens entre biais et variables sont estimées à partir des corrélations de Spearman au pas de temps annuel.
Les résultats sur les deux sous-échantillons de bassins ont été séparés afin d'identifier les liens propres à chaque échantillon, et ainsi dégager les plus robustes ne dépendant pas de la région d'étude.

5.2.1.a. Résultats sur l'échantillon de bassins français

La Figure 5.11 montre les distributions des valeurs de corrélation entre biais et variables climatiques pour l'échantillon français. De manière générale, peu de variables sont associées de manière claire à l'évolution des biais du modèle. Dans le cas de l'évaporation potentielle, le signe de la corrélation avec les erreurs du modèle est assez équivoque et amène à considérer que la variable ne possède aucun lien robuste avec elles. Les précipitations et l'indice d'humidité influencent positivement les biais mais la plupart de ces corrélations ne semblent pas significatives. Seul le rendement des précipitations affiche une relation négative robuste avec les biais : une baisse du rendement en-

traîne une surestimation des débits par tous les modèles et dans la grande majorité des bassins, et inversement une augmentation du rendement entraîne une sous-estimation des débits.

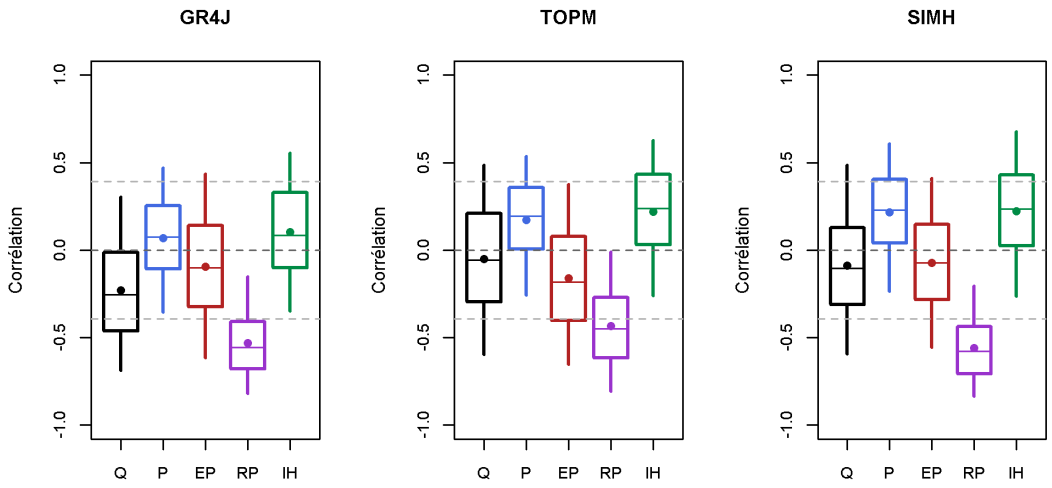


FIGURE 5.11 : Distribution des corrélations des biais annuels des modèles en calage avec les variations de cinq variables hydro-climatiques, sur le sous-échantillon français. Les corrélations calculées avec les variations annuelles du débit moyen (Q), les précipitations moyennes (P), l'évaporation potentielle moyenne (EP), le rendement des précipitations moyen (RP) et l'indice d'humidité moyen (IH) sont respectivement en noir, en bleu, en rouge, en violet, et en vert. Les lignes pointillées grises encadrant la ligne de corrélation nulle dénotent le seuil moyen à partir duquel les corrélations sont statistiquement significatives, à titre indicatif.

Le nombre de bassins où le lien entre rendement et biais est significatif confirme les observations précédentes (Tableau 5.5). GR4J et SimHyd affichent des erreurs corrélées au rendement dans près de deux tiers des bassins de l'échantillon français. Des autres variables, les variations du débit sont le plus fréquemment associées aux erreurs des modèles mais le nombre de bassins concernés reste assez faible en comparaison. Ce constat témoigne des difficultés rencontrées par les modèles lorsque le rendement des précipitations change d'année en année, ce alors que les modèles ont été calés avec l'ensemble les données disponibles.

TABLEAU 5.5 : Pourcentage des bassins du sous-échantillon français où la corrélation entre les biais annuels des modèles et les variables hydro-climatiques sont significatives à 0,1%.

Variable	Modèle hydrologique		
	GR4J	TOPMO	SIMHYD
Débit (Q)	20,9	14,1	14,7
Précipitations (P)	4,5	7,5	13,9
Évap. potentielle (EP)	14,0	16,6	9,6
Rendement précip. (RP)	59,4	41,2	62,9
Indice d'humidité (IH)	9,7	16,1	16,6

5.2.1.b. Résultats sur l'échantillon de bassins australien

La Figure 5.12 montre les distributions des valeurs de corrélation entre biais et variables hydro-climatiques pour l'échantillon australien. Les variations du débit sont dans la plupart des cas négativement corrélées aux variations de l'erreur du modèle GR4J. Cette relation est significative pour une part importante des bassins (Tableau 5.6). Ce n'est cependant pas le cas pour les autres modèles. Hormis le débit dans le cas de GR4J, aucune variable hydro-climatique ne se démarque par une

relation forte avec les erreurs des modèles. Contrairement aux résultats constatés pour le sous-échantillon français, le rendement des précipitations n'est significativement corrélé avec les biais que dans quelques dizaines de bassins, selon les modèles.

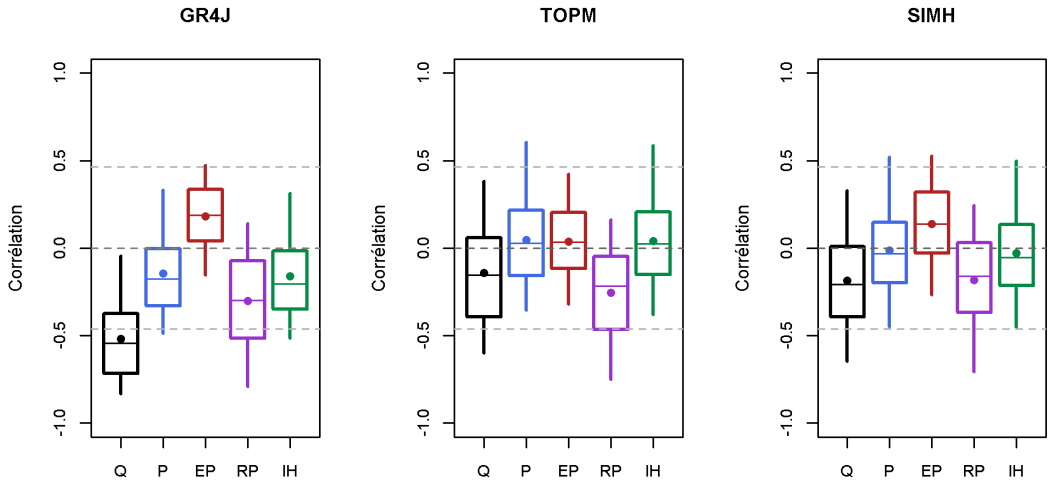


FIGURE 5.12 : Distribution des corrélations des biais annuels des modèles en calage avec les variations de cinq variables hydro-climatiques, sur le sous-échantillon australien. Les corrélations calculées avec les variations annuelles du débit moyen, les précipitations moyennes, l'évaporation potentielle moyenne, le rendement des précipitations moyen et l'indice d'humidité moyen sont respectivement en noir, en bleu, en rouge, en violet, et en vert. Les lignes pointillées grises encadrant la ligne de corrélation nulle dénotent le seuil moyen à partir duquel les corrélations sont significatives.

TABLEAU 5.6 : Part des bassins du sous-échantillon australien où la corrélation entre les biais annuels des modèles et les variables hydro-climatiques sont significatives à 0,1%.

Variable	Modèle hydrologique		
	GR4J	TOPMO	SIMHYD
Débit (Q)	42,5	6,4	8,3
Précipitations (P)	1,7	5,5	4,1
Évap. potentielle (EP)	0,3	0,6	3,6
Rendement précip. (RP)	16,3	11,0	7,2
Indice d'humidité (IH)	1,7	5,5	4,1

Les résultats ne dévoilent pas de corrélations systématiques et robustes entre les biais annuels et les variables hydro-climatiques étudiés. Les liens les plus forts concernent les débits et le rendement des précipitations, mais exhibent des différences régionales qui limitent leur portée générale.

5.2.2. Liens identifiés en extrapolation

5.2.2.a. Description de la méthode

Les liens entre biais des modèles en contrôle et changements hydro-climatiques entre calage et contrôle sont évalués à partir de régressions linéaires réalisées pour chaque bassin sur l'ensemble des biais issus des procédures de calage-contrôle du GSST. Les régressions linéaires modélisent donc l'erreur volumique des modèles en l'expliquant par les variations hydro-climatiques. À cause

de la forme de la procédure du GSST, les points sur lesquels sont calés les modèles linéaires associés à chaque variable ne sont pas strictement indépendants entre eux. En effet, les sous-périodes de calage successivent se chevauchent aux neuf dixièmes, exceptées celles séparées de plusieurs années à cause de lacunes dans les observations. La prise en considération de l'autocorrélation des points nous a donc amenés à recourir à la méthode des moindres carrés généralisés pour l'estimation des paramètres des modèles linéaires (Aitken, 1936). L'estimateur des moindres carrés généralisés est équivalent à un estimateur des moindres carrés ordinaires appliqué aux variables transformées par la transformation d'Aitken, de telle sorte que les résidus soient homoscédastiques et non autocorrélés. L'autocorrélation des résidus des variables non transformées a été fixé à 0,9.
Pour comparer l'explicativité des variations des variables hydroclimatiques étudiées pour les biais en contrôle des modèles hydrologiques, nous avons calculé le logarithme de la vraisemblance (log-vraisemblance) des paramètres des régressions linéaires. La vraisemblance est une mesure de la capacité d'un modèle (ici linéaire) à reproduire une distribution donnée de variables (ici les biais et les variations hydro-climatiques) en fonction des paramètres de ce modèle. En statistiques, les vraisemblances de deux modèles se comparent à partir du ratio de ces vraisemblances, ce qui revient à calculer une différence de leur logarithme. Pour un nombre égal de paramètres, le modèle ayant la log-vraisemblance la plus élevée est le meilleur modèle.
La Figure 5.13 illustre l'application de la méthode sur le bassin de la Meuse à Goncourt (B0220010).

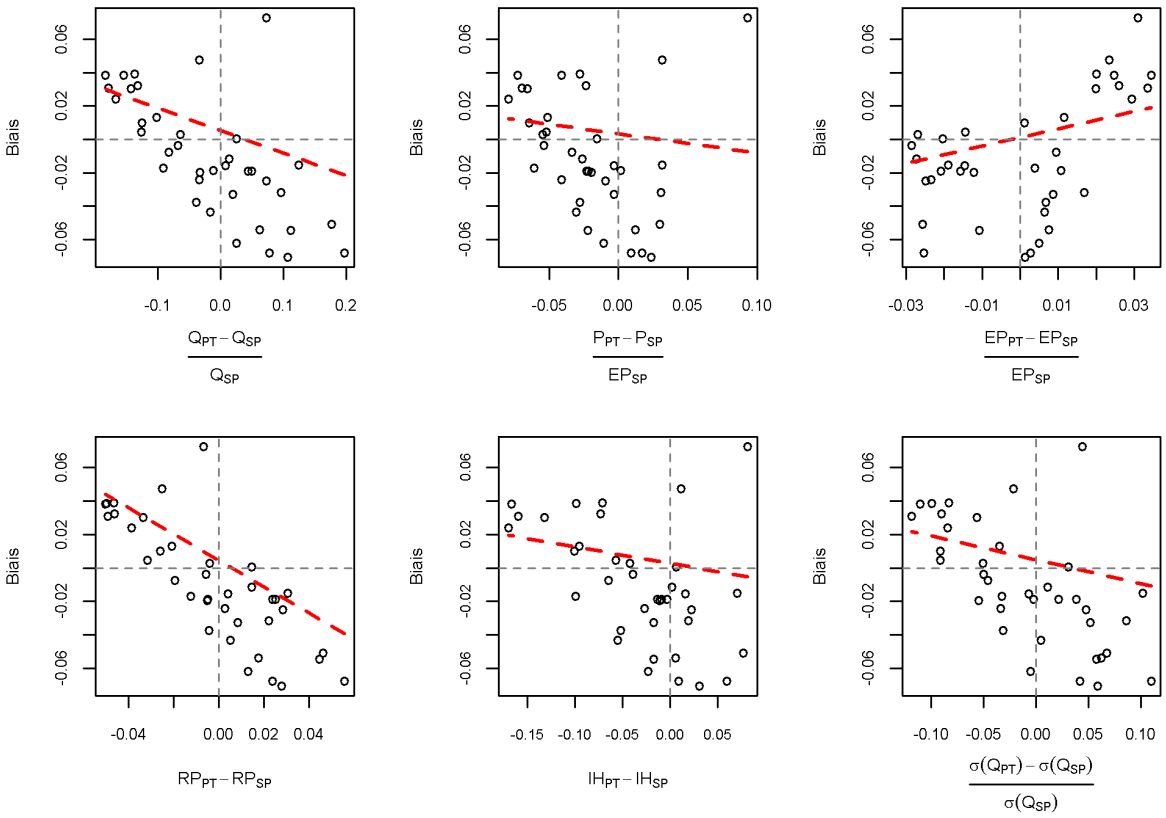


FIGURE 5.13 : Régression des biais en contrôle du modèle GR4J sur les changements hydro-climatiques relatifs entre calage et contrôle pour le bassin de la Meuse à Goncourt entre 1972 et 2018 (un graphe par variable hydro-climatique). Chaque point représente le biais sur la période totale (PT) obtenu en calant sur une des 37 sous-périodes de calage (SP). Les régressions sont calculées pour les variations du débit moyen (Q), des précipitations moyennes (P), de l'évaporation potentielle moyenne (EP), du rendement moyen des précipitations (RP) et de l'indice d'humidité moyen (IH).

Pour chaque bassin, la log-vraisemblance des paramètres de chaque régression linéaire est ensuite comparée à celle des autres afin de déterminer ceux expliquant le mieux les biais, et donc

quels changements climatiques sont les plus susceptibles de provoquer des erreurs de volumes simulés par des modèles hydrologiques. Sur l'exemple de la Meuse à Goncourt, le modèle linéaire reliant les biais aux changements de rendement des précipitations est associé à la meilleure log-vraisemblance, confirmant l'impression visuelle d'alignement des points sur le graphe en bas à gauche de la Figure 5.13. On notera que la forte autocorrélation des résidus des variables produit des régressions linéaires visuellement surprenantes, puisque l'estimateur d'Aitken favorise les points temporellement éloignés.

5.2.2.b. Résultats sur l'échantillon de bassins français

Les distributions des écarts de log-vraisemblance entre les modèles linéaires sur les bassins du sous-échantillon français sont représentées sur la Figure 5.14. Le rendement des précipitations surclasse assez nettement la plupart des autres variables dans l'explication des biais des modèles, en particulier dans le cas de GR4J. Les variations de l'évaporation potentielle ont également un pouvoir explicatif comparable à celles du rendement pour le modèle TOPMO. Quant aux variations des précipitations, le lien qu'elles entretiennent avec les erreurs des modèles est en comparaison plutôt ténu sur la plupart des bassins, alors que c'est généralement à partir de changements dans les précipitations moyennes que sont basées les expériences de DSST dans la littérature scientifique. Enfin, les variations des débits moyens et de la variabilité des débits expliquent généralement assez peu les biais. En ce qui concerne la variabilité des débits, c'est également un résultat intéressant compte tenu des conclusions souvent relevées dans les études de robustesse, qui mentionnent souvent que les modèles produisent des erreurs préoccupantes dans les bassins les plus variables. Ainsi, nos résultats indiqueraient que ces problèmes sont davantage liés à la structure même du régime hydrologique de ces bassins plutôt qu'à des variations temporelles de cette variabilité.
Dans l'ensemble, les résultats obtenus en extrapolation sont assez cohérents avec ceux obtenus précédemment en interpolation sur le sous-échantillon français.

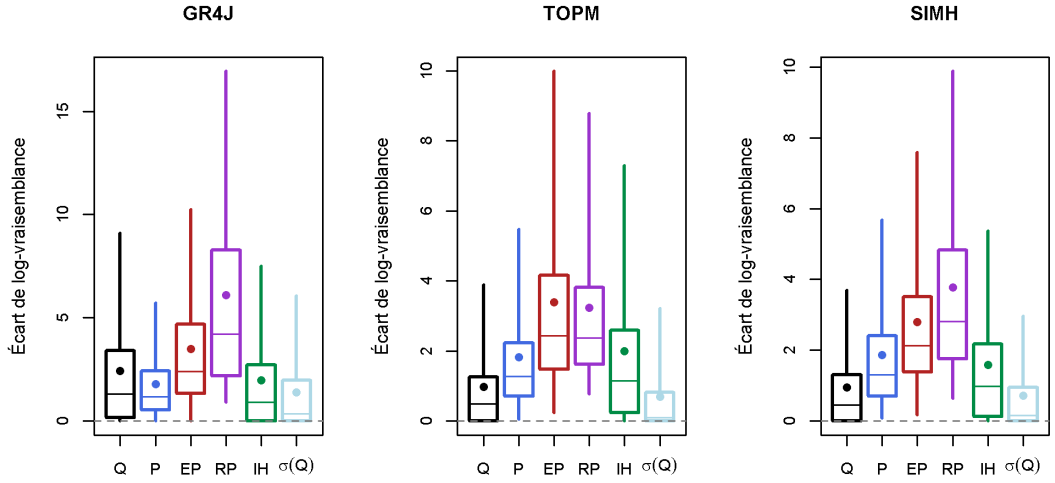


FIGURE 5.14 : Distributions des écarts de log vraisemblance entre chaque régression linéaire et la pire régression linéaire pour chaque bassin du sous-échantillon français. Les log-vraisemblances calculées pour les régressions linéaires liées aux variations relatives du débit moyen (Q), les précipitations moyennes (P), l'évaporation potentielle moyenne (EP), le rendement des précipitations moyen (RP), l'indice d'humidité moyen (IH), et l'écart-type des débits (σ(Q)) sont respectivement en noir, en bleu, en rouge, en violet, en vert, et en bleu clair.

5.2.2.c. Résultats sur l'échantillon de bassins australien

Les distributions des écarts de log-vraisemblance entre les modèles linéaires sur les bassins du sous-échantillon australien sont représentées sur la Figure 5.15. Les résultats sont très similaires à ceux obtenus sur les bassins français : le rendement des précipitations est la meilleure variable explicative des biais des modèles, accompagnée par l'évaporation potentielle dans le cas des modèles TOPMO et SimHyd. L'évaporation potentielle n'a qu'un rôle très faible sur la robustesse du modèle GR4J, dont les erreurs sont mieux expliquées par les précipitations. Ce résultat est d'ailleurs cohérent avec tous les précédents concernant la robustesse de GR4J.

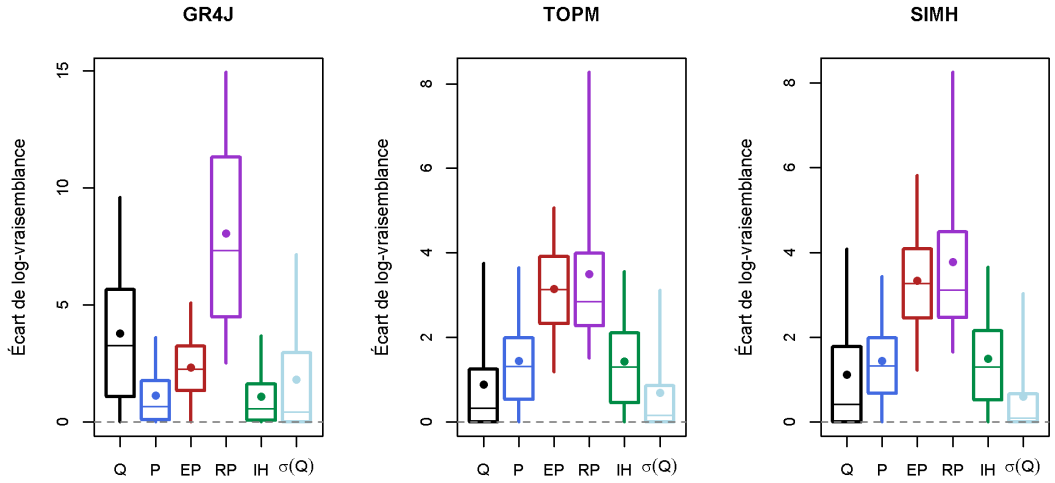


FIGURE 5.15 : Distributions des écarts de log-vraisemblance entre chaque régression linéaire et la pire régression linéaire pour chaque bassin du sous-échantillon australien. Les log-vraisemblances calculées pour les régressions linéaires liées aux variations relatives du débit moyen (Q), les précipitations moyennes (P), l'évaporation potentielle moyenne (EP), le rendement des précipitations moyen (RP), l'indice d'humidité moyen (IH), et l'écart-type des débits (σ(Q)) sont respectivement en noir, en bleu, en rouge, en violet, en vert, et en bleu clair.

5.2.2.d. Discussion des résultats

L'étude des liens entre changements hydro-climatiques et biais en contrôle des modèles hydrologiques révèle avant tout la faible importance des variations des précipitations sur la robustesse des modèles. Ce constat est d'ailleurs valable tant pour les résultats en interpolation que pour les résultats en extrapolation. Il faut bien entendu rappeler que notre analyse est fondée sur un grand échantillon de bassins versants afin d'identifier des tendances générales, et que des liens forts entre certaines variables désignées par nos résultats comme faiblement explicative peuvent être significativement liées aux erreurs de modèles hydrologiques dans des bassins particuliers. Toutefois, en supposant que nos résultats aient une portée générale, il est possible que la robustesse des modèles testés ait été surestimée dans la littérature. En effet, un certain nombre de travaux s'appuie sur des expériences de calage-contrôle entre des périodes définies à partir des différences de précipitations (e.g. Vaze *et al.*, 2010; Li *et al.*, 2012; Bisselink *et al.*, 2016). On peut supposer que la transférabilité des modèles testés aurait pu être encore moindre si ces expériences avaient été réalisées à partir du rendement des précipitations.

Le rôle joué par le rendement des précipitations dans les problèmes de robustesse des modèles hydrologiques montre la difficulté que revêt la modélisation de changements dans la relation précipitations-débit des bassins versants. Le meilleur pouvoir explicatif des modèles linéaires

mettant en relation biais et variations de cette relation indique d'ailleurs que les erreurs des modèles suivent une loi grossièrement proportionnelle à ces variations, en comparaison de celles des autres variables. Bien que les résultats ne soient pas montrés, le signe de la relation entre biais en contrôle et variations du rendement observé est très majoritairement négatif, en adéquation avec les résultats obtenus en interpolation.
Les raisons de la difficulté qu'éprouvent les modèles hydrologiques à simuler des changements du rendement du bassin ne sont pas claires à ce niveau d'analyse. Toutefois, on peut estimer que, sur un échantillon aussi large de bassins versants, des changements de relation précipitations-débit ne peuvent pas être uniquement le fait d'erreurs systématiques de mesure des débits ou des précipitations et que nombre d'entre elles sont le résultat de processus physiques. En outre, on peut noter que les liens plus ténus entre variations des erreurs volumiques annuelles et variations du rendement annuel observés en interpolation par rapport à ceux relevés en extrapolation, notamment sur l'échantillon australien, signifient peut-être que les modèles sont capables de moduler la réponse simulée du bassin versant aux forçages à des échelles temporelles interannuelles, mais que les principales difficultés concernent les changements de long terme. Ces conclusions corroborent les travaux de Fowler *et al.* (2020) sur le manque de mémoire de long terme dans les modèles hydrologiques, ainsi que ceux de Saft *et al.* (2016a).
Il convient de remarquer qu'il n'est pas aisé voire pertinent d'anticiper l'amplitude des erreurs commises par les modèles hydrologiques en projection climatique à partir du rendement des précipitations projeté, puisque celui-ci ne peut pas être calculé autrement que par des modèles hydrologiques. En plus du problème de circularité du raisonnement que cela poserait, il est possible que cela conduirait à sous-estimer les incertitudes liées aux modèles hydrologiques dans la mesure où ces modèles auraient tendance à simuler un rendement des précipitations très stable. Le recours aux estimations des modèles de surfaces continentales inclus dans les modèles de circulation générale pourraient donner une estimation exploitable pour les modèles hydrologiques de bassins versants, à condition que les échelles spatiales considérées soient comparables. Face à la difficulté d'estimer les variations futures du rendement des précipitations, il n'apparaît pas non plus envisageable de corriger *a posteriori* les simulations des modèles au regard des changements climatiques subis par les bassins simulés, comme l'avait proposé Coron (2013).

L'influence des changements d'évaporation potentielle moyenne sur les erreurs interroge, notamment au regard du faible lien trouvé avec les précipitations. Les travaux de Oudin *et al.* (2006b) indiquent en effet que les entrées d'évaporation potentielle ont moins d'influence sur les débits modélisés par les modèles que les entrées de précipitations. L'importance hétérogène du lien entre biais et changements de l'évaporation potentielle moyenne constatée pour les trois modèles invite à considérer que leur structure rend les modèles plus ou moins robustes à ces changements. Dans le cas de GR4J par exemple, la possibilité de boucler le bilan hydrique par une fonction d'échanges souterrains en plus de l'évaporation explique peut-être la plus faible sensibilité des biais du modèle aux variations temporelles de l'évaporation potentielle. Cette hypothèse sera évaluée dans le Chapitre 6 suivant.

Enfin, le choix de procéder à des régressions linéaires pour analyser le comportement des erreurs vis-à-vis des variations hydro-climatiques revient à formuler une hypothèse implicite sur la structure de ces erreurs, qu'il convient d'expliciter. En cherchant à expliquer les biais centrés autour de zéro par des variations relatives des conditions climatiques, il est attendu que des variations dans le sens positif provoquent des biais d'un certain signe (négatif dans le cas du rendement des précipitations) et que des variations négatives provoquent des erreurs du signe inverse du signe précédent (positif dans le cas du rendement des précipitations). Nous considérons que ce parti pris est justifié par la mention récurrente dans la littérature que des changements de signes opposés entraînent des erreurs de signes opposés (e.g. Coron *et al.*, 2012).
Cependant, un cadre d'analyse neutre, c'est-à-dire sans hypothèse préalable sur les signes des

erreurs des modèles, a d'ailleurs déjà été mis en place dans les travaux de Coron (2013), bien qu'appliqué à une liste de variables climatiques plus restreinte que la nôtre. Afin d'éviter de négliger le rôle de certains changements hydro-climatiques sur la robustesse des modèles, il serait intéressant de reproduire ce genre d'étude neutre avec une liste plus fournie de variables hydrologiques, moins évidentes que celles utilisées dans cette thèse ou par Coron (2013).

5.3. Synthèse

Objectifs

Les travaux menés dans ce Chapitre avaient pour objectif de déterminer l'influence des conditions de calage sur la robustesse de modèles hydrologiques. La comparaison de plusieurs fonctions objectif dans le cadre du GSST constituait la question centrale afin de garantir la meilleure transférabilité aux paramètres des modèles calés dans le reste des développements de la thèse. La pertinence de l'algorithme « pas-à-pas » a également été évaluée par rapport à un algorithme de référence en modélisation hydrologique. Enfin, les dernières analyses présentées consistaient en l'exploration des liens entre erreurs volumiques des modèles hydrologiques et changements des conditions hydro-climatiques.

Comparaison des fonctions objectif

Les quatre fonctions objectif testées, choisies pour leur simplicité et pour représenter un panel divers de critères populaires et de critères prometteurs, ont démontré une certaine diversité de comportements en calage. Aucune redondance évidente n'a été détectée entre elles, de sorte que l'on pouvait s'attendre à des capacités différenciées à estimer des jeux des paramètres robustes. Toutefois, l'analyse des performances des modèles en calage a dévoilé une sensibilité assez faible de la robustesse des modèles au choix de la fonction objectif, bien que des différences statistiquement robustes aient été dégagées.
Ainsi, le KGE racine ressort comme la fonction objectif permettant d'estimer les jeux de paramètres les plus polyvalents. La fonction objectif obtient les simulations les moins biaisées, les plus performantes en étiage, et les plus proches des variations mensuelles de la réponse des bassins versants aux précipitations. Le principal défaut du KGE s'est avéré être la simulation des étiages, conformément aux attentes, mais également une moins bonne capacité à simuler les variations du rendement des précipitations. Le SKGE n'a pas montré d'avantages particuliers comparé au KGE. Le SKGE racine, l'équivalent du SKGE appliqué à la transformation racine carrée des débits, a permis d'obtenir des performances similaires à celles obtenues avec le KGE racine. L'étude de la sensibilité des fonctions objectif au choix de la période de calage a toutefois conduit à préférer le KGE racine au SKGE racine, en privilégiant la fonction objectif garantissant la plus indépendante des conditions de calage.

Évaluation de l'algorithme de calage

L'efficacité de l'algorithme « pas-à-pas » a été comparée à celle du SCE-UA, très populaire en modélisation hydrologique, pour vérifier que les résultats précédents n'étaient pas altérés par une estimation incorrecte des jeux de paramètres optimaux. Les résultats ont montré une grande hétérogénénité entre les modèles hydrologiques. GR4J bénéficie légèrement de l'utilisation de la méthode « pas-à-pas », tandis que SimHyd a semblé relativement insensible à l'algorithme employé. Les plus grands écarts ont été observés pour le modèle TOPMO, pour lequel nous avons estimé que la méthode « pas-à-pas » était sous-optimale dans 30% des bassins modélisés, contre 8% avec l'al-

gorithme SCE-UA. Cependant, la comparaison des performances en contrôle de TOPMO obtenues avec les deux algorithmes a révélé que les implications sur la robustesse des modèles hydrologiques était assez faibles en comparaison de l'influence des fonctions objectif. Nous avons donc pris la décision de conserver l'algorithme « pas-à-pas » pour la suite des travaux de thèse.

Analyse des liens entre changements hydro-climatiques et erreurs volumiques des modèles

L'analyse des liens entre erreurs volumiques des modèles hydrologiques et variations des conditions climatiques observées a enfin démontré les difficultés importantes rencontrées par les modèles lors de changements dans la relation précipitations-débit. À l'échelle interannuelle, les biais des modèles sont très bien corrélés avec les variations de cette relation pour les bassins français. Bien que ce résultat ne soit pas retrouvé pour le sous-échantillon australien à l'échelle interannuelle, l'analyse des biais en contrôle montre une forte dépendance aux changements de long terme dans le rendement des précipitations. Ainsi, nos résultats suggèrent que la simulation des modifications durables du comportement des bassins est un enjeu majeur de la robustesse des modèles hydrologiques. Des autres variables étudiées, nous avons relevé le rôle des variations de la demande évaporatoire moyenne sur les erreurs des modèles, bien que son importance varie selon les modèles hydrologiques considérés. Le chapitre suivant est construit autour de l'analyse de l'influence du calcul de l'évaporation potentielle sur la robustesse des modèles hydrologiques.

Chapitre 6.

Étude de l'influence de l'évaporation potentielle sur la robustesse des modèles hydrologiques

Introduction

Ce chapitre présente les travaux menés pour tenter de répondre à la question suivante : peut-on identifier un modèle d'évaporation potentielle qui maximise la robustesse des modèles hydrologiques ?
Pour y répondre, nous avons dans un premier temps comparé les performances des modèles forcés par quatre formules d'évaporation potentielle différentes dans le cadre du GSST. Les quatre formules d'évaporation potentielle sont celles d'Oudin, de Morton, de Penman et de Penman-Monteith. Les résultats sont présentés de la même manière que ceux concernant les conditions de calage dans le Chapitre 5 précédent, c'est-à-dire en analysant en premier lieu les résultats agrégés puis en raffinant l'analyse des comportements régionaux et des différents modèles hydrologiques.
La deuxième partie des travaux s'articule autour d'une tentative d'identification des formules expliquant le mieux les variations interannuelles des débits à partir de régressions linéaires simples, afin de s'affranchir de l'effet de compensation provoqué par le calage des paramètres des modèles hydrologiques.

6.1. Comparaison de formules d'évaporation potentielle à travers la modélisation hydrologique

La comparaison des formules d'évaporation potentielle a dans un premier temps été réalisée par le biais de la modélisation hydrologique, à partir des mêmes critères d'évaluation que ceux utilisés au Chapitre 5. Les modèles hydrologiques ont été calés avec le KGE racine.

6.1.1. Comparaison générale des formules d'évaporation

6.1.1.a. Performances générales

Les distributions des performances des modèles forcés avec les quatre formulations d'évaporation potentielle sur les sous-périodes du GSST, moyennées par bassin, sont montrées sur la Figure 6.1 pour le sous-échantillon français et sur la Figure 6.2 pour le sous-échantillon australien. La distinction des deux sous-échantillons est justifiée par l'impossibilité de calculer les évaporations potentielles de Penman et de Penman-Monteith sur les bassins australiens. De même qu'au chapitre précédent, les performances sont agrégées par bassin de telle sorte que tous les bassins possèdent le même poids dans les distributions affichées.

Dans l'ensemble, les performances des modèles hydrologiques apparaissent peu sensibles aux formules d'évaporation utilisées, bien que de différences légères soient observables. Dans le sous-échantillon de bassins français (Figure 6.1), les distributions de biais en contrôle sont par exemple assez proches les unes des autres exceptée celle obtenue avec la formule de Penman-Monteith, pour laquelle le quart supérieur des biais est nettement moins satisfaisant que les autres. En comparaison de la sensibilité aux fonctions objectif de la robustesse des modèles hydrologiques, telle que calculée avec le PMR (cf. Figures 5.3 et 5.4), la sensibilité aux formules d'évaporation semble plus élevée. Les formules d'Oudin et de Morton affichent à ce titre les meilleurs résultats. Néanmoins, le test de Wilcoxon n'aboutit qu'à des valeurs p de 6% entre les

distributions liées à la formule de Morton et de Penman-Monteith, ce qui n'est pas suffisant pour conclure à une différence statistiquement robuste au vu de la taille de l'échantillon. De même, les distributions de valeurs de NSE diffèrent, à l'avantage de la formule d'Oudin, de sorte que la valeur p associée à l'hypothèse nulle entre les performances obtenues avec la formule d'Oudin (respectivement Morton) et celles obtenues avec la formule de Penman-Monteith atteint 2% (respectivement 6%). Le constat est identique pour les performances en étiage évaluées avec le MARE.

En ce qui concerne la capacité des modèles à simuler les variations mensuelles du rendement observé des précipitations, la hiérarchie des formules diverge selon que l'on considère la corrélation ou le rapport des variabilités. Les formules d'Oudin et Morton atteignent les meilleures valeurs de corrélation et sont significativement meilleures que celle de Penman. La formule de Morton permet de reproduire le mieux la variance du rendement, suivie par celles d'Oudin et de Penman. Les pires performances sont obtenues avec la formule de Penman-Monteith. L'écart entre la distribution des valeurs atteintes avec la formule de Morton et celles obtenues avec celle de Penman-Monteith est statistiquement significatif. Celui entre les formules de Morton et d'Oudin a une valeur p de 9%, ce qui n'est pas suffisant pour conclure que la formule Morton permet d'obtenir les meilleures performances.
Il ressort nettement de ces résultats que la formule de Penman-Monteith amène les modèles hydrologiques à sous-performer en comparaison des autres, en particulier par rapport aux formules de Morton et d'Oudin. Ces deux formules semblent tout aussi appropriées l'une que l'autre pour la modélisation hydrologique.

L'analyse des performances sur le sous-échantillon de bassins autraliens (Figure 6.2) montre toutefois que la formule de Morton y est mieux adaptée que celle d'Oudin. Les distributions obtenues avec la formule de Morton sont visuellement presque toutes meilleures que celles obtenues avec la formule d'Oudin. Cette impression n'est néanmoins confirmée par l'analyse statistique que pour le biais des modèles, leur valeur de NSE et le rapport des variabilités du rendement mensuel des précipitations. Une discussion de ces résultats est proposée aux Sections 6.1.2 et 6.1.3.

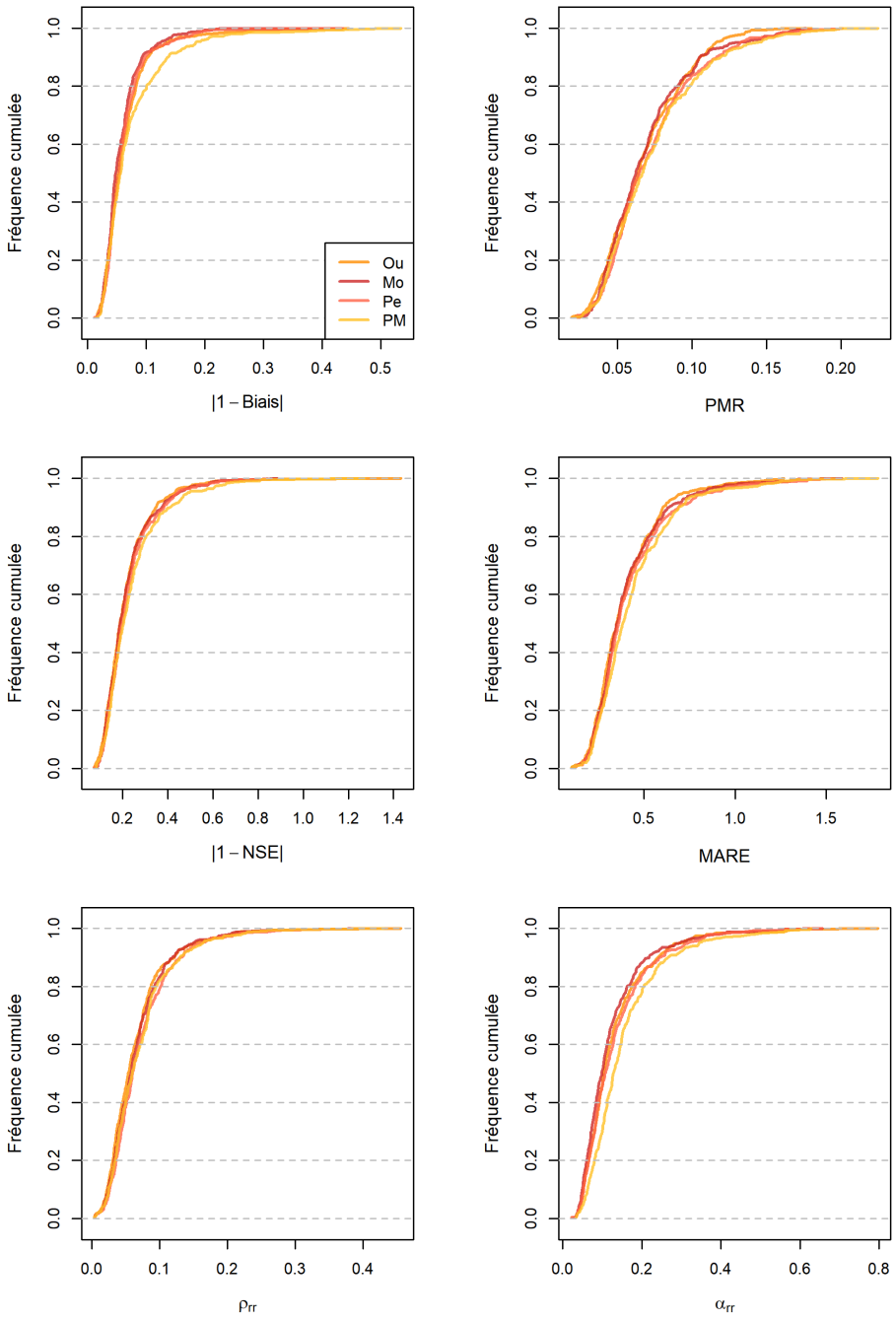


FIGURE 6.1 : Distributions des valeurs des six critères d'évaluation (un par graphe), moyennées par bassins sans distinction des modèles, atteintes en contrôle dans le cadre du GSST avec chacune des quatre formules d'évaporation potentielle sur les bassins du sous-échantillon français. Les couleurs du jaune au rouge foncé ont été choisies de telle sorte à indiquer les évaporations moyennes dans l'ordre croissant des valeurs moyennes d'évaporation potentielle sur le sous-échantillon. Les valeurs des critères ont été calculées à partir de l'écart à la valeur optimale, de telle sorte qu'une simulation parfaite serait située tout à gauche des graphes.

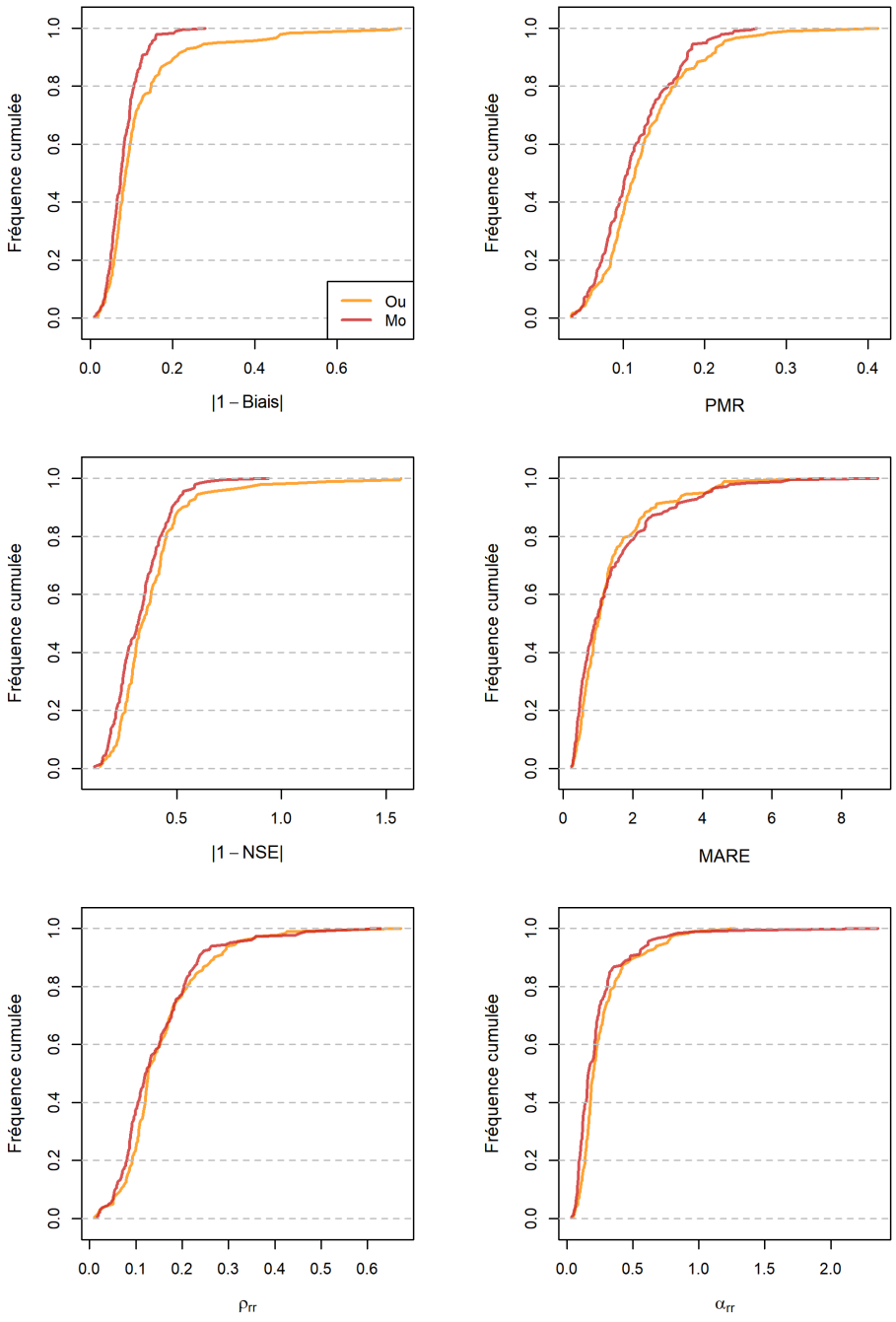


FIGURE 6.2 : Distribution des valeurs des six critères d'évaluation (un par graphe), moyennées par bassins sans distinction des modèles, atteintes en contrôle dans le cadre du GSST avec les deux formules d'évaporation potentielle sur les bassins du sous-échantillon australien. Les couleurs orange et rouge foncé indiquent respectivement les évaporations d'Oudin et de Morton. Les valeurs des critères ont été calculées à partir de l'écart à la valeur optimale, de telle sorte qu'une simulation parfaite serait située tout à gauche des graphes.

6.1.1.b. Stabilité des performances des modèles hydrologiques selon la formule d'évaporation potentielle

L'analyse des performances des modèles hydrologiques moyennées par bassin a montré l'existence de différences entre les formules d'évaporation potentielle. L'analyse est complétée ici par l'évaluation de la variabilité des performances des modèles en fonction de l'évaporation potentielle dans chaque bassin, c'est-à-dire l'évaluation de la stabilité des performances des modèles selon la période de calage (Section 5.1.3.a). En un sens, cela revient à estimer la cohérence temporelle de la pertinence des formules d'évaporation potentielle. En effet, certaines formules peuvent être bien adaptées sous certaines conditions climatiques et moins dans d'autres. Il est ainsi possible que le domaine de validité de certaines formules dépende par exemple de l'écart moyen des températures journalières, des conditions moyennes de vent ou d'humidité, ou encore de la couverture nuageuse moyenne, et que ces caractéristiques varient dans le temps. À travers l'analyse de la stabilité des performances des modèles hydrologiques, nous avons cherché à identifier les formules d'évaporation potentielle qui sont les plus robustes face à ces changements de conditions pour la modélisation hydrologique. Nous pourrions résumer en indiquant que les résultats précédents concernaient plutôt la cohérence spatiale des formules d'évaporation, tandis que ceux de ce paragraphe concernent leur cohérence temporelle.

Les distributions d'écarts-types des performances sur les bassins du sous-échantillon français sont visibles sur la Figure 6.3. On constate en premier lieu que la stabilité des performances peut être relativement sensible au choix de la formule d'évaporation potentielle, et que les résultats sont assez hétérogènes suivant le critère d'évaluation considéré. La formule d'Oudin est à ce titre significativement meilleure que les autres formules pour la variabilité des valeurs de PMR, en particulier par rapport à la formule de Penman-Monteith. La formule d'Oudin, avec celle de Morton, est également avantagée pour la corrélation du rendement des précipitations. Dans les deux cas précédents, cette stabilité s'accompagne de performances moyennes également supérieures, ce qui renforce le constat de qualité de la formule d'Oudin pour la simulation des variations interannuelles des débits (PMR) et pour celles du rendement mensuel des précipitations.
Toutefois, certaines formules aboutissant à des performances plutôt médiocres par rapport aux autres sont tout de même caractérisées par une meilleure stabilité de ces performances. C'est le cas par exemple de la formule de Penman-Monteith pour le biais des modèles, pour la simulation des étiages, et à moindre égard pour la simulation des crues. Pour expliquer ce phénomène, on peut supposer au regard des descriptions de la Section 3.3.4.a que la formule de Penman-Monteith produit une demande évaporatoire trop faible en été, ce qui conduit à des simulations biaisées, mais qu'en revanche les variations interannuelles qu'elle modélise sont proches de celles de la demande évaporatoire réelle. Une hypothèse plus pessimiste pourrait être que la formule de Penman-Monteith est fortement inadaptée à la modélisation hydrologique sur une partie des bassins du sous-échantillon, de telle sorte que les paramètres des modèles hydrologiques ne puissent compenser les erreurs dans les forçages qu'en convergeant vers une certaine configuration, qui varierait très peu d'une période de calage à une autre. Le rôle des données d'entrée dans la formule de Penman-Monteith et les raisons pour lesquelles l'évaporation potentielle calculée avec la formule de Penman-Monteith est faible en période estivale ont été examinés en détail dans la suite du chapitre.

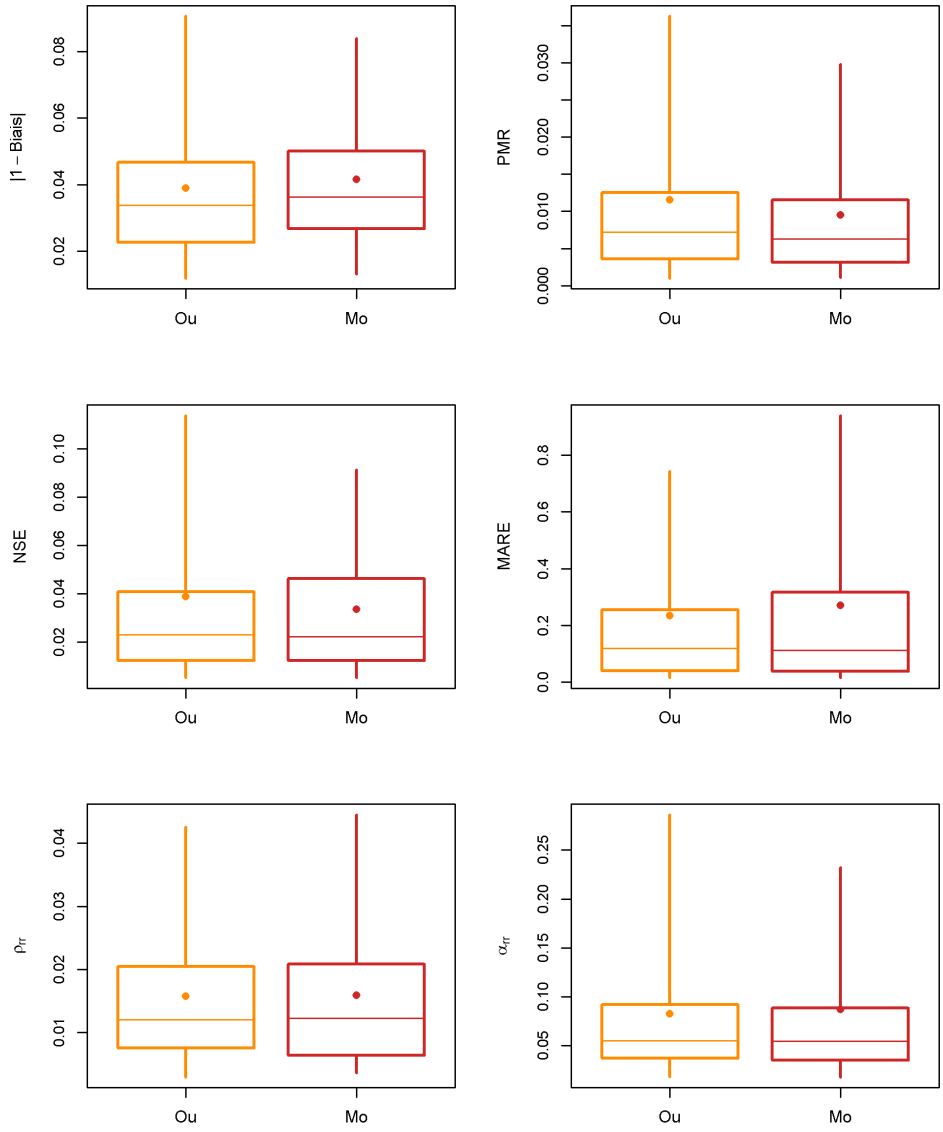


FIGURE 6.4 : Distribution des écarts-types des six critères d'évaluation (un par graphe), calculés pour chaque bassin sur l'ensemble des performances obtenues en contrôle dans le cadre du GSST, avec les deux formules d'évaporation potentielle sur les bassins du sous-échantillon australien. Les couleurs orange et rouge foncé indiquent respectivement les évaporations potentielles d'Oudin et de Morton. À chaque procédure de calage-contrôle, les performances des trois modèles hydrologiques ont été agrégées en une seule valeur pour ne considérer que la variabilité liée au choix de la période de calage.

Enfin, la formule de Penman est souvent associée aux scores les moins stables. Ajouté à des performances moyennes plutôt intermédiaires par rapport aux autres formules, ce résultat montre que cette formule n'est pas la plus avantageuse pour la modélisation hydrologique en France. Dans l'ensemble, les résultats analysés jusqu'ici indiquent que les meilleures formules d'évaporation potentielle sont celles d'Oudin et de Morton.

Ces deux formules se départagent de manière équivoque sur le sous-échantillon de bassins australiens (Figure 6.4). Les performances obtenues avec la formule d'Oudin sont moins sensibles à la période de calage que celles obtenues avec la formule de Morton pour les critères de biais et de MARE. Les rôles s'inversent dans le cas des autres critères, ce qui rend ces résultats peu cohérents et difficiles à interpréter à l'aune des performances moyennes décrites précédemment.

Il est intéressant de rappeler que la comparaison des caractéristiques des évaporations potentielles produites par les deux formules en Australie effectuée au paragraphe 3.3.4.b montre que la demande évaporatoire modélisée par la formule d'Oudin est nettement inférieure à celle modélisée par la formule de Morton. Nous proposons donc d'interpréter les résultats obtenus pour la formule d'Oudin sur les bassins australiens de la même façon que précédemment pour la formule de Penman-Monteith, c'est-à-dire comme une manifestation possible d'une sous-estimation de la demande évaporatoire par la formule d'Oudin sur une partie du sous-échantillon de bassins australiens.

6.1.1.c. Sensibilité des modèles hydrologiques à la formule d'évaporation potentielle

Avant d'analyser plus en profondeur le comportement des modèles forcés par la formule de Penman-Monteith en France et par la formule d'Oudin en Australie, nous avons effectué une estimation de la sensibilité des trois modèles aux entrées d'évaporation potentielle. Pour cela, nous avons calculé pour chaque bassin et chaque modèle hydrologique l'écart-type des performances moyennes en contrôle obtenues pour les formules d'évaporation potentielle disponibles. Le calcul d'un écart-type sur quatre valeurs a certes peu de sens en statistiques, et sur deux valeurs encore moins, mais nous avons considéré qu'il permettait d'apprécier grossièrement les écarts de performances entre les formules d'évaporation. Par le moyen de cette analyse, nous avons évalué la capacité des différents modèles à s'adapter à des erreurs dans les entrées d'évaporation potentielle (ou au moins à des différences, puisqu'il n'est pas possible de critiquer directement les chroniques d'évaporation potentielle). Les résultats sont présentés sur la Figure 6.5
Il est clair que le modèle GR4J a une sensibilité nettement inférieure aux entrées d'évaporation potentielle que les modèles TOPMO et SimHyd. Ce constat est valide pour l'ensemble des critères d'évaluation étudiés. De manière générale, SimHyd est le modèle le plus sensible à la formule d'évaporation potentielle choisie pour la majorité des critères de performance.

Dans la suite, nous étudions les raisons pour lesquelles la formule d'Oudin est moins adaptée que la formule de Morton pour les bassins australiens.

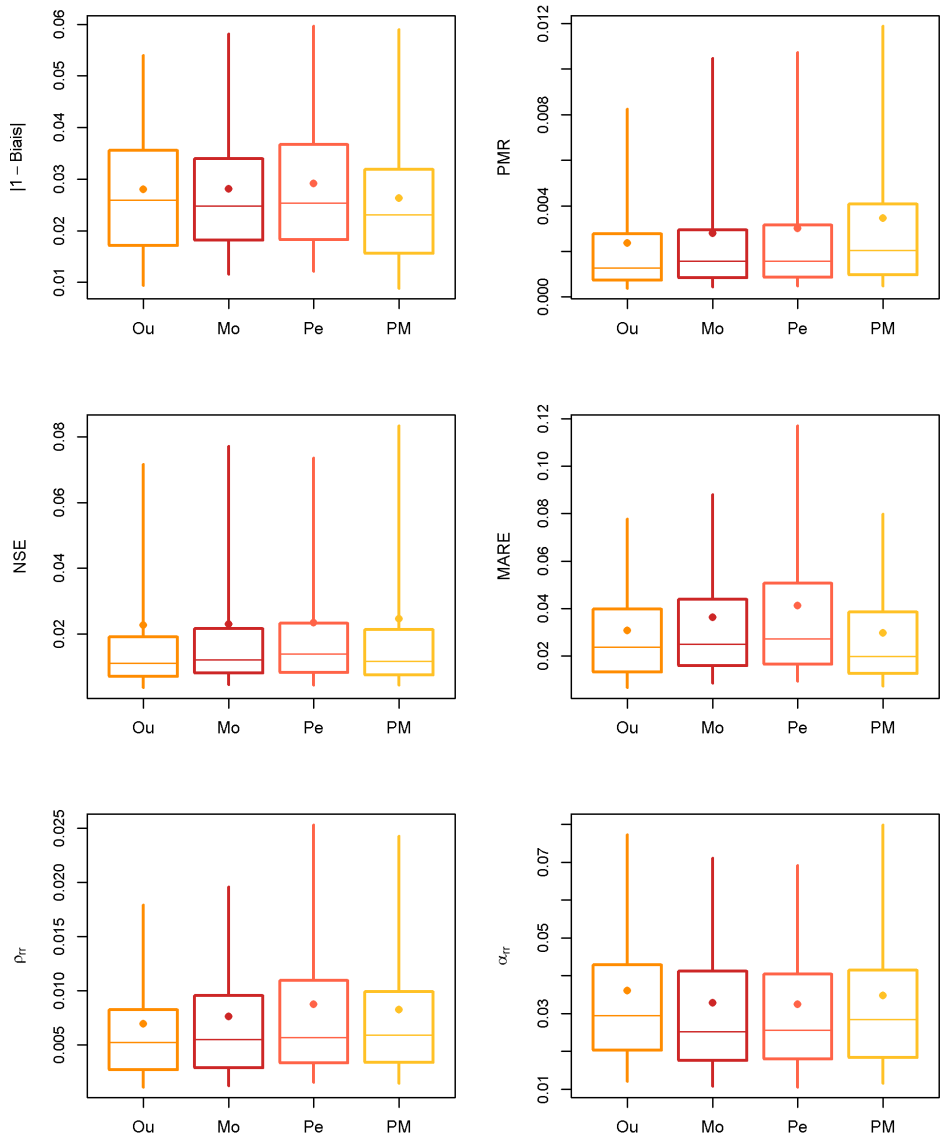


FIGURE 6.3 : Distribution des écarts-types des six critères d'évaluation (un par graphe), calculés pour chaque bassin sur l'ensemble des performances obtenues en contrôle dans le cadre du GSST, avec chacune des quatre formules d'évaporation potentielle sur les bassins du sous-échantillon français. Les couleurs du jaune au rouge foncé ont été choisies de telle sorte à indiquer les évaporations moyennes dans l'ordre croissant des valeurs moyennes d'évaporation potentielle sur le sous-échantillon. À chaque procédure de calage-contrôle, les performances des trois modèles hydrologiques ont été agrégées en une seule valeur pour ne considérer que la variabilité liée au choix de la période de calage.

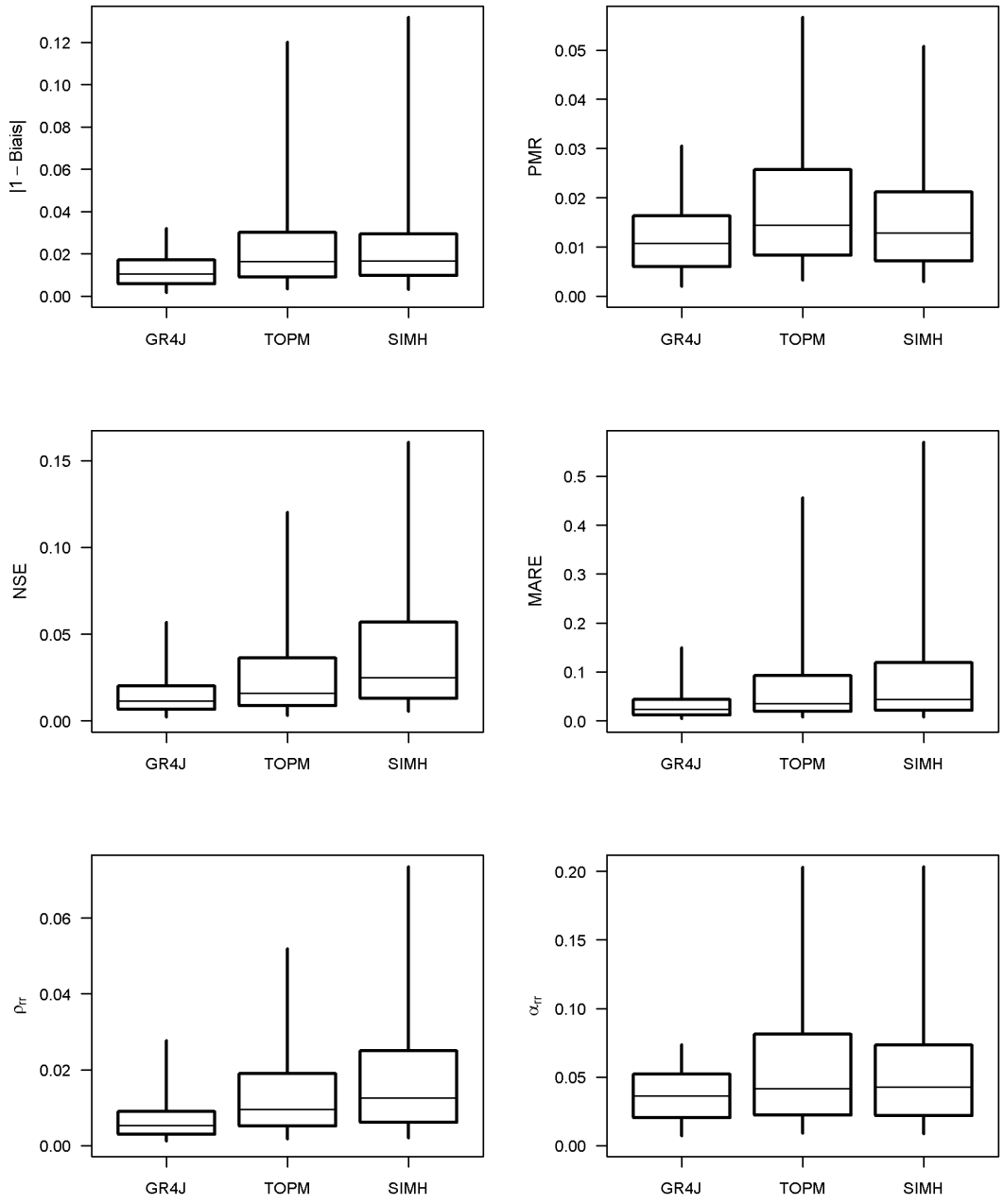


FIGURE 6.5 : Distribution des écarts-types des six critères d'évaluation (un par graphe), calculés pour chaque bassin sur l'ensemble des performances moyennes obtenues en contrôle avec les différentes formules d'évaporation potentielle, pour chacun des trois modèles sur les bassins de l'échantillon. Les écarts-types sont donc calculés à partir de quatre valeurs sur les bassins français et à partir de deux valeurs sur les bassins australiens.

6.1.2. Interprétation des résultats obtenus avec la formule d'Oudin sur les bassins australiens

6.1.2.a. Exemple de la Nambucca River à Bowraville

Pour mieux comprendre le comportement des modèles forcés par la formule d'Oudin sur l'échantillon de bassins australiens, nous proposons d'examiner l'exemple de la Nambucca River à Bowraville (AU205006). Sur ce bassin, les biais moyens sur les débits en contrôle de GR4J, TOPMO et SimHyd atteignent respectivement 4%, 27% et 24% lorsque les modèles sont forcés avec la formule d'Oudin, contre 4%, 6% et 4% avec celle de Morton. Nous avons représenté sur la Figure 6.6 le cycle annuel moyen de l'évaporation réelle mensuelle simulée par les modèles hydrologiques. Les simulations ont été réalisées en calant le modèle sur toute la chronique disponible.

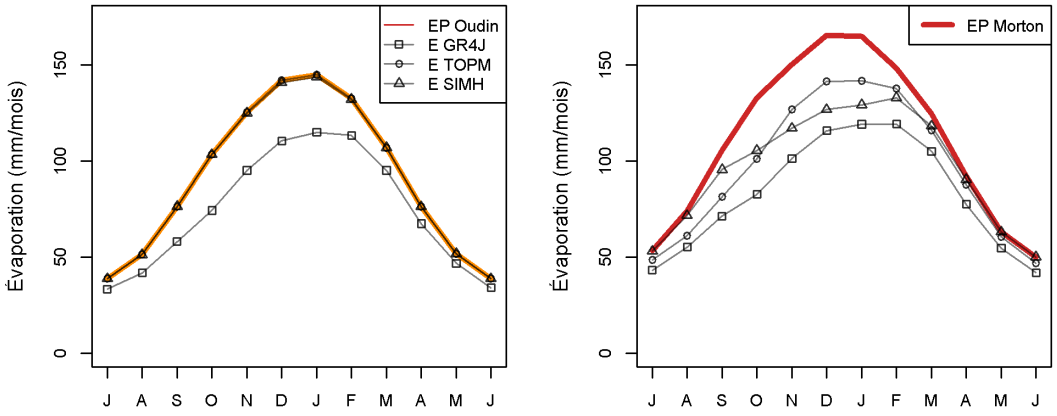


FIGURE 6.6 : Cycles annuels moyens de l'évaporation potentielle et des évaporations réelles simulées par les trois modèles hydrologiques, pour les formules d'Oudin (orange) et de Morton (rouge foncé). Les lignes noires représentent les évaporations réelles mensuelles simulées par GR4J (carrés), TOPMO (cercles) et SimHyd (triangles). Les cycles annuels sont représentés par rapport au calendrier hydrologique australien.

Forcés avec l'évaporation potentielle d'Oudin, TOPMO et SimHyd simulent une évaporation réelle égale à la demande évaporatoire, c'est-à-dire au maximum de ce qu'il est possible d'évaporer (évaporation potentielle de 1089 mm sur une année en moyenne avec la formule d'Oudin, contre 1325 mm avec la formule de Morton). Ces modèles sont justement ceux pour lesquels le biais est problématique, y compris en calage sur toutes les données disponibles. Tout se passe comme si les paramètres des modèles s'étaient organisés de telle sorte que les réservoirs depuis lesquels l'eau est attaquée par l'évaporation soient toujours suffisamment remplis pour évaporer à taux maximal. L'analyse de SimHyd montre en effet que la taille du réservoir d'interception atteint sa valeur maximale autorisée, c'est-à-dire 11 mm, et que son réservoir d'humidité a une capacité également très élevée (1130 mm) par rapport à celle du modèle sur les autres bassins. Associés à un paramètre d'infiltration très élevé, ces deux réservoirs évaporent effectivement à taux potentiel. Dans le cas de TOPMO, le paramètre de taux d'évaporation est fortement négatif en comparaison du taux de remplissage moyen du réservoir d'infiltration, et conduit à évaporer à taux maximal.
L'écart entre les évaporations potentielles simulées par les formules d'Oudin et de Morton n'est pas très important, mais dans le deuxième cas aucun des trois modèles n'évapore constamment à taux potentiel. SimHyd parvient à évaporer suffisamment durant le printemps austral (entre juillet et septembre) pour régler son bilan sans évaporer à taux maximal pendant l'été austral. TOPMO évapore à taux très élevé durant l'été et l'automne austral. Ainsi, forcés avec l'évaporation de

Morton, les trois modèles simulent une évaporation réelle supérieure en moyenne à celle qu'ils simulent avec la formule d'Oudin (respectivement 988 mm, 1152 mm et 1155 mm pour GR4J, TOPMO et SimHyd avec la formule de Morton sur une année en moyenne, contre respectivement 886 mm, 1089 mm et 1086 mm avec la formule d'Oudin).

6.1.2.b. Analyse du rendement mensuel de l'évaporation potentielle

Le constat formulé par l'analyse du bassin de la Nambucca River est généralisé en évaluant la moyenne du rendement mensuel de l'évaporation potentielle, c'est-à-dire le ratio entre l'évaporation réelle mensuelle simulée par les modèles et l'évaporation potentielle mensuelle, sur le reste des bassins du sous-échantillon australien (Figure 6.7).
Bien que le rendement de l'évaporation potentielle du modèle GR4J soit peu affecté par le choix de la formule d'évaporation potentielle en comparaison des autres modèles, ce rendement est tout de même plus élevé lorsque le modèle est forcé avec l'évaporation potentielle d'Oudin. Le rendement de l'évaporation potentielle de TOPMO et SimHyd est nettement plus sensible que celui de GR4J. Ces modèles évaporent à un taux supérieur à 0,95 dans près d'un quart des bassins du sous-échantillon australien.

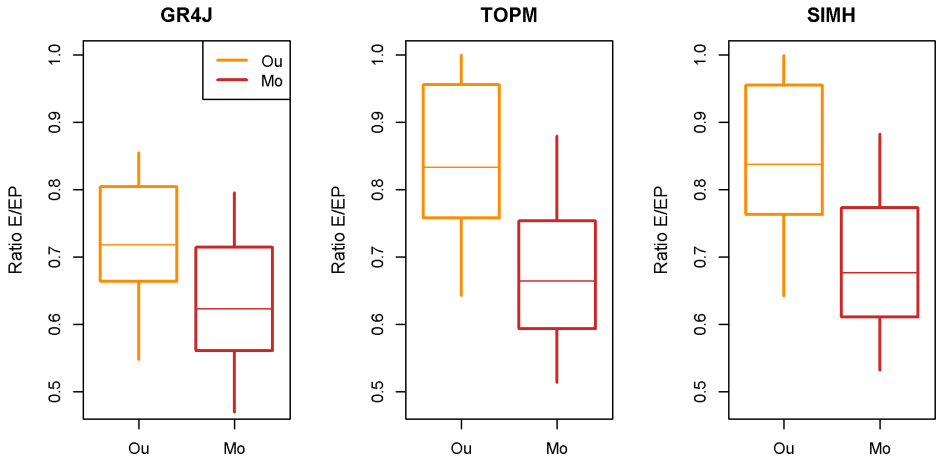


FIGURE 6.7 : Distribution des rendements mensuels moyens de l'évaporation potentielle des trois modèles hydrologiques sur les bassins du sous-échantillon australien, avec les deux formules d'évaporation potentielle. Les couleurs orange et rouge foncé indiquent respectivement les évaporations d'Oudin et de Morton.

La Figure 6.8 montre qu'il existe un lien entre biais des modèles et ratio moyen du rendement mensuel de l'évaporation. Les distributions des biais en contrôle des trois modèles y sont représentées pour chacun des quartiles de rendement de l'évaporation potentielle simulée par les modèles avec la formule d'Oudin. Alors que les simulations qu'ils produisent sont dans l'ensemble peu biaisées, les modèles TOPMO et SimHyd surestiment largement les volumes écoulés pour les bassins du dernier quartile, c'est-à-dire pour les bassins où ils simulent des rendements supérieurs à 0,95. En considérant que les mesures de précipitations et de débits de ces bassins sont non biaisées, ce qui n'est pas certain, la demande évaporatoire modélisée par la formule est sous-estimée d'au moins la différence entre débits observés et débits simulés moyens.
Afin de confirmer l'hypothèse précédente, nous avons vérifié la cohérence des biais des modèles TOPMO et SimHyd sur le sous-échantillon de bassins où le rendement de l'évaporation potentielle des deux modèles était supérieure à 0,98. Les biais des deux modèles y sont comparés sur la Figure 6.9.

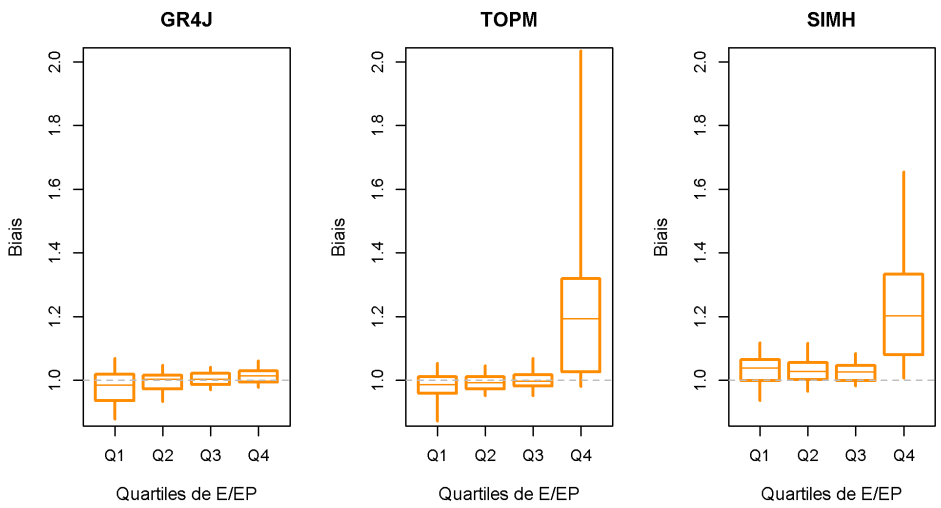


FIGURE 6.8 : Distributions des biais moyens en contrôle des trois modèles hydrologiques avec la formule d'Oudin sur les bassins du sous-échantillon australien, pour chacun des quartiles des rendements moyens de l'évaporation potentielle de la Figure 6.7. Les biais n'ont pas été normalisés.

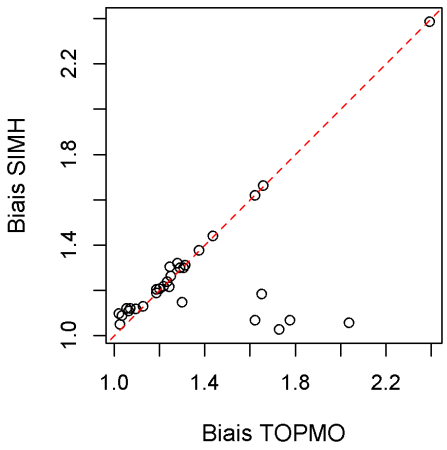


FIGURE 6.9 : Comparaison des biais moyens en contrôle des modèles SimHyd et TOPMO avec l'évaporation d'Oudin, sur les bassins australiens où le rendement de l'évaporation potentielle simulé par les modèles TOPMO et SimHyd avec la formule d'Oudin est supérieur à 0,98.

Dans la plupart des bassins, les deux modèles commettent les mêmes erreurs sur les volumes moyens, ce qui confirme l'hypothèse selon laquelle ces erreurs sont liées à la même raison, c'est-à-dire une évaporation potentielle sous-estimée par la formule d'Oudin. Pour une poignée de bassins cependant, le débit simulé par SimHyd n'est pas biaisé au contraire de celui simulé par TOPMO. Cela est dû à un paramétrage particulier du réservoir d'humidité de sol de SimHyd, qui lui permet de garder une très grande quantité d'eau en rétention tout en réduisant les pertes par écoulement hypodermique et par recharge. L'eau stockée au cours du temps ne s'écoulant pas tant que la capacité maximale du réservoir n'est pas atteinte, ce comportement permet de fermer le bilan hydrique en jouant sur le ΔS de l'équation du bilan de masse, c'est-à-dire sur les variations du stock du bassin versant. Ce paramétrage n'a rien de physique et est simplement exploité par SimHyd afin de compenser artificiellement le déficit de demande évaporatoire. Ces paramètres ne conviendraient évidemment pas pour une étude de changement climatique si tôt que la capacité maximale de rétention du réservoir serait atteinte. Le paramétrage de TOPMO ne lui permet d'adopter ce genre de

comportement.

6.1.2.c. Interprétation physique à partir des caractéristiques des bassins concernés

La carte des bassins où le rendement de l'évaporation potentielle de TOPMO et SimHyd est supérieur à 0,98 est affichée en Figure 6.10.

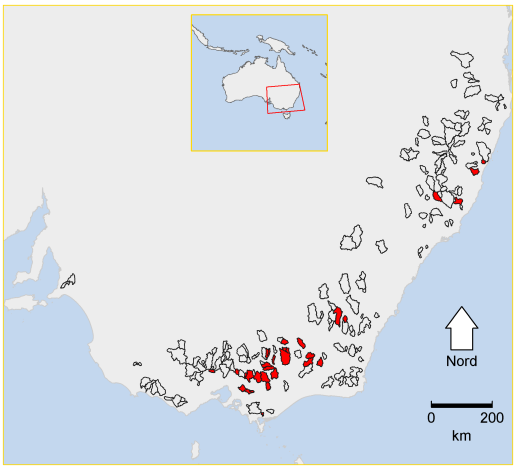


FIGURE 6.10 : Carte des bassins australiens où le rendement de l'évaporation potentielle simulé par les modèles TOPMO et SimHyd avec la formule d'Oudin est supérieur à 0,98 (en rouge).

La majorité de ces bassins est située au Nord-Est de Melbourne, qui correspond à la partie Sud des Alpes australiennes. Cette région est couverte de forêts et est en moyenne plus pentue que le reste de la zone géographique de notre échantillon. La corrélation de Spearman entre la moyenne par bassin des rendements de l'évaporation potentielle des deux modèles et le couvert forestier (respectivement la pente moyenne) des bassins australiens atteint 0,58 (respectivement 0,61). Il est donc possible que la raison pour laquelle la formule d'Oudin sous-estime l'évaporation potentielle sur une partie des bassins australiens soit liée à une absence de lien entre température moyenne et demande évaporatoire en forêt. À conditions climatiques équivalentes, les forêts évaporent en moyenne plus que les espaces ouverts alors que la température y est généralement plus basse. L'interception contribue à une partie importante de l'évaporation des forêts, et on peut supposer que la demande évaporatoire subie par l'eau interceptée au niveau de la canopée est liée au rayonnement. Cette hypothèse est peut-être une piste pour expliquer la raison pour laquelle la formule de Morton, qui est une formule fondée sur une approche radiative, y est plus adaptée. Il convient toutefois de nuancer en rappelant que la complexité de la formule de Morton associée à la présence de quelques paramètres empiriques pourrait aussi agir de sorte à obtenir la bonne réponse pour les mauvaises raisons en compensant des hypothèses physiques invalides.

Face à l'impossibilité d'arbitrer définitivement sur les causes des meilleures estimations réalisées par la formule de Morton, nous pouvons au moins conclure que la qualité des simulations des modèles hydrologiques (excepté GR4J) est altérée par une évaporation potentielle sous-estimée, en particulier lorsqu'ils représentent le bilan hydrique des bassins avec l'évaporation comme unique source de perte.

6.1.3. Interprétation des résultats obtenus avec les formules de Penman-Monteith et d'Oudin sur les bassins français

6.1.3.a. Analyse du rendement mensuel de l'évaporation potentielle

La comparaison générale des performances des modèles forcés par les différentes formules d'évaporation a montré que la formule de Penman-Monteith se positionnait par rapport aux autres sur les bassins français de la même manière que la formule d'Oudin se positionnait par rapport à celle de Morton sur les bassins australiens. Nous proposons ici de vérifier que ces moins bonnes performances sont également liées à une demande évaporatoire sous-estimée.

Nous examinons la distribution des rendements de l'évaporation potentielle simulés par les modèles hydrologiques en Figure 6.11. Des motifs similaires à celui observé pour l'évaporation potentielle d'Oudin en Australie sont identifiables pour l'évaporation potentielle d'Oudin et de Penman-Monteith sur le sous-échantillon français. De fait, au moins un quart des bassins nécessitent d'être modélisés au moyen d'un rendement très élevé, de plus de 0,95, par TOPMO et SimHyd.

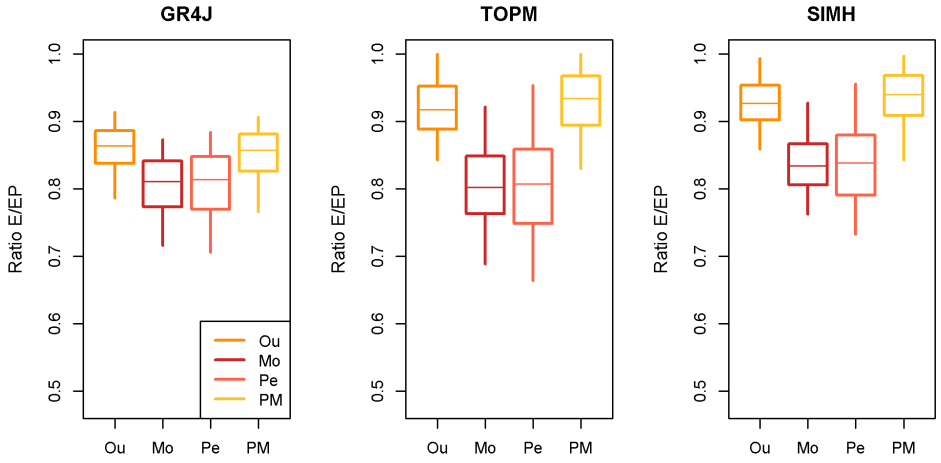


FIGURE 6.11 : Distributions des rendements mensuels moyens de l'évaporation potentielle des trois modèles hydrologiques sur les bassins du sous-échantillon français, avec chacune des quatre formules d'évaporation potentielle. Les couleurs du jaune au rouge foncé ont été choisies de telle sorte à indiquer les évaporations moyennes dans l'ordre croissant des valeurs moyennes d'évaporation potentielle sur le sous-échantillon.

Conformément à ce qui était attendu, les biais en contrôle sont très élevés sur les bassins où le rendement de l'évaporation potentielle est proche de 1, en comparaison des biais observés sur le reste du sous-échantillon (Figure 6.12). C'est donc le signe que l'évaporation potentielle est insuffisante pour que le bilan hydrique puisse être fermé par les modèles, même lorsque leurs paramètres sont choisis de sorte à évaporer constamment à taux maximal. Les erreurs des modèles sont plus importantes avec la formule de Penman-Monteith qu'avec celle d'Oudin, pour la raison que cette dernière produit une évaporation potentielle légèrement plus forte que la première.
On note que dans le quartile des bassins où le rendement de l'évaporation potentielle est le plus faible (Q1 sur la Figure 6.12), les biais sont à quelques exceptions près centrés autour de 1. Dans l'éventualité où l'évaporation potentielle calculée sur certains de ces bassins serait surestimée, de sorte que les paramètres doivent s'ajuster pour limiter l'accès de la demande évaporatoire à l'eau, on peut tout de même constater que les débits ne sont dans l'ensemble pas sous-estimés. Cela invite à considérer qu'en modélisation hydrologique il est préférable de surestimer l'évaporation potentielle plutôt que de la sous-estimer.

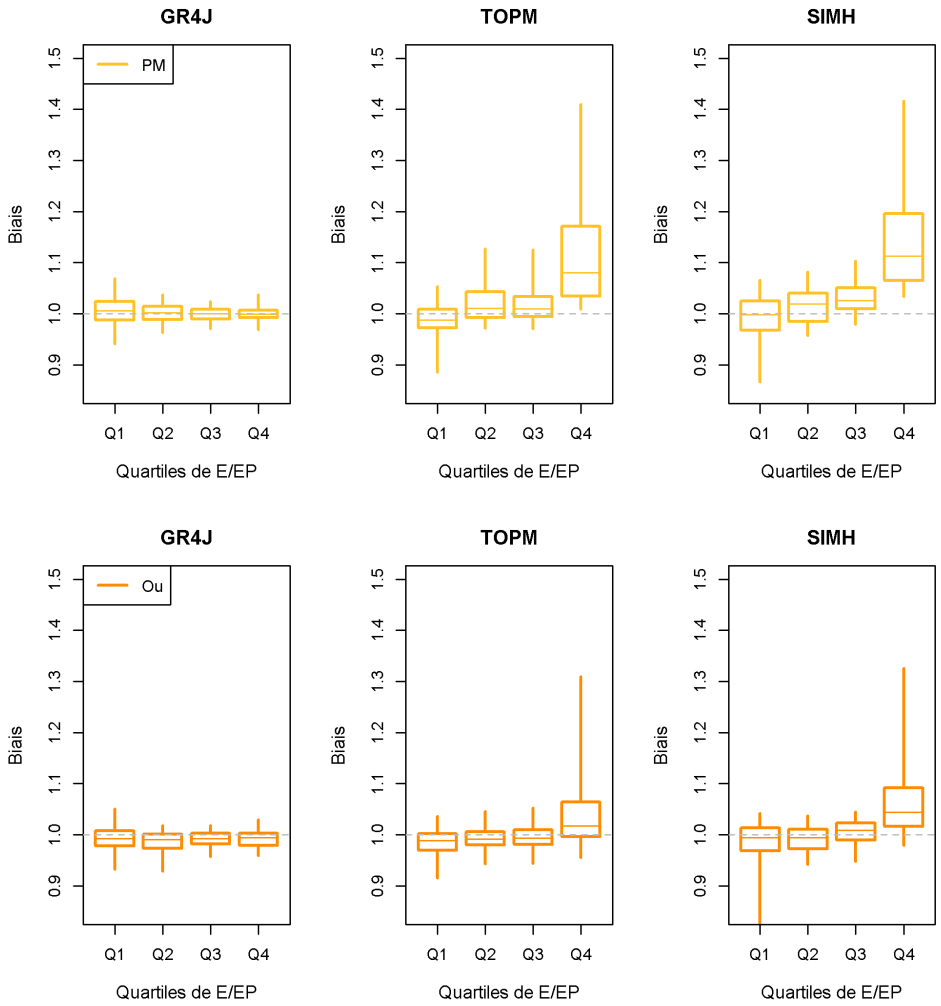


FIGURE 6.12 : Distributions des biais moyens en contrôle des trois modèles hydrologiques avec les formules de Penman-Monteith (en jaune) et d'Oudin (en orange) sur les bassins du sous-échantillon australien, pour chacun des quartiles des rendements moyens de l'évaporation potentielle de la Figure 6.11. Les biais n'ont pas été normalisés.

6.1.3.b. Interprétation physique à partir des caractéristiques des bassins concernés

Les bassins du sous-échantillon français où le rendement de l'évaporation potentielle de TOPMO et Simhyd est supérieur à 0,98 sont situés à l'Est de la France, où le climat est le plus continental (Figure 6.13). L'évaporation potentielle produite par la formule de Penman sur ces bassins étant suffisante pour que les modèles puissent fermer le bilan hydrique, nous avons dans un premier temps suspecté un rôle de la résistance aérodynamique, qui interagit avec la résistance stomatique au dénominateur de l'équation de Penman-Monteith. Il n'existe cependant aucune corrélation probante entre ces bassins et la répartition spatiale de la résistance aérodynamique moyenne calculée dans la formule de Penman-Monteith.
L'analyse des corrélations du rendement de l'évaporation potentielle avec des critères topographiques et hydro-climatiques montre un lien significatif du rendement avec la température moyenne des bassins (corrélation de Spearman de 0,34) et avec le volume de précipitations solides (corrélation de Spearman de 0,34). Les précipitations solides influencent justement l'albédo des mailles SAFRAN dans nos calculs de l'évaporation potentielle avec les formules de Penman et de Penman-Monteith, de telle sorte que les bassins de l'Est de la France ont un albédo moyen parfois supérieur à 0,28, et réfléchissent donc une part supérieure du rayonnement haute fréquence (Figure 6.13.b). La corrélation du rendement mensuel moyen de l'évaporation potentielle avec

l'albédo moyen des bassins est de 0,37.
Bien que de telles valeurs d'albédo soient certainement irréalistes, il ne semble pas raisonnable de conclure que le calcul de l'albédo est le seul facteur responsable de la sous-estimation de l'évaporation potentielle par la formule de Penman-Monteith. La formule de Penman est en effet également affectée par ce calcul et n'aboutit pas à ces mêmes erreurs. Ces résultats soulignent donc davantage la nécessité de faire preuve de vigilance quant aux téléscopages possibles entre les nombreux paramétrages de l'équation de Penman-Monteith sur l'évaporation potentielle estimée. On peut supposer que les erreurs constatées proviennent de l'action combinée de la résistance aérodynamique fixée à une valeur inadéquate, de la mise à jour de l'albédo sur les surfaces enneigées, et peut-être d'erreurs dans les données d'entrée dans les formules d'évaporation (par exemple le vent journalier).

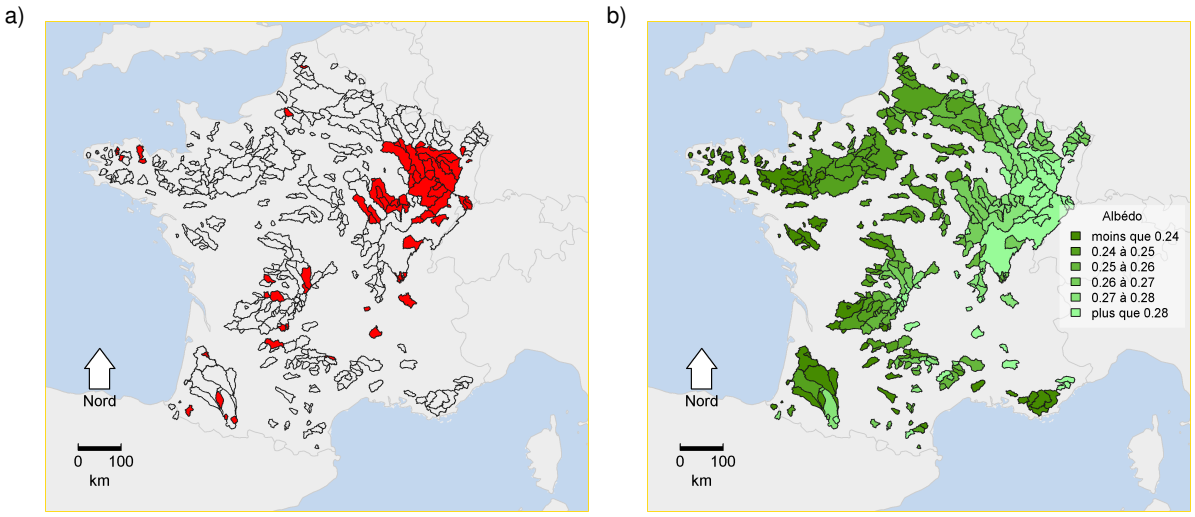


FIGURE 6.13 : a) Carte des bassins français où le rendement de l'évaporation potentielle simulé par les modèles TOPMO et SimHyd avec la formule de Penman-Monteith est supérieur à 0,98 (en rouge). b) Carte des valeurs moyennes de l'albédo des bassins français tel que calculé avec la formule de Penman-Monteith. Les couleurs du vert foncé au vert clair indiquent les valeurs croissantes d'albédo.

Le rendement moyen de l'évaporation potentielle simulé en forçant TOPMO et SimHyd avec l'évaporation d'Oudin est également fortement liée à la température moyenne et au volume moyen des précipitations solides des bassins du sous-échantillon français (corrélations de Spearman de –0,36 et 0,41 respectivement). Toutefois, à l'inverse du cas précédent, la répartition spatiale des bassins où les deux modèles évaporent à taux potentiel n'est pas caractérisée par une concentration remarquable de bassins versants dans une région spécifique (Figure 6.14). On peut cependant remarquer qu'une partie de ces bassins se situe sur des massifs forestiers connus, tels que les Vosges, les Cévennes et l'Aubrac. Bien que significative, les corrélations du rendement mensuel de l'évaporation potentielle avec la couverture forestière et la pente moyenne des bassins du sous-échantillon français restent toutefois plus faibles que celles calculées pour les bassins australiens (respectivement 0,20 et 0,19).
Ce constat pourrait constituer un point de départ pour des travaux spécifiques à la formule d'Oudin et à son comportement en forêt et sur les massifs. On peut considérer que le rôle important joué par l'interception en forêt puisse justifier la prise en compte d'une variable supplémentaire, comme le rayonnement observé, pour évaluer la demande évaporatoire au niveau de la canopée, ou au paramétrage d'une correction dépendant de l'indice foliaire du terrain.

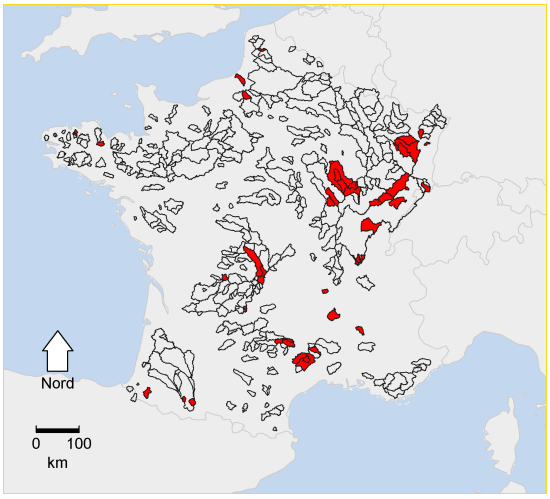


FIGURE 6.14 : Carte des bassins français où le rendement de l'évaporation potentielle simulé par les modèles TOPMO et SimHyd avec la formule d'Oudin est supérieur à 0,98 (en rouge)

6.1.4. Synthèse

Les travaux de comparaison des formules d'évaporation potentielle ont conduit jusqu'ici à identifier l'évaporation de Morton comme la plus indiquée pour la modélisation hydrologique sur les bassins australiens, et celles d'Oudin et de Morton sur les bassins français. Ces formules permettent à la fois de maximiser les performances moyennes en contrôle des modèles hydrologiques et d'améliorer leur robustesse face à la diversité des conditions de calage.

Les développements menés dans cette section mènent enfin au constat trivial qu'il est essentiel pour la modélisation hydrologique que l'évaporation potentielle soit suffisante pour fermer le bilan hydrique. Cette assertion est particulièrement vraie pour les modèles représentant l'évaporation comme unique facteur de pertes d'eau dans l'équation du bilan en eau. On peut arguer que l'absence d'échanges souterrains dans les modèles TOPMO et SimHyd constitue évidemment une hypothèse de modélisation qui peut s'avérer délétère dans certains contextes géologiques. Le modèle GR4J évacue au contraire le surplus d'eau en le modélisant comme une perte par le souterrain. Les bassins du sous-échantillon français sont tous peu influencés et sont situés en amont des bassins versants majeurs qu'ils alimentent, et il est à ce titre raisonnable de considérer que ces bassins puissent perdre de l'eau vers l'aval par le sous-sol. Toutefois, le comportement de GR4J n'est pas nécessairement réaliste dans l'ensemble des bassins où l'évaporation potentielle a été jugée insuffisante. Par ailleurs, c'est grâce à l'absence d'échanges souterrains dans les modèles TOPMO et SimHyd qu'il nous a été possible d'identifier les problèmes relatifs au calcul de l'albédo pour l'estimation de l'évaporation potentielle dans les formules de Penman et de Penman-Monteith, ainsi que les difficultés que rencontrait la formule d'Oudin dans les bassins forestiers australiens.

Au vu de ces résultats, nous avons constaté qu'un aspect important de la modélisation de l'évaporation potentielle n'était pas ressorti de la comparaison des différentes formules. Dans la perspective d'évaluer la robustesse des modèles hydrologiques dans un climat changeant, il semble en effet inévitable d'évaluer le rôle de l'évolution des forçages dans l'explication des variations interannuelles du débit des rivières. Or, en dehors de l'analyse des distributions du PMR atteint par les modèles forcés par les formules d'évaporation potentielle, les discussions ont avant tout été centrées autour de la question de la moyenne de l'évaporation potentielle plutôt qu'autour de ses variations.

6.2. Comparaison de formules d'évaporation potentielle au travers des variations interannuelles du débit

Dans cette partie, nous avons tenté d'éclairer les différences de quelques formulations d'évaporation potentielle couramment utilisées en hydrologie en revenant aux fondamentaux de la modélisation. L'utilisation de modèles hydrologique complique en effet la critique des formules d'évaporation potentielle en compensant leurs défauts par l'ajustement de leurs paramètres en calage. Pour éviter cet écueil, nous avons implémenté des régressions linéaires reliant les anomalies interannuelles de précipitations et d'évaporation potentielle à celles des débits par une relation linéaire. Un autre modèle linéaire reliant le rendement hydrologique des précipitations et l'indice d'humidité, inspiré de la formule proposée par Oldekop [1911], a été utilisé en complément.

6.2.1. Description de la méthode

6.2.1.a. Variables explicatives

La réalité des processus hydrologiques est loin d'être linéaire et supposer que les débits annuels sont liés aux variables climatiques par des relations affines est à l'évidence une hypothèse simplificatrice forte. Cette hypothèse est néanmoins intéressante pour la simplicité d'analyse qu'elle offre. Les coefficients de la régression du débit sur les précipitations et l'évaporation potentielle prennent en effet par construction les valeurs de l'élasticité du débit vis-à-vis de ces variables, valeurs facilement interprétables et déjà abordées par d'autres études (Sankarasubramanian *et al.*, 2001; Andréassian *et al.*, 2016). L'élasticité indique le taux de variation d'une variable expliquée résultant d'une variation de 1% d'une variable explicative. À une date donnée, l'élasticité des débits Q par rapport à une variable X est définie par la formule suivante :

$$\epsilon_{Q/X} = \frac{\Delta Q/\overline{Q}}{\Delta X/\overline{X}} \tag{6.1}$$

Où ΔQ et ΔX représentent respectivement l'anomalie annuelle de Q et X à une date donnée par rapport à leur moyenne de long terme $\overline{Q}$ et $\overline{X}$. Lorsque la variable expliquée varie de manière proportionnelle à la variable explicative, l'élasticité est positive ; lorsque la variable expliquée varie de manière inversement proportionnelle à la variable explicative, l'élasticité est négative. Il convient de noter qu'une valeur d'élasticité est supérieure à 1 ne signifie pas que le débit augmente plus que le volume des précipitations, puisque l'élasticité fait intervenir des variations en pourcentage et donc dépend des volumes moyens précipités et écoulés, qui sont généralement différents sur la plupart des bassins.
D'après la formule, l'élasticité est dépendante de l'année considérée : c'est une élasticité « instantanée ». Dans la suite, nous appelons élasticité la grandeur résultant de la prise en compte des élasticités instantanées de toute la chronique disponible. Cette élasticité peut être estimée à partir de la médiane des élasticités instantanées, ou bien à partir du coefficient de pente d'un modèle affine reliant anomalies des débits et anomalies d'une variable explicative. C'est la deuxième méthode qui est choisie ici.

La facilité d'interprétation de l'élasticité, c'est-à-dire des pentes des modèles affines, permet à la fois d'évaluer l'influence des précipitations et de l'évaporation sur le débit et de comparer les différentes formules d'évaporation potentielle entre elles. La linéarité des équations présente également l'avantage de moins diluer ces différences dans la complexité d'un modèle, qui autoriserait les com-

pensations entre paramètres. Pour résumer succinctement, formuler l'hypothèse linéaire permet paradoxalement de ne pas avoir à en formuler d'autres.
Malgré les avantages garantis par l'usage de régressions linéaires, cette technique présente néanmoins quelques défauts. Certains d'entre eux sont liés au choix de variables explicatives purement climatiques : celui de ne pas inclure les changements d'usage du sol et de l'eau d'origines anthropiques, et celui de ne pas requérir la fermeture du bilan hydrique et donc d'évacuer la question des échanges souterrains entre bassins. Ce dernier point peut toutefois être considéré comme utile ici puisqu'il évite la compensation des différences entre les formules d'évaporation potentielle constatées précédemment pour GR4J. D'autres limitations existent, liées au caractère linéaire des modèles employés, notamment l'absence de représentation de la rétroaction de l'état hydrologique du bassin sur la relation qui lie le débit au climat. Néanmoins, pour pallier cette faille et inclure dans les modèles une représentation de la mémoire des bassins, les variables climatiques des années précédentes ont été ajoutées aux variables explicatives, suivant cette structure :

$$\frac{\Delta Y_n}{\overline{Y}} = \sum_k \epsilon_{X_k,n} \frac{\Delta X_{k,n}}{\overline{X_k}} + \sum_k \epsilon_{X_k,n-1} \frac{\Delta X_{k,n-1}}{\overline{X_k}} \tag{6.2}$$

Où Y est la variable expliquée (par exemple le débit) et X_k représente l'ensemble des variables climatiques explicatives du débit annuel. Le terme $\epsilon_{k,n}$ correspond à l'élasticité de la variable expliquée Y vis-à-vis de la variable explicative X_k à la même année n. Le terme $\epsilon_{k,n-1}$ correspond à l'élasticité de la variable expliquée Y vis-à-vis de la variable explicative X_k l'année précédente $(n-1)$. Dans la suite, on appellera « mémoire » l'influence du climat des années passées sur l'hydrologie d'un bassin à une date donnée.

6.2.1.b. Modèles linéaires testés

Quatre modèles ont été construits et testés, conçus comme deux paires comprenant un modèle de référence sans mémoire et un modèle avec mémoire. La première paire de modèles, appelés « N0 » et « N1 », est basée sur un modèle naïf, strictement linéaire, liant débits et variables climatiques. La deuxième paire de modèles, appelés « B0 » et « B1 », s'inspire de ce que les hydrologues ont coutume d'appeler « l'hypothèse de Budyko » en reliant (ici linéairement) le rendement des précipitations (Q/P) à l'indice d'humidité (pour mémoire, l'indice d'humidité P/EP est l'inverse de l'indice d'aridité EP/P). Dans le deuxième cas, les anomalies de débit sont donc reliées aux anomalies de précipitations et d'évaporation potentielle par une relation non linéaire.

$$\frac{\Delta Q_n}{\overline{Q}} = \epsilon^0_{P,n} \frac{\Delta P_n}{\overline{P}} + \epsilon^0_{EP,n} \frac{\Delta EP_n}{\overline{EP}} \qquad \text{« N0 », (6.3)}$$

$$\frac{\Delta Q_n}{\overline{Q}} = \epsilon^1_{P,n} \frac{\Delta P_n}{\overline{P}} + \epsilon^1_{EP,n} \frac{\Delta EP_n}{\overline{EP}} + \epsilon^1_{P,n-1} \frac{\Delta P_{n-1}}{\overline{P}} + \epsilon^1_{EP,n-1} \frac{\Delta EP_{n-1}}{\overline{EP}} \qquad \text{« N1 », (6.4)}$$

$$\frac{\Delta (Q/P)_n}{\overline{Q/P}} = \epsilon^0_{P/EP,n} \frac{\Delta (P/EP)_n}{\overline{P/EP}} \qquad \text{« B0 », (6.5)}$$

$$\frac{\Delta (Q/P)_n}{\overline{Q/P}} = \epsilon^1_{P/EP,n} \frac{\Delta (P/EP)_n}{\overline{P/EP}} + \epsilon^1_{P/EP,n-1} \frac{\Delta (P/EP)_{n-1}}{\overline{P/EP}} \qquad \text{« B1 », (6.6) ?}$$

Caler quatre paramètres sur des séries d'une vingtaine d'observations peut produire des résultats très incertains et en ce sens les deux modèles « B » diminuent ce risque par rapport aux modèles « N », à mémoire constante, en offrant une formulation plus parcimonieuse des processus hydrologiques. Toutefois, ils ne permettent pas d'arbitrer entre les rôles respectifs des précipitations et de l'évaporation dans le bilan hydrologique. L'utilisation des deux types de modèle est donc complémentaire et devrait permettre de juger plus finement la pertinence de l'ajout d'une forme de mémoire

aux modèles.
Nous avons également testé des modèles intégrant jusqu'aux variables climatiques de l'année $n-2$, dont il sera fait brièvement mention dans l'évaluation des résultats. Pour rappel, les années sont comptabilisées à partir du calendrier hydrologique propre à chaque pays.

6.2.1.c. Estimation des paramètres

La méthode présentée jusqu'à ce point possède deux problèmes d'ordre purement statistique. Utiliser une simple régression pour expliquer une variable auto-corrélée peut perturber l'estimation des paramètres des variables explicatives. Par ailleurs, l'ajout des variables climatiques décalées d'un an invalide l'hypothèse d'indépendance des variables explicatives de la régression, puisque chaque terme intervient deux fois dans l'explication des anomalies de débit.
Pour remédier à ces deux problèmes, la méthode des moindres carrés généralisés a donc été préférée à la régression linéaire simple (Aitken, 1936). Les résultats présentés dans la suite sont ceux issus de ce modèle.
Les résultats ont été confirmés par l'implémentation d'un modèle linéaire simple s'appuyant sur une méthode de ré-échantillonnage de type Monte Carlo (résultats non présentés). Cette technique permet également de s'affranchir de l'autocorrélation des débits annuels, en formulant l'hypothèse que l'ordre (i.e. le temps) n'a pas de valeur explicative. Le ré-échantillonnage aléatoire des observations a été construit de sorte à ne pas sélectionner d'années consécutives, afin de ne pas réutiliser deux fois la même valeur d'une variable explicative. Les résultats obtenus ainsi sont similaires à ceux produits par la méthode des moindres carrés généralisés et ne sont pas présentés dans la suite.

On notera que l'évaluation de la robustesse des modèles linéaires n'est pas l'objet de cette analyse, et que de ce fait les modèles ont été calés sur l'ensemble des chroniques disponibles sans procéder à un *Split-Sample Test*, pourvu qu'elles contiennent un nombre suffisant d'années consécutives sans lacunes. Un sous-ensemble de 338 bassins versants français et de 161 bassins australiens a donc été formé en ne gardant que les bassins possédant au moins 20 triplets d'années consécutives ayant chacune moins de 10% de débits lacunaires.
Les analyses présentées dans la suite ne concernent d'abord que les bassins versants français. Ce choix résulte de la non disponibilité des évaporations potentielle de Penman et de Penman-Monteith sur les bassins australiens. Le cas des bassins versants australiens est seulement abordé ensuite.

6.2.2. Élasticité du débit vis-à-vis des variables climatiques annuelles

6.2.2.a. Élasticité des modèles « N »

Les valeurs d'élasticité du débit vis-à-vis des précipitations et de l'évaporation potentielle obtenues avec les modèles « N » sont présentées en Figure 6.15.
Les précipitations de l'année en cours ont logiquement une influence forte sur les écoulements. L'étroitesse des distributions suggère que cette influence est spatialement homogène. Curieusement, les valeurs d'élasticité du débit vis-à-vis de l'évaporation potentielle centrées autour de zéro semblent indiquer que l'évaporation potentielle de l'année en cours n'a pas de réelle incidence sur les écoulements pour la très grande majorité des bassins, et cela quelle que soit la formule d'évaporation potentielle utilisée. Un test de Student confirme la faible significativité de la plupart des valeurs d'élasticité.
Des différences sont en revanche visibles au regard des valeurs de l'élasticité des débits vis-à-vis de l'évaporation potentielle de l'année précédente, avec une élasticité généralement négative, conformément aux attentes. Les valeurs correspondant à la formule d'Oudin sont nettement supérieures en valeur absolue et réparties sur spectre plus large que les valeurs issues des autres formules.

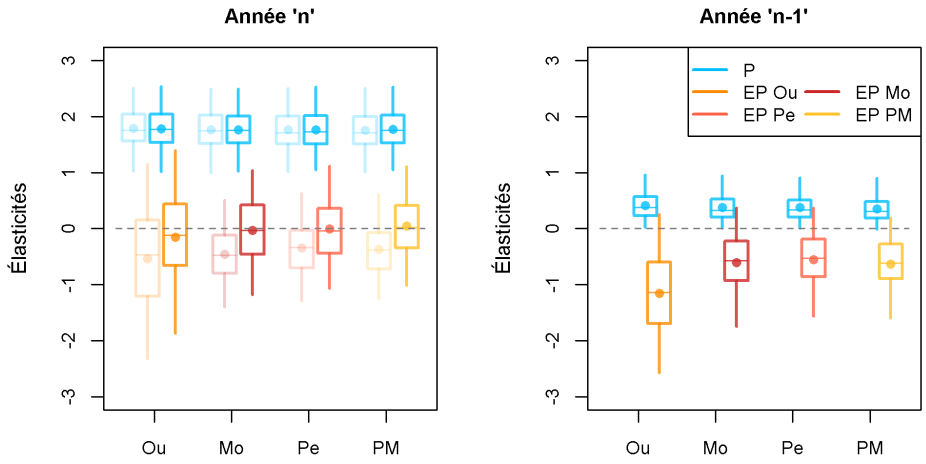


FIGURE 6.15 : Distributions des paramètres associés aux variables des régressions linéaires « N0 » et « N1 » aux années n et $n-1$, pour chacune des formules d'évaporation potentielle. Les paramètres des régressions linéaires sont équivalents aux élasticités du débit vis-à-vis des précipitations ou de l'évaporation potentielle. Les boxplots aux contours transparents représentent les paramètres des régressions linéaires « N0 » et les boxplots aux contours opaques représentent les paramètres des régressions linéaires « N1 ». Les paramètres des précipitations sont en bleu. Les paramètres de l'évaporation potentielle sont en couleurs chaudes, du jaune au rouge foncé, de sorte à indiquer les évaporations moyennes dans l'ordre croissant des valeurs moyennes d'évaporation potentielle sur le sous-échantillon.

L'évaporation potentielle d'Oudin varie peu à l'échelle interannuelle (voir Chapitre 3), ce qui amène vraisemblablement le modèle linéaire à compenser par une valeur absolue élevée des coefficients de la régression. La dispersion des valeurs d'élasticité obtenues avec cette formule est néanmoins surprenante, alors qu'on attendrait une certaine cohérence dans la relation qui lie l'évaporation potentielle au débit au regard de la cohérence des liens qu'entretient l'écoulement avec ses autres déterminants (précipitations, géologie). Toutefois, rapportées à leur élasticité moyenne (i.e. en calculant le coefficient de variations), les valeurs d'élasticité des quatre formules sont assez proches.

6.2.2.b. Élasticité des modèles « B »

Les valeurs d'élasticité du rendement des précipitations vis-à-vis de l'indice d'humidité (P/EP) obtenues avec les modèles « B » sont montrées sur la Figure 6.16.
Les résultats obtenus avec les modèles linéaires inspirés de « l'hypothèse de Budyko » confirment l'importance de la situation climatique pendant l'année hydrologique $n-1$. L'indice d'humidité des années n et $n-1$ est positivement relié aux variations du rendement des précipitations, et ce lien est significatif au moins pour les trois quarts des bassins de l'étude. La positivité du lien entre rendement des précipitations et indice d'humidité sur la plupart des bassins de l'échantillon français témoigne surtout de l'influence des précipitations sur le débit de l'année en cours constatée dans le cas des modèles « N », dans la mesure où celle de l'évaporation potentielle ne semblait pas déterminante.
Au contraire des conclusions précédentes, les quatre formulations de l'évaporation potentielle ne conduisent pas à des résultats différents. Les valeurs d'élasticité obtenues avec la formule d'Oudin sont seulement légèrement supérieures à celles des trois autres formules.
Il est possible que la faiblesse du lien entre variations du débit et variations de l'évaporation potentielle puisse être expliquée en partie par la définition de l'année hydrologique en France métropolitaine, qui commence au début de la saison froide et humide et termine à la fin de la saison chaude et sèche. En effet, les anomalies annuelles d'évaporation potentielle, dont les valeurs sont très stables d'hiver en hiver, sont principalement portées par ses variations estivales, c'est-à-dire à la fin de l'année hydrologique. Les précipitations saisonnières sont plus aléatoires d'une année à

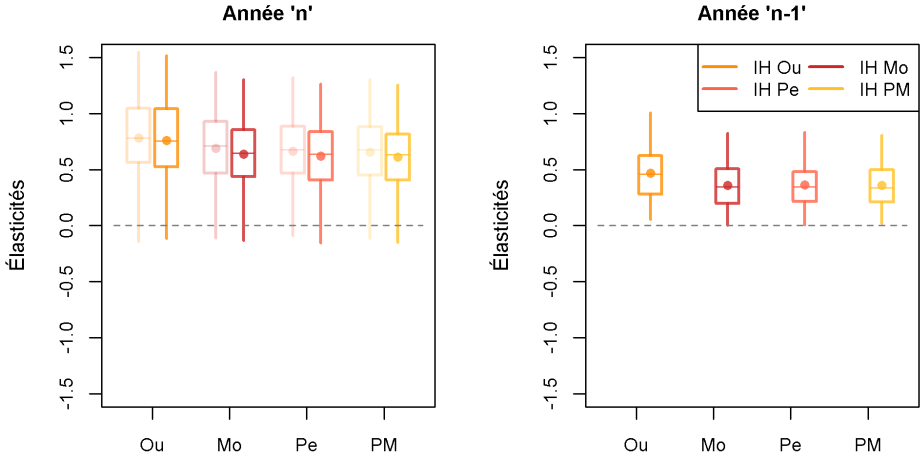


FIGURE 6.16 : Distributions des paramètres associés aux variables des régressions linéaires « B0 » et « B1 » aux années n et $n-1$, pour chacune des formules d'évaporation potentielle. Les paramètres des régressions linéaires sont équivalents aux élasticités du rendement des précipitations vis-à-vis de l'indice d'humidité IH (P/EP). Les boxplots aux contours transparents représentent les paramètres des régressions linéaires « B0 » et les boxplots aux contours opaques représentent les paramètres des régressions linéaires « B1 ». Les paramètres de l'indice d'humidité sont en couleurs chaudes, du jaune au rouge foncé, de sorte à indiquer les évaporations moyennes dans l'ordre croissant des valeurs moyennes d'évaporation potentielle sur le sous-échantillon.

l'autre et donc plus susceptibles d'expliquer l'écoulement moyen pendant une année hydrologique complète.
Pour étayer cette hypothèse, nous avons reproduit l'analyse des variations interannuelles des débits au moyen des régressions linéaires en considérant cette fois les années calendaires plutôt qu'hydrologiques. De fait, les résultats obtenus sont différents : dans ce cas les valeurs de l'élasticité du débit vis-à-vis de l'évaporation potentielle de l'année en cours sont majoritairement négatives, tandis que celles de l'année précédente sont plus proches de zéro (résultats non montrés). La méthode aboutit donc à des résultats contradictoires selon qu'elle soit appliquée aux années hydrologiques ou calendaires.
Pour expliquer cette contradiction, nous suggérons que l'agrégation des variables explicatives au pas de temps annuel conduit à une perte d'information trop importante des mécanismes saisonniers, qui sont inclus dans l'année en cours ou dans l'année précédente selon le calendrier choisi. Dans la suite, nous avons donc repris l'analyse à partir des modèles linéaires divisant l'année en saisons afin de considérer les effets des variations intra-annuelles de la réponse des bassins versants aux forçages climatiques.

6.2.3. Élasticité du débit vis-à-vis des variables climatiques saisonnières

6.2.3.a. Implémentation des modèles saisonniers

Les résultats précédents invitent à distinguer plusieurs saisons au sein d'une année pour l'estimation des relations entre débit annuel et forçages saisonniers. Toutefois, il serait déraisonnable d'alourdir les modèles avec un nombre trop important de variables explicatives (et donc de paramètres calés). Nous avons donc fait le choix de séparer l'année en seulement deux saisons, adaptées à chaque bassin à partir d'une classification des mois de l'année selon le rendement

mensuel des précipitations. Le choix du rendement des précipitations est justifié par la nécessité de définir deux saisons caractérisées par des réponses hydrologiques aussi hétérogènes que possible aux forçages climatiques. Le rendement des précipitations renseigne sur la part des précipitations effectivement transformée en écoulement, directement liée à l'élasticité du débit vis-à-vis des précipitations, et témoigne de l'état de sécheresse des sols et donc du ratio d'évaporation réelle. En procédant ainsi, nous avons donc cherché à valoriser au mieux l'information climatique pour l'explication des écoulements. La saison regroupant les mois au plus fort rendement des précipitations a été appelée saison « productive », tandis que la saison de bas rendement a été simplement qualifiée de saison « non productive ». L'analyse présentée dans la suite a été effectuée à partir des années hydrologiques et non plus calendaires.

La définition des saisons pour chaque bassin a été réalisée au moyen d'une classification ascendante hiérarchique, c'est-à-dire une méthode d'agrégation itérative des mois de l'année les plus similaires en termes de rendement des précipitations. À chaque mois de l'année a été associé un vecteur correspondant à l'ensemble des rendements observés à ce mois sur la chronique disponible. La distance de Canberra (Lance et Williams, 1967) a été choisie pour calculer la dissimilarité entre deux mois. Ce choix résulte d'un processus d'essais erreurs pour que l'algorithme reproduise une classification proche de celle réalisée manuellement par l'expérimentateur sur un échantillon de bassins tests. À chaque itération de la classification, la dissimilarité entre deux groupes de mois déjà formés a été évaluée comme la distance maximale entre les paires constituées d'un mois du premier groupe et d'un mois du deuxième groupe. Une fois la dissimilarité entre chaque groupe calculée, les deux groupes les moins dissimilaires ont été fusionnés, et ainsi de suite jusqu'à l'obtention de deux groupes distincts. La classification a été réalisée de sorte à ne pouvoir fusionner que des groupes de mois consécutifs dans l'année, afin de garantir la cohérence temporelle des saisons ainsi définies.
Le résultat de l'application de cette procédure est illustré en Figure 6.17 pour le bassin de la Meuse à Goncourt. Les mois définissant la saison productive se démarquent nettement par un fort rendement mensuel des précipitations. Certains mois montrent une plus grande variabilité interannuelle que d'autres (par exemple le mois de novembre). Ces mois sont généralement ceux situés à l'interface entre les deux saisons.

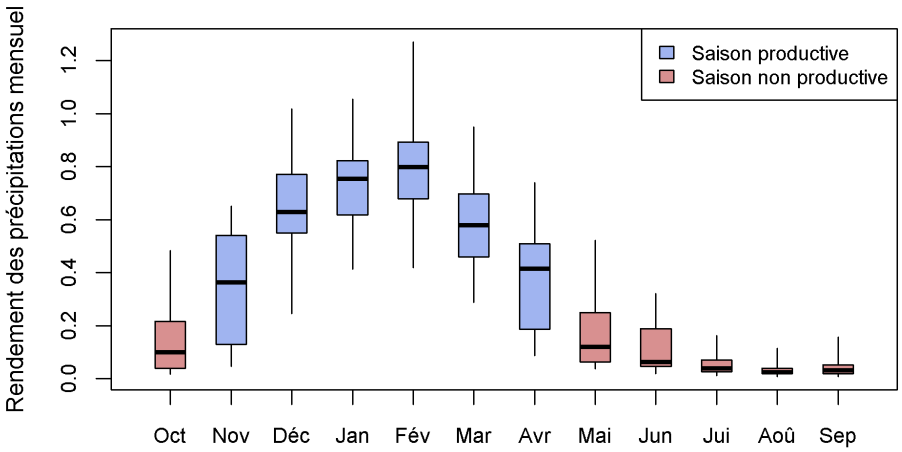


FIGURE 6.17 : Distributions des rendements mensuels des précipitations de la Meuse à Goncourt (B0220010). Les boîtes à moustaches bleues correspondent aux mois inclus dans la saison « productive », les boîtes à moustaches rouges correspondent aux mois inclus dans la saison « non productive ».

La répartition des mois en fonction des saisons sur le sous-échantillon de bassins français est récapitulée dans le Tableau 6.1. On identifie facilement les mois d'hiver froids et pluvieux entre

décembre et avril, et les étés chauds et secs entre juin et octobre.

TABLEAU 6.1 : Répartition (en %) des mois calendaires dans les saisons définies à partir de la classification ascendante hiérarchique, sur l'échantillon de bassins versants français.

Saison	Mois											
	Jan	Fév	Mar	Avr	Mai	Jun	Jui	Aoû	Sep	Oct	Nov	Déc
Productive	97.3	98.5	99.1	99.7	54.0	41.7	8.9	2.7	0.3	7.4	37.6	87.9
Non productive	2.7	1.5	0.9	0.3	46.0	58.3	91.1	97.1	99.7	92.6	62.4	12.1

Il convient de noter que les saisons ainsi définies ne s'ajustent pas nécessairement au calendrier hydrologique. Ainsi, en France, la saison « productive » commence au mois de décembre dans une majorité de bassins, c'est-à-dire deux mois après le début de l'année hydrologique. Nous avons fait le choix de respecter la définition de référence de l'année hydrologique en France et en Australie en agrégeant dans une même saison les données de mois non temporellement consécutifs dans les bassins où la saison « productive » ne commence pas en octobre en France, ou en juillet en Australie. Dans ces cas, une même saison a donc été constituée de quelques mois au début de l'année hydrologique et des quelques mois à la fin de l'année hydrologique.
Nous avons vérifié a posteriori l'influence de cette décision sur les résultats en reprenant les calculs pour l'année calendaire : ce choix n'a qu'une influence mineure sur les performances des modèles linéaires et n'a aucune influence sur les résultats de la comparaison des formules d'évaporation potentielle.

La division de l'année hydrologique en deux saisons nous a permis de distinguer les conditions climatiques des saisons productives et non productives comme des variables explicatives indépendantes des variations du débit :

$$\frac{\Delta Q_n}{\overline{Q}} = \epsilon^0_{P,n,p}\frac{\Delta P_{n,p}}{\overline{P}} + \epsilon^0_{EP,n,p}\frac{\Delta EP_{n,p}}{\overline{EP}} + \epsilon^0_{P,n,i}\frac{\Delta P_{n,i}}{\overline{P}} + \epsilon^0_{EP,n,i}\frac{\Delta EP_{n,i}}{\overline{EP}} \quad \text{« N0 », (6.7)}$$

$$\begin{aligned}\frac{\Delta Q_n}{\overline{Q}} &= \epsilon^1_{P,n,p}\frac{\Delta P_{n,p}}{\overline{P}} + \epsilon^1_{EP,n,p}\frac{\Delta EP_{n,p}}{\overline{EP}} + \epsilon^1_{P,n,i}\frac{\Delta P_{n,i}}{\overline{P}} + \epsilon^1_{EP,n,i}\frac{\Delta EP_{n,i}}{\overline{EP}} \\ &+ \epsilon^1_{P,n-1,p}\frac{\Delta P_{n-1,p}}{\overline{P}} + \epsilon^1_{EP,n-1,p}\frac{\Delta EP_{n-1,p}}{\overline{EP}} + \epsilon^1_{P,n-1,i}\frac{\Delta P_{n-1,i}}{\overline{P}} + \epsilon^1_{EP,n-1,i}\frac{\Delta EP_{n-1,i}}{\overline{EP}}\end{aligned} \quad \text{« N1 », (6.8)}$$

$$\frac{\Delta (Q/P)_n}{\overline{Q/P}} = \epsilon^0_{P/EP,n,p}\frac{\Delta (P/EP)_{n,p}}{\overline{P/EP}} + \epsilon^0_{P/EP,n,i}\frac{\Delta (P/EP)_{n,i}}{\overline{P/EP}} \quad \text{« B0 », (6.9)}$$

$$\begin{aligned}\frac{\Delta (Q/P)_n}{\overline{Q/P}} &= \epsilon^1_{P/EP,n,p}\frac{\Delta (P/EP)_{n,p}}{\overline{P/EP}} + \epsilon^1_{P/EP,n,i}\frac{\Delta (P/EP)_{n,i}}{\overline{P/EP}} \\ &+ \epsilon^1_{P/EP,n-1,p}\frac{\Delta (P/EP)_{n-1,p}}{\overline{P/EP}} + \epsilon^1_{P/EP,n-1,i}\frac{\Delta (P/EP)_{n-1,i}}{\overline{P/EP}}\end{aligned} \quad \text{« B1 », (6.10)}$$

Où les indices p et i indiquent respectivement les variables des saisons productives et improductives. La symbole $\epsilon^1_{EP,n-1,i}$ représente par exemple l'élasticité du débit vis-à-vis de l'évaporation potentielle de la saison non productive de l'année précédente, pour la régression linéaire « N1 ».

6.2.3.b. Élasticité des modèles « N » saisonniers

La distribution des élasticités du débit vis-à-vis des précipitations et de l'évaporation potentielle des années n et $n-1$ différenciées selon les deux saisons est représentée sur la Figure 6.18.

Comme prévu, le débit est plus sensible aux précipitations de la saison productive de l'année en cours qu'à celles de la saison non productive. Les précipitations des deux saisons de l'année $n-1$ ont néanmoins des élasticités assez similaires sur les débits, de même ordre de grandeur que celles estimées pour les régressions linéaires précédentes sur les variables annuelles. L'évaporation potentielle de la saison productive de l'année en cours joue un rôle important sur les débits, ce qui se comprend à l'aune de la saturation en eau élevée des sols en hiver, propice à un fort rendement de l'évaporation potentielle même lorsque la demande évaporatoire est faible. En revanche, l'élasticité du débit vis-à-vis de l'évaporation potentielle de la saison non productive est contre-intuitivement positive sur la majorité des bassins. Ces élasticités ne sont toutefois significatives que sur moins d'une dizaine de bassins de l'échantillon. L'évaporation potentielle de la saison non productive de l'année $n-1$ joue un rôle un peu plus important sur l'évolution interannuelle des débits, avec des valeurs d'élasticité majoritairement négatives. Malgré cela, cette influence est significative pour seulement une trentaine de bassins.

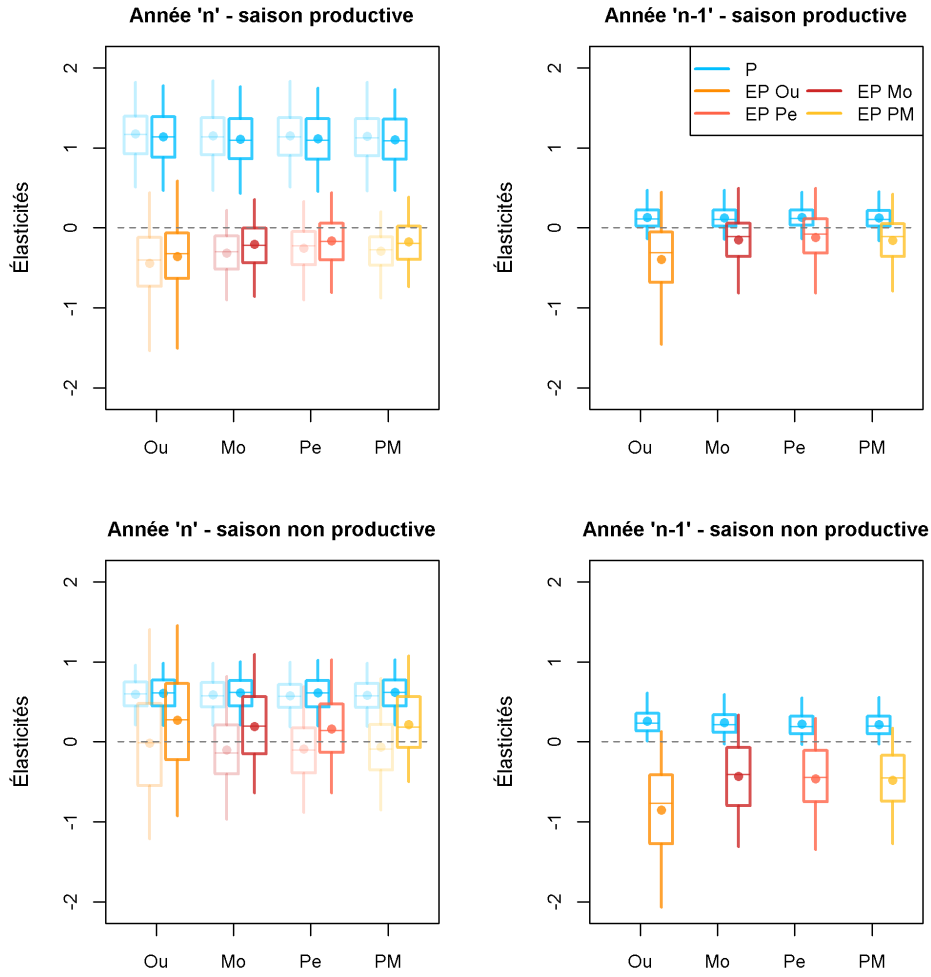


FIGURE 6.18 : Distributions des paramètres associés aux variables des régressions linéaires « N0 » et « N1 » aux années n et $n-1$ et aux saisons productive et non productive, pour chacune des formules d'évaporation potentielle. Les paramètres des régressions linéaires sont équivalents aux élasticités du débit vis-à-vis des précipitations ou de l'évaporation potentielle. Les boxplots aux contours transparents représentent les paramètres des régressions linéaires « N0 » et les boxplots aux contours opaques représentent les paramètres des régressions linéaires « N1 ». Les paramètres des précipitations sont en bleu. Les paramètres de l'évaporation potentielle sont en couleurs chaudes, du jaune au rouge foncé, de telle sorte à indiquer les évaporations moyennes dans l'ordre croissant des valeurs moyennes d'évaporation potentielle sur le sous-échantillon.

En définitive, les variations de l'évaporation potentielle ne semblent avoir qu'une faible influence sur les variations des débits, en comparaison de celles des précipitations.

6.2.3.c. Élasticité des modèles « B » saisonniers

Les élasticités obtenues avec les modèles « B » reliant le rendement annuel des précipitations à l'indice d'humidité des saisons précédentes montrent des motifs similaires à ceux observés avec les modèles « N » (Figure 6.19).
La saison productive de l'année en cours joue un rôle majeur dans le rendement annuel observé des précipitations, au contraire de la saison non productive, où l'influence de l'indice d'humidité sur le rendement n'est significative que sur une quarantaine de bassins. Les conditions climatiques de la même saison l'année précédente sont toutefois plus déterminantes : les valeurs d'élasticité sont clairement positives et significatives sur la moitié des bassins.
Il est intéressant de constater que la formulation du modèle « B1 » donne un poids plus conséquent aux conditions climatiques de la saison non productive de l'année $n-1$, en comparaison de l'influence des précipitations et de l'évaporation potentielle de cette saison dans le modèle « N1 ». Il est possible que la parcimonie de cette formulation limite les redondances entre les variables ayant un lien de corrélation, comme les précipitations et l'évaporation potentielle d'une même saison. Une augmentation des précipitations est en effet généralement associée à une baisse de la demande évaporatoire liée à la couverture nuageuse plus fréquente, à l'humidité de l'air, au rafraichissement du sol, etc. L'indice d'humidité combine les influences antagonistes des variations des précipitations et de l'évaporation potentielle, et donc évite que l'influence d'une variable soit effacée au profit de l'autre.

En résumé, nous avons constaté que les résultats des modèles linéaires de type « N » et « B » étaient cohérents dans l'ensemble, et que leur comportement était à peu près physiquement interprétable. Il a été nécessaire de détailler les forçages à l'échelle saisonnière pour réaliser que l'évaporation potentielle influence le débit de l'année hydrologique principalement pendant la saison productive, qui marque généralement le début de l'année hydrologique, et pendant la saison non productive de l'année précédente, généralement à la fin de l'année hydrologique. L'influence de l'évaporation potentielle sur les écoulements semble donc décalée d'une saison dans le temps en moyenne. Ce constat est peut-être lié au fait que les réserves d'eau du sol attaquées par l'évaporation jouent un rôle de filtre passe-bas de la réponse du bassin aux précipitations, et que les déficits ou surplus d'eau accumulée pendant une saison se ressentent généralement lors de la saison suivante.
Très peu d'écarts ont été observés entre les élasticités liées aux différentes formules d'évaporation, hormis pour celle d'Oudin, plus fortes et spatialement moins cohérentes du fait de sa faible variabilité interannuelle.

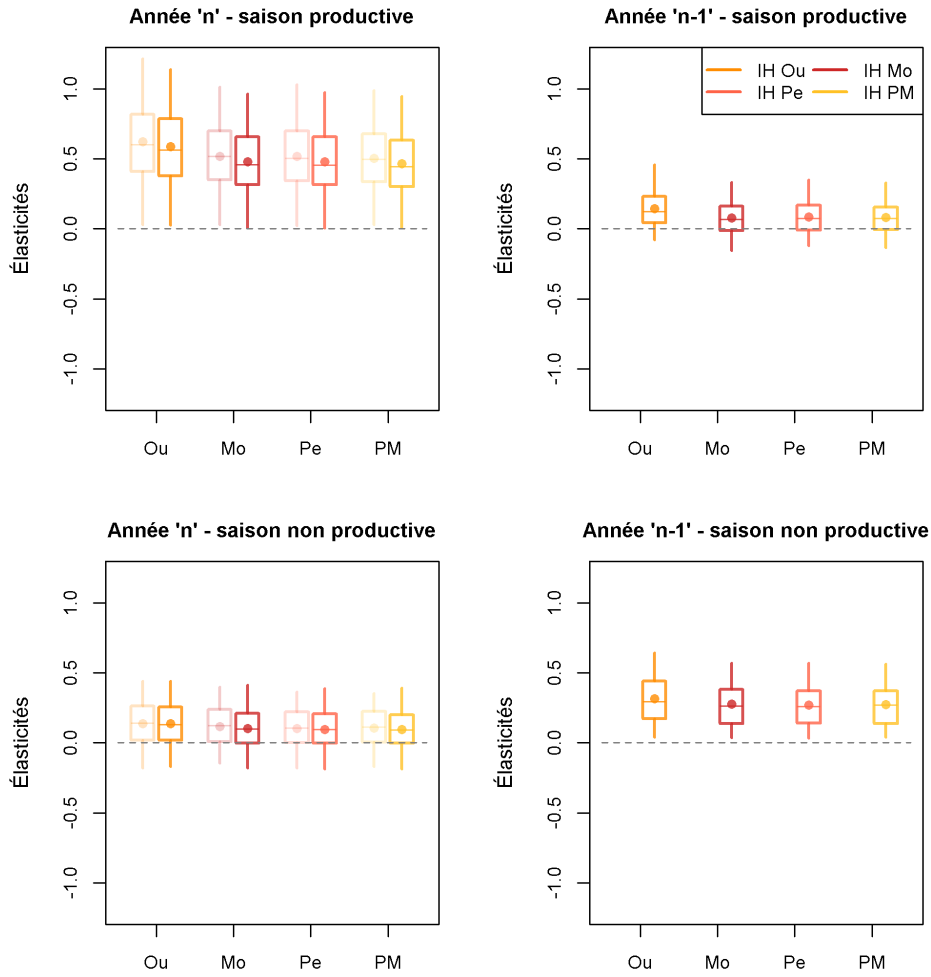


FIGURE 6.19 : Distributions des paramètres associés aux variables des régressions linéaires « B0 » et « B1 » aux années n et $n-1$ et aux saisons productive et non productive, pour chacune des formules d'évaporation potentielle. Les paramètres des régressions linéaires sont équivalents aux élasticités du rendement des précipitations vis-à-vis de l'indice d'humidité. Les boxplots aux contours transparents représentent les paramètres des régressions linéaires « N0 » et les boxplots aux contours opaques représentent les paramètres des régressions linéaires « N1 ». Les paramètres de l'indice d'humidité sont en couleurs chaudes, du jaune au rouge foncé, de telle sorte à indiquer les évaporations moyennes dans l'ordre croissant des valeurs moyennes d'évaporation potentielle sur le sous-échantillon.

6.2.4. Comparaison des performances des modèles linéaires saisonniers

Dans la suite, les performances des modèles testés sont discutées afin d'évaluer le potentiel explicatif des différentes formulations d'évaporation potentielle et de l'apport d'une mémoire aux modèles. Pour qu'elles soient comparables, les performances sont évaluées au regard du débit pour les deux types de modèles. On notera qu'exprimés de façon à expliquer le débit au lieu du rendement des précipitations, les modèles de type « B » ne sont pas des modèles linéaires.
Alors que les résultats n'ont jusque là été présentés que pour les bassins français, nous avons inclus en seconde partie de cette section ceux résultant de l'application de la méthode à l'échantillon de bassins australiens.

6.2.4.a. Bassins français

Les performances des deux types de modèles confirment l'importance de l'information portée par l'année hydrologique précédant l'année en cours (Figure 6.20). L'ajout de variables explicatives supplémentaires augmente nécessairement la part de variance expliquée et donc le coefficient de détermination d'un modèle linéaire, néanmoins on constate que cette augmentation est en moyenne de 10% de variance expliquée, ce qui n'est pas négligeable. Les tests réalisés sur des modèles ayant une mémoire jusqu'à $n-2$ n'ont pas révélé d'amélioration perceptible des performances (résultats non montrés).
En comparaison de l'ajout d'une mémoire aux modèles, le choix de la formule d'évaporation potentielle n'a presqu'aucune influence sur la distribution des performances des modèles linéaires sur l'échantillon de bassins versants.

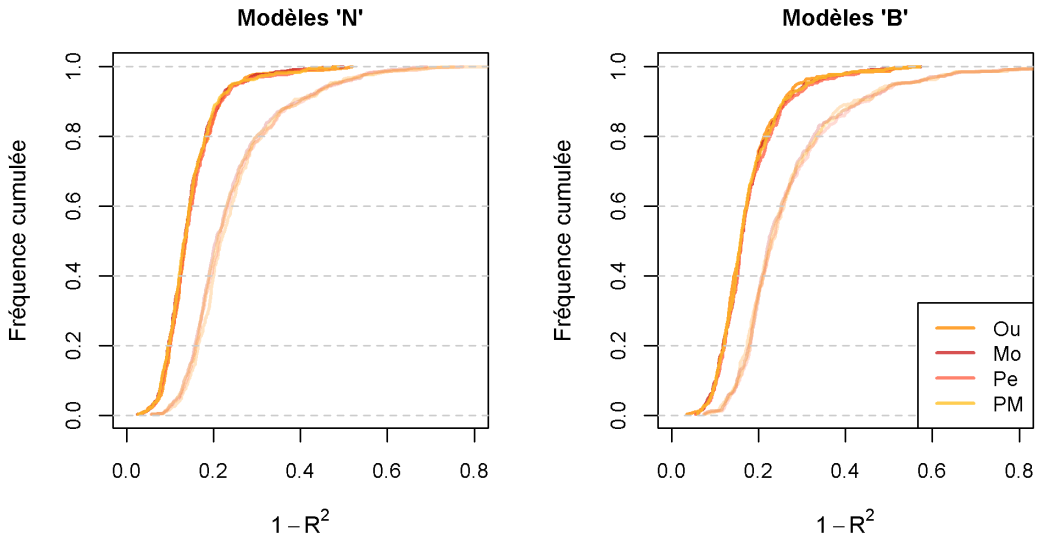


FIGURE 6.20 : Distributions des valeurs du coefficient de détermination des régressions linéaires « N » et « B », avec chacune des quatre formules d'évaporation potentielle sur les bassins du sous-échantillon français. Les couleurs du jaune au rouge foncé ont été choisies de telle sorte à indiquer les évaporations moyennes dans l'ordre croissant des valeurs moyennes d'évaporation potentielle sur le sous-échantillon. Les lignes transparentes représentent les performances des régressions linéaires « N0 » (ou « B0 ») et les lignes opaques représentent les performances des régressions linéaires « N1 » (ou « B1 »). Les valeurs du critère ont été calculées à partir de l'écart à la valeur optimale, de telle sorte qu'une simulation parfaite serait située tout à gauche des graphes.

Pour tenir compte de l'effet sur les performances de l'augmentation du nombre de variables expli-

catives dans les modèles à mémoire, les performances des modèles linéaires ont été jugées au regard du critère d'Akaike (1973), dénoté AIC, pour *Akaike Information Criterion*. Le critère d'Akaike est construit comme une combinaison linéaire de la log-vraisemblance des paramètres d'un modèle et du nombre de ses paramètres de sorte à pénaliser, à performances égales, les modèles les plus complexes. Les valeurs d'AIC les plus négatives indiquent les meilleurs modèles. Les distributions des AIC des modèles sur l'échantillon sont affichées en Figure 6.21.

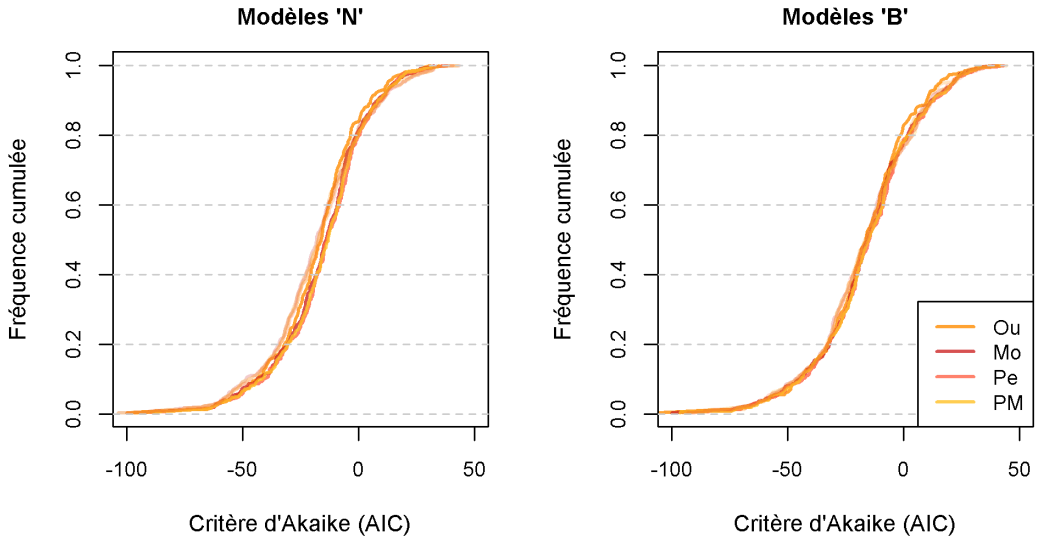


FIGURE 6.21 : Distributions des valeurs du critère d'Akaike des régressions linéaires « N » et « B », avec chacune des quatre formules d'évaporation potentielle sur les bassins du sous-échantillon français. Les couleurs du jaune au rouge foncé ont été choisies de telle sorte à indiquer les évaporations moyennes dans l'ordre croissant des valeurs moyennes d'évaporation potentielle sur le sous-échantillon. Les lignes transparentes représentent les performances des régressions linéaires « N0 » (ou « B0 ») et les lignes opaques représentent les performances des régressions linéaires « N1 » (ou « B1 »). Les valeurs d'AIC les plus négatives indiquent les meilleurs modèles.

Les distributions d'AIC ne permettent pas de mieux distinguer des écarts entre les formules d'évaporation. En revanche, les distributions des modèles sans mémoire chevauchent cette fois-ci celles des modèles à mémoire. Dans le cas des modèles « N », les moins parcimonieux, les modèles sans mémoire « N0 » surperforment même en moyenne légèrement par rapport aux modèles à mémoire « N1 ».

Constatant que des écarts n'étaient pas discernables dans les distributions pour l'identification des formules d'évaporation potentielle les plus adaptées à la représentation des dynamiques interannuelles du débit, nous avons comparé les performances bassin par bassin plutôt qu'en moyenne. Pour cela, nous avons calculé pour chaque bassin l'écart entre la log-vraisemblance du modèle linéaire « N1 » forcé par les quatre formules par rapport à la plus faible log-vraisemblance parmi ces quatre réalisations (Figure 6.22).
Les motifs obtenus sont les mêmes pour les modèles « N1 » et pour les modèles « B1 ». La formule d'Oudin est généralement associée aux meilleures performances. Celles-ci se distinguent d'ailleurs bien de celles de la formule de Morton, tandis que les écarts entre la formule de Morton, celle de Penman-Monteith et celle de Penman sont plus faibles. Lorsqu'elle n'est pas associée aux plus faibles valeurs de log-vraisemblance, la formule de Penman ne se démarque pas vraiment des formules concurrentes, en particulier d'après les résultats obtenus pour les modèles « N1 ». Il convient de remarquer que la formule d'Oudin est associée à la plus faible log-vraisemblance dans moins de 5% des bassins pour les modèles « N1 » et dans près de 25% pour les modèles « B1 », ce qui montre que cette formule est la pire dans une portion relativement restreinte des bassins de

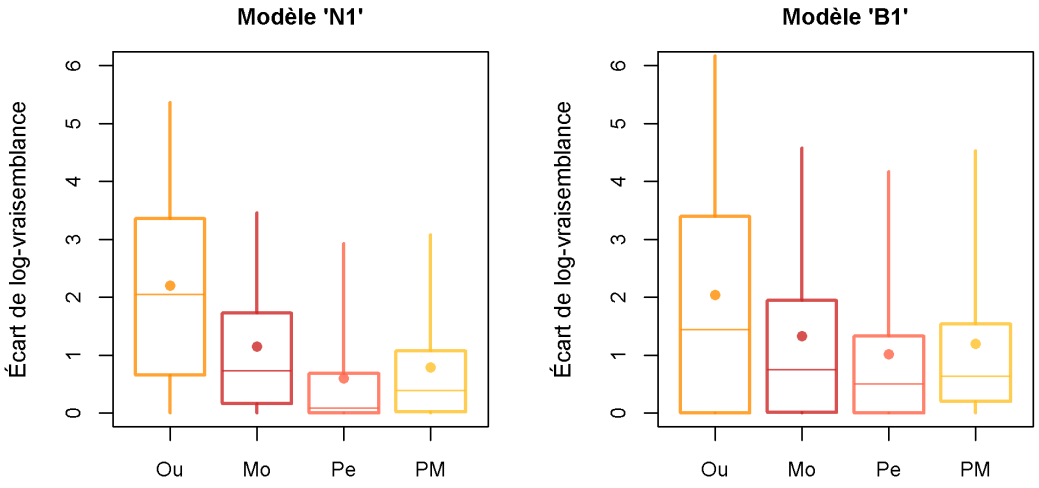


FIGURE 6.22 : Distributions des écarts de log-vraisemblance des régressions linéaires « N1 » et « B1 », entre chacune des quatre formules d'évaporation potentielle et la pire formule d'évaporation potentielle sur les bassins du sous-échantillon français. Les couleurs du jaune au rouge foncé ont été choisies de telle sorte à indiquer les évaporations moyennes dans l'ordre croissant des valeurs moyennes d'évaporation potentielle sur le sous-échantillon.

l'échantillon.
Nous pouvons donc conclure que la formule d'Oudin est la plus adaptée pour expliquer les évolutions interannuelles du débit sur le sous-échantillon français.

6.2.4.b. Bassins australiens

La méthode a également été appliquée à 161 bassins du sous-échantillon australien, choisis de sorte à pouvoir caler les modèles linéaires sur au moins 20 triplets d'années consécutives.
De même que pour les bassins français, deux saisons ont été définies pour chaque bassin à partir de l'algorithme de classification ascendante hiérarchique. La répartition des mois de l'année dans la saison productive et dans la saison non productive est montrée dans le Tableau 6.2.

TABLEAU 6.2 : Répartition des mois calendaires dans les saisons définies à partir de la classification ascendante hiérarchique, sur l'échantillon de bassins versants australiens.

Saison	Mois											
	Jan	Fév	Mar	Avr	Mai	Jun	Jui	Aoû	Sep	Oct	Nov	Déc
Productive	19.9	27.3	31.7	37.9	47.2	68.3	94.4	94.4	86.3	65.8	54.7	44.7
Non productive	80.1	72.7	68.3	62.1	52.8	31.7	5.6	5.6	13.7	34.2	45.3	55.3

Les performances des modèles linéaires « N » et « B », sans mémoire ou avec une année de mémoire, ont été évaluées avec le critère d'Akaike (Figure 6.23). À l'instar des résultats obtenus pour les bassins français, les distributions obtenues avec les deux formules d'évaporation se chevauchent sans qu'il soit possible de visualiser des différences claires. En revanche, les performances des modèles à mémoire dépassent clairement celles des modèles sans mémoire, en particulier pour les modèles « N ». Par ailleurs, les valeurs du critère d'Akaike des modèles « N1 » sont en moyenne plus faibles que celles associées aux modèles « B1 ». Cela indique donc que malgré le plus grand nombre de variables explicatives du débit pour les modèles « N1 » (8 variables contre 4 variables), caractère pénalisé par le critère d'Akaike, ceux-ci sont suffisamment

plus vraisemblables que les modèles « B1 » pour que leurs performances restent meilleures. Ce résultat suggère qu'il est important de décorréler les variations des précipitations et de l'évaporation potentielle pour expliquer les variations des débits australiens, plutôt que de les associer à travers les variations de l'indice d'humidité.

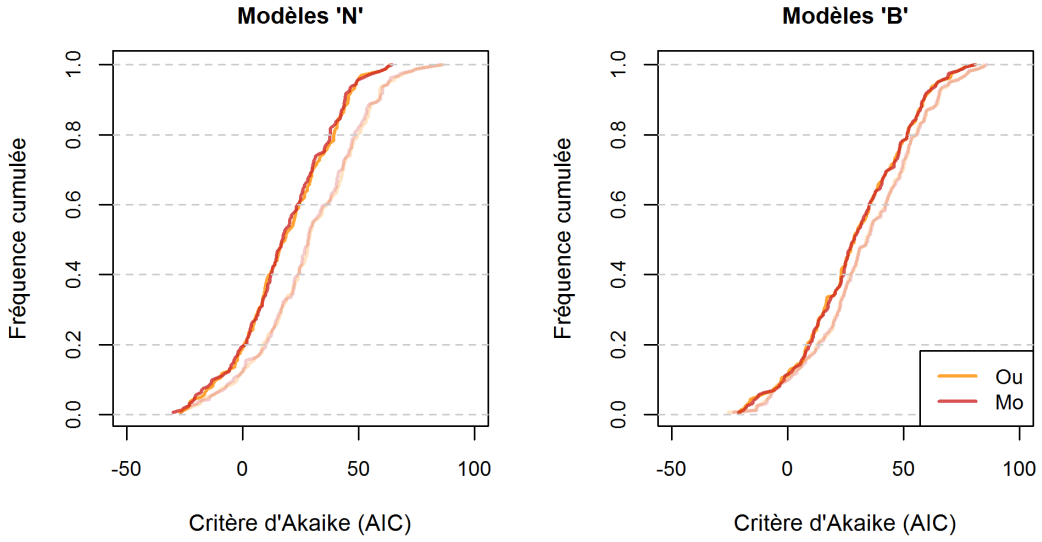


FIGURE 6.23 : Distributions des valeurs du critère d'Akaike des régressions linéaires « N » et « B », avec les deux formules d'évaporation potentielle sur les bassins du sous-échantillon australien. Les couleurs orange et rouge foncé indiquent respectivement les évaporations d'Oudin et de Morton. Les lignes transparentes représentent les performances des régressions linéaires « N0 » (ou « B0 ») et les lignes opaques représentent les performances des régressions linéaires « N1 » (ou « B1 »). Les valeurs d'AIC les plus négatives indiquent les meilleurs modèles.

La comparaison du potentiel explicatif des formules d'évaporation potentielle d'Oudin et de Morton est réalisée à travers l'évaluation des écarts de log-vraisemblance entre le meilleur et le moins bon modèle à une année de mémoire pour chaque bassin (Figure 6.24). Les deux formules se départagent clairement avec le modèle « N1 ». Contrairement au constat formulé dans le cas des bassins français, la formule de Morton obtient ici les meilleures performances. Celle d'Oudin est la moins bonne formule dans plus de la moitié des bassins étudiés.
Les résultats du modèle « B1 » ne sont pas caractérisés par une distinction aussi nette entre les deux formules. Généralement, les résultats sont plutôt à l'avantage de la formule d'Oudin. Nous suggérons que la pertinence limitée des modèles « B » en comparaison des modèles « N » sur les bassins australiens invite à donner moins de crédit à ces résultats.
Nous pouvons donc conclure que la formule de Morton est la plus adaptée pour expliquer les évolutions interannuelles du débit sur le sous-échantillon de bassins australien.

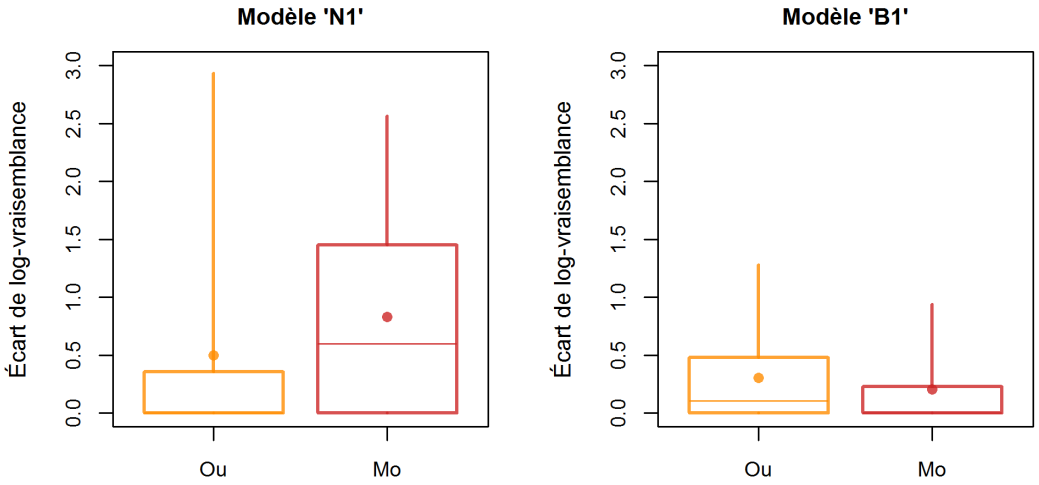


FIGURE 6.24 : Distributions des écarts de log-vraisemblance des régressions linéaires « N1 » et « B1 », entre chacune des deux formules d'évaporation potentielle et la pire formule d'évaporation potentielle sur les bassins du sous-échantillon australien. Les couleurs orange et rouge foncé indiquent respectivement les évaporations potentielles d'Oudin et de Morton.

6.2.5. Discussion des résultats concernant les formules d'évaporation potentielle

6.2.5.a. Interprétation de la robustesse des modèles hydrologiques forcés par les formules d'évaporation potentielle

Les conclusions de cette analyse concernant la comparaison des formules d'évaporation potentielle ne sont pas incohérentes avec celles dégagées par l'étude des performances générales des modèles hydrologiques, et en particulier vis-à-vis des distributions de PMR (Figure 6.1). Toutefois, la comparaison des formules d'évaporation potentielle à travers le prisme de la robustesse des modèles hydrologiques ne permettait pas de déterminer si les meilleurs résultats obtenus avec les formules d'Oudin et de Morton étaient uniquement liées aux problèmes de sous-estimation de l'évaporation potentielle par les formules de Penman-Monteith et d'Oudin. Notre analyse suggère ici que la meilleure robustesse des modèles hydrologiques forcés avec les formules d'Oudin et de Morton sur les bassins français, ou de Morton seulement sur les bassins australiens, peut également être expliquée par les meilleures co-variations des évaporations potentielles produites par ces formules avec les débits annuels.

Malgré ce résultat, il est possible que le facteur déterminant des performances des modèles hydrologiques soit en premier lieu l'amplitude de la demande évaporatoire plutôt que de la qualité de ses variations interannuelles.
Un premier argument à cette assertion est l'écart important entre les distributions de PMR des modèles hydrologiques sur les bassins australiens (Figure 6.2) selon qu'ils sont forcés avec la formule d'Oudin ou celle de Morton, à l'avantage de la deuxième. Alors que les deux formules ne se départagent pas nettement au regard des co-variations interannuelles avec le débit (Figure 6.24), il est en revanche apparu assez clairement que la formule d'Oudin sous-estimait la demande évaporatoire dans certains bassins. Étant donné que les valeurs de PMR ne dépendent pas du bilan hydrique moyen des modèles mais plutôt de ses variations, on peut considérer que cette sous-estimation est responsable de la moindre transférabilité des paramètres calés dans le cadre du GSST.
Dans le même registre, la supériorité des performances sur les bassins français des modèles hydro-

logiques forcés avec la formule de Penman par rapport à la formule de Penman-Monteith (Figure 6.1) n'est pas cohérente avec les meilleures co-variations interannuelles de l'évaporation de Penman-Monteith avec le débit par rapport à celle de Penman (Figure 6.22).
Enfin, les travaux récents de Jayathilake et Smith (2020a), où les auteurs comparent quatre formules d'évaporation potentielle de résolution temporelle variable, corroborent l'hypothèse selon laquelle l'amplitude de l'évaporation potentielle moyenne détermine plus fortement les paramètres de GR4J que la résolution temporelle des formules. En procédant à une validation croisée des formules d'évaporation, les auteurs suggèrent d'ailleurs qu'il est préférable de caler un modèle hydrologique avec des données d'évaporation potentielle n'intégrant pas les mécanismes ayant lieu à des échelles temporelles infra-mensuelles en validant avec des données d'évaporation de résolution temporelle plus fine, plutôt que l'inverse. Ces travaux ont l'intérêt de séparer ce qui relève de la phase de calage des paramètres de ce qui relève de la qualité des données utilisées pour évaluer les performances des modèles hydrologiques.

6.2.5.b. Liens entre performances et complexité des formules d'évaporation potentielle

La raison qui semble la plus évidente pour expliquer la répartition des performances des formules d'évaporation potentielle est le nombre de variables météorologiques nécessaires à leur calcul. La formule d'Oudin ne nécessite que la température moyenne et se démarque nettement de la formule de Morton qui requiert trois variables, c'est-à-dire seulement une de moins que les formules de Penman et de Penman-Monteith. On peut émettre l'hypothèse que la multiplication des variables météorologiques accroît le risque que des erreurs de mesure influencent l'estimation finale de l'évaporation potentielle. La non-linéarité des équations reliant ces variables est en outre susceptible d'accentuer les effets des biais systématiques présents dans les observations. Bien qu'il ne soit pas possible d'identifier précisément les variables les moins pertinentes, on peut suggérer que la très grande variabilité spatiale du vent dans les zones de relief suscite un doute légitime quant à l'intérêt de l'estimation de sa moyenne sur des mailles de plusieurs dizaines de kilomètres carrés.

On peut malgré tout nuancer ce propos en notant dans un premier temps que la supériorité de la formule de Morton sur la formule d'Oudin en Australie dément cette interprétation. Il a justement été pointé précédemment que la prise en compte d'autres variables que la température pourrait être bénéfique pour la qualité des estimations de la formule d'Oudin sur certains bassins. Il convient cependant de rappeler que les paramètres de la formule de Morton ont fait l'objet d'un calage sur les bassins australiens, ce qui leur confère un avantage certain.
Une seconde nuance importante à apporter à l'interprétation précédente sur le rôle du nombre des variables sur les performances des formules d'évaporation potentielle concerne la question des échelles de temps considérées. En effet, nos travaux sur les variations interannuelles des débits ont été réalisés à des échelles saisonnières et leurs conclusions ne sauraient certainement pas être extrapolés à des échéances plus courtes.
Enfin, il convient de remarquer que la formule de Penman-Monteith obtient de meilleurs résultats que la formule de Penman dans l'explication des variations interannuelles des débits, alors que la première se distingue justement par une dépendance accrue aux observations de vent à travers le ratio des résistances stomatique et aérodynamique.
En réalité, l'approche mise en œuvre manque d'un niveau de détail suffisant sur les mécanismes de l'évaporation potentielle pour qu'il soit possible d'approfondir ces questions à partir de nos résultats. Dans cette perspective, il serait par exemple intéressant de caractériser l'influence des précipitations sur l'humidité du sol et donc sur le rendement de l'évaporation potentielle, éventuellement à des échelles de temps plus fines que l'échelle saisonnière. Cela conduirait toutefois à formuler des hypothèses supplémentaires sur les processus hydrologiques, ce que l'on cherchait justement à éviter en implémentant cette méthode pour s'affranchir des effets de compensation des paramètres

des modèles hydrologiques.

6.2.5.c. Divergence des conclusions sur les deux sous-échantillons de bassins

Malgré les performances légèrement plus satisfaisantes obtenues sur les bassins français avec la formule de Morton comparée aux formules de Penman et de Penman-Monteith, notre étude aboutit à des résultats contradictoires sur les sous-échantillons de bassins français et australien. Les régimes hydrologiques très différents que connaissent les bassins des deux sous-échantillons invitent à ne pas considérer cette conclusion comme surprenante. L'échelle temporelle de prise en considération des forçages, de l'ordre du semestre, est vraisemblablement trop grossière pour tenir compte des spécificités régionales du processus d'évaporation. De ce fait, l'agrégation de la demande évaporatoire à l'échelle de la saison peut dans certains contextes représenter convenablement les processus à l'œuvre, grâce à la présence ou non de certains mécanismes dans l'estimation de l'évaporation potentielle, et ne pas être pertinente dans d'autres contextes. Les couverts végétaux typiques des bassins australiens et des bassins français sont par exemple susceptibles d'introduire des différences dans la contribution de l'interception à l'évaporation réelle totale. On peut supposer que dans des régions où l'interception joue un rôle prépondérant, il soit plus important de représenter convenablement les co-variations des précipitations et de la demande évaporatoire puisque l'évaporation réelle est directement dépendante des précipitations interceptées.

Nous suggérons également de nuancer les résultats concernant la formule de Morton en France et en Australie, dans la mesure où les paramètres de la formule ont fait l'objet d'un calage sur des données de bacs d'évaporation pour les bassins australiens, tandis les paramètres de référence estimés par Morton (1983) ont été retenus pour les bassins français. Il serait intéressant de procéder à un calage des paramètres sur les bassins français à partir de données spécifiques pour que la comparaison soit plus rigoureuse.
Il convient de rappeler que la formule d'Oudin étant empirique, elle a également fait l'objet d'un calage pour la modélisation hydrologique sur un échantillon de bassins français, australiens et américains (Oudin *et al.*, 2005a). Néanmoins, on peut considérer que la complexité des formulations de Morton, de Penman et plus encore de Penman-Monteith, rend un paramétrage unique pour chacune de ces formules compliqué à réaliser tant leurs paramètres interagissent entre eux (voir par exemple en Section 6.1.3.b ou la présentation de Besson *et al.* (2020) sur l'harmonisation par Météo-France du calcul de l'évaporation potentielle, diapositives 15 à 18.). Idéalement, une partie de ces paramètres devrait pouvoir être fixée au cas par cas à partir des caractéristiques physiques des bassins modélisés (émissivité et albédo de la surface, résistance stomatique). Mais ces caractéristiques sont difficiles à estimer, notamment parce qu'elles ne sont pas uniformes sur un bassin versant.

6.2.5.d. Difficulté d'évaluation des formules d'évaporation potentielle

Il est particulièrement difficile de choisir une approche pertinente pour évaluer les formules d'évaporation potentielle pour la modélisation hydrologique.
Une démarche extrême consiste à procéder de façon purement empirique en comparant directement les performances de modèles hydrologiques forcés avec des formules différentes : c'est l'approche mise en œuvre au début de ce chapitre. L'intérêt de cette démarche est qu'elle permet de tenir compte des interactions entre les forçages et les modèles hydrologiques, son inconvénient est l'interprétabilité limitée des résultats.
Une démarche opposée consisterait à évaluer les formules à partir d'observations de terrain de l'évaporation potentielle. L'intérêt principal de cette approche est qu'elle permet la comparaison directe des sorties des modèles d'évaporation potentielle à la grandeur observée en question. Néanmoins, il n'est pas garanti que la formule d'évaporation potentielle la plus réaliste soit également la plus adaptée pour la modélisation hydrologique, comme le suggèrent les travaux d'Oudin *et al.*

(2005b) déjà évoqués au Chapitre 2. Par ailleurs, il a été remarqué au Chapitre 1 que la mesure de l'évaporation potentielle telle qu'elle est définie dans les formules d'évaporation potentielle est très difficile et pose de nombreux problèmes méthodologiques. De ce point de vue, il n'est pas certain que les observations à utiliser pour valider une formule soient appropriées pour en valider une autre.

6.2.6. Tentative d'interprétation physique des effets de mémoire

L'étude des variations interannuelles du débit des bassins français a montré la faible importance du choix de la formule d'évaporation potentielle comparativement à la prise en considération de la mémoire des bassins. Nous avons cherché à interpréter ce résultat à partir du contexte hydrogéologique des bassins du sous-échantillon français (non disponibles pour les bassins australiens).
Pour cela, nous avons établi la liste des bassins pour lesquels le critère d'Akaike des modèles sans mémoire (N0 et B0) était meilleur que celui des modèles considérant une année de mémoire (N1 et B1), et avons reporté la position de ces bassins sur la carte hydrogéologique des zones très peu perméables du BRGM (Margat, 1980).

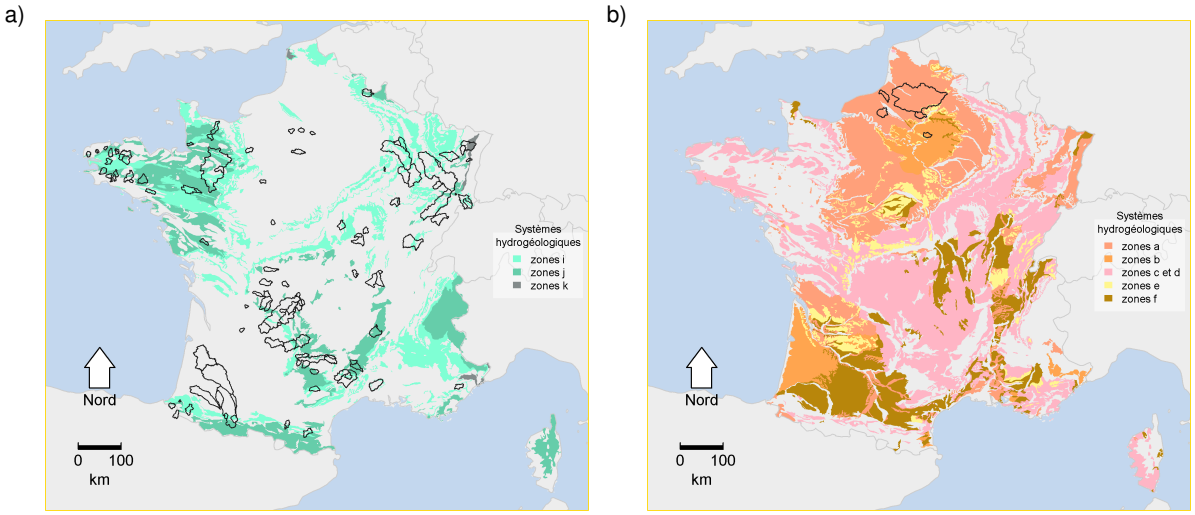


FIGURE 6.25 : a) Carte des 172 bassins « sans mémoire », d'après les régressions linéaires. La zone *i* indique les couvertures d'aquifères captifs profonds éventuels. La zone *j* indique les domaines sans nappe étendue, libre ou captive (socles schisteux, socles sous couverture non aquifère, massifs cristallins et schisteux des régions montagneuses à ruissellement très prédominant). La zone *k* indique les zones très complexes non divisibles à l'échelle de la carte (aquifères très locaux possibles). **b)** Carte des 4 bassins « à mémoire », d'après les régressions linéaires. La zone *a* indique les aquifères continus, à nappe libre, assimilables à une monocouche. La zone *b* indique les aquifères multicouche, à nappe supérieure libre. Les zones *c* et *d* indiquent les aquifères discontinus à surface libre, karstique ou volcanique, de structure tabulaire, ou de structure plissée, compartimentée, ou encore de roches cristallines fracturées. La zone *e* indique les couvertures semi-perméables, pouvant comporter des nappes perchées, des aquifères monocouches continus ou discontinus à nappe libre ou semi-captive. La zone *f* indique les formations semi-perméables capacitives encaissant des aquifères locaux, alluviaux surtout, qu'elles contribuent à alimenter, et/ou couverture semi-perméable de système multicouche sans nappe supérieure libre. D'après Margat (1980).

Nous avons procédé à l'identique pour l'établissement de la liste des bassins possédant une grande mémoire, en sélectionnant les bassins où les modèles considérant deux années de mémoire (N2 et B2) ont un meilleur critère d'Akaike que les modèles considérant seulement une année de mémoire (N1 et B1). La répartition géographique de ces bassins « à mémoire » a été confrontée à la localisation des systèmes aquifères à aire d'alimentation (Margat, 1980). Ces comparaisons

ont servi à tester l'hypothèse selon laquelle le climat d'une année influence l'hydrologie d'un bassin d'autant plus que la connectivité entre sa surface et ses capacités de stockage souterraines est grande. Les cartes obtenues sont présentées en Figure 6.25.

Les 121 bassins sans mémoire de l'échantillon sont situés dans des régions variées du territoire français. Une partie importante d'entre eux est effectivement localisée dans des zones très faiblement perméables, telles que le massif armoricain ou l'Aubrac. Les sous-bassins de l'Adour ne sont pas des zones très peu perméables mais plutôt semi-perméables encaissant quelques aquifères locaux. L'identification de ces groupes de bassins par notre méthode reste donc physiquement assez logique. Les frontières de certains bassins chevauchent plusieurs types de systèmes hydrogéologiques, dont la contribution respective à la physionomie des écoulements n'est pas évidente. Enfin, une poignée de petits bassins censés être « sans mémoire » sont présents dans des régions dont l'hydrologie est notoirement dominée par l'influence d'aquifères, comme le bassin parisien. Malgré cela, nos résultats tendent à corroborer l'hypothèse selon laquelle les bassins peu influencés par les conditions climatiques passées sont des bassins peu perméables et peu propices à l'accumulation de déficits et de surplus d'eau dans le souterrain.

Les 5 bassins identifiés comme des bassins « à mémoire » sont quant à eux regroupés dans le nord de la France, où affleure la craie. Le suitement de la craie alimente un écoulement de base très fort, dont il n'est pas surpenant qu'il dépende des apports pluviométriques cumulés des années précédentes. On peut finalement s'étonner qu'une portion si faible de l'échantillon des bassins français ait été identifiée comme caractérisant une influence marquée des conditions climatiques passées, compte tenu de l'étendue du territoire où la surface contribue à l'alimentation d'aquifères. Nous suggérons d'expliquer ce constat par la simplicité des modèles linéaires mis en œuvre, qui n'ont pas vocation à représenter toute la complexité des phénomènes souterrains. Le but de notre analyse étant avant tout de chercher à expliquer les variations des débits par celles des forçages climatiques, la méthodologie mise en place nous a conduits à passer à côté de nombreux bassins où une mémoire existe mais s'exprime de manière non triviale. On notera néanmoins qu'une analyse complémentaire des coefficients d'autocorrélation des débits annuels inclus dans la méthode des moindres carrés généralisés fait apparaître des bassins où les débits suivent clairement des variations de basse fréquence, comme le bassin de l'Eure à Louviers (H9501010), qui est également dominé par l'influence de nappes libres (résultats non montrés).

Enfin, nous ne disposions pas de suffisamment d'information sur l'hydrogéologie des bassins australiens pour mener une étude approfondie de la mémoire de ces bassins telle qu'estimée par les modèles linéaires. Nous avons toutefois remarqué que sur cet échantillon une part nettement plus grande de bassins versants (39 sur 161) était considérée comme possédant une mémoire des conditions climatiques des deux dernières années, tandis qu'une poignée seulement ne montrait pas de signes d'influence des conditions passées (4 sur 161). Nous attribuons ce résultats à l'occurence de la *Millenium Drought*, caractérisée par une baisse nette des précipitations et conséquemment des débits.

6.3. Synthèse

Conclusions du Chapitre 6

Objectifs

Les travaux menés dans ce chapitre avaient pour objectif de déterminer la meilleure formule d'évaporation potentielle pour la modélisation hydrologique en conditions climatiques changeantes. Les quatre formules d'évaporation potentielle testées ont été décrites au Chapitre 3, et incluent les formulations d'Oudin, de Morton, de Penman et de Penman-Monteith. Seules les deux premières ont pu être calculées pour les bassins australiens. Ces formules représentent un échantillon d'approches de complexité variée, parmi les plus utilisées pour la modélisation hydrologique en France et en Australie.
La comparaison a dans un premier temps été réalisée à partir des résultats du GSST afin d'examiner les performances des modèles hydrologiques dans des expériences de calage-contrôle. Une méthode simple fondée sur des régressions linéaires a été mise en œuvre dans une deuxième partie pour évaluer la capacité des différentes formules à expliquer les variations interannuelles du débit.

Comparaison des évaporations potentielles à travers les performances des modèles hydrologiques

Au regard des critères d'évaluation étudiés, la robustesse des modèles hydrologiques est relativement peu sensible au choix de la formule d'évaporation potentielle. Certaines formules se démarquent néanmoins par des performances moyennes plus faibles. Ainsi, la formule de Penman-Monteith est associée sur le sous-échantillon français aux simulations les plus biaisées et les moins proches des observations en hautes eaux comme en étiages. Les formules d'Oudin et de Morton permettent au contraire d'avoir les meilleurs résultats pour la majorité des critères analysés. Sur le sous-échantillon australien, les performances atteintes avec la formule de Morton sont cependant nettement meilleures que celles obtenues avec la formule d'Oudin.
L'étude de la sensibilité aux conditions de calage des performances en contrôle tend à corroborer les conclusions précédentes en France : les formules d'Oudin et de Morton minimisent la dépendance des performances des modèles au choix de la période de calage. Les résultats constatés en Australie sont cependant moins univoques.

Interprétation des résultats au regard du bilan hydrique moyen

Pour comprendre les résultats précédents, nous avons d'abord analysé l'influence des valeurs moyennes différentes des évaporations. La formule de Penman-Monteith produit en effet les évaporations moyennes les plus faibles en France, suivies par celles estimées avec la formule d'Oudin. La formule d'Oudin produit les évaporations les plus faibles en Australie.
Nous avons examiné en premier lieu les rendements mensuels de l'évaporation potentielle simulés par les modèles, c'est-à-dire les ratios mensuels de l'évaporation réelle rapportée à l'évaporation potentielle. Sur une partie non négligeable du sous-échantillon de bassins français, ces rendements sont constamment égaux ou très proches de 1 pour les modèles SimHyd et TOPMO forcés avec l'évaporation de Penman-Monteith et d'Oudin. Cela signifie que leurs paramètres ont été calés de telle sorte que les modèles évaporent à taux potentiel à chaque pas de temps. Ce comportement s'explique par une évaporation potentielle moyenne insuffisante pour pouvoir fermer le bilan hydrique des bassins simulés : les modèles tentent donc de se rapprocher de l'équilibre en

évaporant à un taux très proche du taux potentiel. Ce phénomène n'est pas constaté pour GR4J, qui dispose d'une fonction d'échanges souterrains pour équilibrer son bilan. Ainsi, dans l'hypothèse où le recours aux échanges souterrains ne saurait être justifié pour tous les bassins concernés, la formule de Penman-Monteith, et dans un moindre mesure la formule d'Oudin, sous-estiment l'évaporation potentielle sur certains bassins français. Le constat est similaire pour la formule d'Oudin sur l'échantillon de bassins australiens.
Dans le cas de la formule de Penman-Monteith, les effets de la correction de l'albédo lors des chutes de neige et de la résistance stomatique s'additionnent et semblent être responsables de cette sous-estimation. La formule d'Oudin apparaît quant à elle trop faible dans certains bassins forestiers.

Interprétation des résultats au regard des variations interannuelles du débit

Nous avons également cherché à évaluer la capacité des formules d'évaporation à expliquer les variations interannuelles du débit sans recourir à la modélisation hydrologique. Pour cela, nous avons comparé les résultats de régressions linéaires des variations interannuelles des débits par les variations saisonnières des forçages, pour les différentes formules d'évaporation potentielle. Deux saisons ont été définies pour chaque bassin à partir d'une classification des rendements mensuels des précipitations. Les régressions linéaires ont été formulées en intégrant ou non les conditions climatiques de l'année précédente afin d'introduire les effets de l'accumulation des surplus ou des déficits en eau.
Nos résultats montrent que les performances des modèles linéaires sont assez insensibles au choix de la formule d'évaporation potentielle. Ces performances semblent au contraire sensibles à la prise en compte des conditions climatiques des années précédentes. La formule d'Oudin explique néanmoins légèrement mieux les variations interannuelles des écoulements en France que les formules de Morton, de Penman et de Penman-Monteith. Inversement, en Australie la formule de Morton est plus pertinente que la formule d'Oudin.
Ces résultats sont cohérents avec ceux observés à travers les performances des modèles hydrologiques dans le cadre du GSST, ce qui indique que les meilleures formules produisent des estimations de l'évaporation potentielle à la fois d'un ordre de grandeur suffisant pour que tous les modèles puissent fermer leur bilan hydrique et dont les variations interannuelles sont cohérentes avec celles du débit.

Hydrogéologie des bassins ayant une mémoire des conditions climatiques passées

L'analyse des variations interannuelles du débit a révélé quelques caractéristiques intéressantes concernant la mémoire des bassins versants. Les distributions des coefficients d'élasticité du débit vis-à-vis de l'évaporation potentielle montrent que l'évaporation potentielle influence les débits annuels des bassins français pendant la saison productive de l'année en cours et pendant la saison non productive de l'année précédente. Les bassins versants de l'échantillon français intègrent donc les variations saisonnières de l'évaporation potentielle avec une saison de retard en moyenne.
Tous les bassins n'ont en revanche pas la même mémoire des conditions climatiques passées. Les écoulements interannuels de certains bassins versants sont ainsi mieux représentés par les modèles sans mémoire que les modèles avec une année de mémoire. D'autres au contraire bénéficient d'une prise en compte de la mémoire des deux années précédentes. L'examen de la répartition des bassins versants français associés aux modèles sans ou avec mémoire montre une certaine cohérence avec leurs caractéristiques hydrogéologiques. Ainsi, les bassins sans mémoire sont généralement situés dans des zones de très faible perméabilité, tandis que les bassins bénéficiant de la prise en compte de la mémoire sont plutôt localisés dans des régions où les surfaces alimentent des aquifères libres.

Choix d'une formule d'évaporation potentielle unique pour le reste des travaux

Au regard des résultats et des discussions menées dans ce chapitre, le choix d'une formule générique d'évaporation potentielle pour la modélisation hydrologique semble être la formule d'Oudin pour le sous-échantillon français et la formule de Morton pour le sous-échantillon australien. Néanmoins, les développements de la suite de la thèse impliquent uniquement le modèle GR4J, dont nous avons constaté que les performance étaient peu affectées par les entrées d'évaporation potentielle. À ce titre, il peut être avantageux d'employer une formule unique pour les deux sous-échantillons de bassins versants, de sorte à réduire le nombre des facteurs d'interprétation des résultats des chapitres suivants.
Dans toute la suite, la formule d'Oudin a donc été retenue comme la formule de référence pour calculer l'évaporation potentielle. La formule est en effet la plus indiquée pour la modélisation hydrologique sur les bassins versants français. Sur les bassins, notamment australiens, où la formule sous-estime la demande évaporatoire, l'effet de cette sous-estimation est compensé par la capacité de GR4J à régler son bilan par la fonction d'échanges souterrains. Étant donné les bons résultats obtenus avec la formule de Morton sur les bassins français, nous aurions pu faire le choix de cette formule plutôt que celle d'Oudin. Toutefois la formule de Morton n'a à notre connaissance presque jamais été utilisée en France hormis par Oudin (2004), ce qui est susceptible de nuire à l'appropriation de nos travaux par d'autres hydrologues. Par ailleurs, les paramétrages différents de la formule en France et en Australie diminuent l'intérêt de l'usage d'une formule unique sur l'ensemble de l'échantillon.

Conclusions de la Partie II

Le diagnostic de la robustesse de trois modèles hydrologiques mis en œuvre dans la Partie II a été réalisé en évaluant l'influence qu'avaient certains choix de simulation sur leurs performances des modèles en calage-contrôle.

L'analyse a d'abord porté sur le choix de la fonction objectif utilisée pour réaliser le calage, celui de l'algorithme de calage, et sur les types de changements climatiques rencontrés dans les bassins versants étudiés. Le choix de la fonction objectif a clairement été identifié comme un facteur non négligeable affectant la robustesse des modèles considérés. Le KGE appliqué aux débits non transformés, régulièrement utilisé dans la littérature scientifique concernant la robustesse des modèles hydrologiques, est notamment apparu comme assez peu robuste. La simple adaptation du KGE à la racine carrée des débits (« KGE racine ») s'est révélé au contraire un choix pertinent.
Néanmoins, l'emploi du KGE racine ne bénéficie que marginalement aux performances des modèles en contrôle, ce qui suggère que les perspectives d'amélioration de la robustesse des modèles hydrologiques par des techniques de calage plus adaptées restent limitées.

Les travaux réalisés dans cette Partie II sur le choix de l'évaporation potentielle aboutissent à une conclusion similaire : il est possible d'identifier des formules d'évaporation potentielle plus appropriées que d'autres pour la modélisation hydrologique en climat changeant, mais la marge d'amélioration semble également assez limitée. En outre, contrairement aux résultats concernant les fonctions objectif, la hiérarchie des formules d'évaporation potentielle est assez dépendante du contexte de modélisation. La formule d'Oudin semble dans l'ensemble bien convenir à la modélisation des bassins français, mais sous-estime la demande évaporatoire dans les bassins forestiers australiens. Bien que la formule de Morton produise des résultats satisfaisants à la fois dans les conditions françaises et australiennes, il est difficile de conclure définitivement sur son potentiel d'utilisation en général puisque les bonnes performances constatées pour les bassins australiens sont au moins en partie dûes au calage spécifique de ses paramètres pour le contexte

australien.

Les résultats obtenus dans cette partie suggèrent donc que la recherche de la réduction des incertitudes liées aux paramètres et aux entrées d'évaporation potentielle n'est pas suffisante pour améliorer la robustesse des modèles face à des conditions climatiques variables. Il semble de fait inévitable de questionner les limites des modèles eux-mêmes.
La suite des développements a été focalisé sur le modèle GR4J, qui est à la fois le moins complexe et le plus robuste des modèles étudiés dans les chapitres précédents. Dans un premier temps, nous avons conçu et appliqué une méthode de diagnostic censée identifier les lacunes structurelles limitant la robustesse du modèle. Les résultats de cette analyse ont ensuite servi à proposer un amendement de la structure du modèle.

Partie III

Évaluation des limites structurelles du modèle GR4J

Chapitre 7.

Méthode d'évaluation des limites structurelles d'un modèle hydrologique

Introduction

Ce chapitre présente une méthode d'évaluation des limites structurelles des modèles hydrologiques conceptuels, destinée à aider l'identification des composantes réduisant la capacité du modèle à satisfaire plusieurs critères de performance et de robustesse simultanément. La méthode est fondée sur une tentative de rationalisation des approches multi-critères visant à permettre leur application à des grands échantillons de bassins versants.
Après une description de son principe, la méthode est mise en œuvre pour le modèle GR4J avec un ensemble de critères de performance déterminant la capacité du modèle à simuler simultanément et sans erreur le bilan hydrique des bassins versants et les variations interannuelles de ce bilan sur plusieurs gammes de débit. L'évaluation porte sur la capacité de GR4J à être à la fois non biaisé et robuste sur ces différentes gammes de débit, et est suivie par une analyse des paramètres et des flux internes du modèle afin de comprendre le comportement du modèle et d'identifier les composantes et paramétrages limitants.

7.1. Principes de la méthode

7.1.1. Intérêt de la méthode

7.1.1.a. Problématique

L'évaluation des modèles et de leurs paramètres repose généralement sur la capacité des modèles à reproduire différents aspects du régime hydrologique d'un bassin. De nombreux critères d'évaluation cherchent à estimer cette capacité de façon synthétique en agrégeant l'erreur journalière du modèle. Ces critères se comportent néanmoins à certains égards comme des boîtes noires et masquent certaines faiblesses des simulations. Certaines stratégies d'évaluation sont donc fondées explicitement sur un ensemble de critères distincts caractérisant chacun un aspect plus précis des performances du modèle. Un bon modèle est censé obtenir des performances satisfaisantes au regard de tous les critères considérés. Une approche souvent adoptée consiste à explorer l'espace des paramètres d'un modèle, par exemple à partir de tirages Monte Carlo, et à vérifier si certains jeux de paramètres respectent des limites d'acceptabilité de performances prédéfinies pour l'ensemble de ces critères.
Il arrive toutefois que la structure d'un modèle soit telle qu'il n'est pas possible d'optimiser les performances vis-à-vis d'un critère sans que cela se fasse au détriment d'un autre. Dans les situations où certains critères rentrent en forte compétition les uns par rapport aux autres, la recherche d'un jeu de paramètres optimisant tous ces critères simultanément se heurte donc à l'inexistence d'un tel jeu de paramètres. Dans ces conditions, l'optimisation des paramètres du modèle nécessite la détermination d'un compromis entre ces critères. En d'autres termes, caler le modèle revient à choisir entre les jeux de paramètres optimaux au sens de Pareto, c'est-à-dire tels qu'aucun jeu de paramètres ne les surclasse vis-à-vis de tous les critères à la fois (Section 2.2.2.b).

Pour améliorer les structures des modèles, il peut être intéressant de repérer quels critères sont en compétition et identifier les paramètres et composantes structurelles impliquées dans ces compétitions. Certains auteurs ont ainsi proposé des méthodes d'évaluation multi-critères afin de discriminer plusieurs structures de modèles, et déterminer les types de structures les plus performants (Euser *et al.*, 2013; Coxon *et al.*, 2014). D'autres, tels que Kavetski et Fenicia (2011), ont procédé

de même en complexifiant de manière incrémentale une structure de base pour déterminer les composantes minimales nécessaires. Leurs résultats suggèrent notamment qu'aucune structure unique n'est adaptée à toutes les situations de modélisation à la fois.
À notre connaissance, il n'existe pas de méthode permettant de critiquer et identifier les composantes les moins adaptées d'une structure de modèle unique sur un grand nombre de bassins versants. Si les méthodes multi-critères peuvent être appliquées à un modèle sur quelques bassins versants, elles supposent une analyse au cas par cas des compétitions entre les critères considérés.

7.1.1.b. Objectifs

La méthode que nous proposons est une tentative de rationalisation de l'évaluation multi-critères. L'objectif principal est d'aider l'identification robuste des composantes limitant la polyvalence d'un modèle, c'est-à-dire ses performances au regard de plusieurs critères simultanément. Pour cela, la méthode repose d'une part sur l'estimation d'un indicateur de la capacité du modèle à satisfaire plusieurs critères de performance simultanément, appelée « polyvalence du modèle », et d'autre part sur l'estimation de la possibilité d'optimiser simultanément des groupes de critères, afin d'identifier les critères les moins « compatibles » entre eux.

L'approche mise en œuvre peut ainsi être divisée en plusieurs étapes successives :

- Explorer l'espace des paramètres pour dessiner les fronts de Pareto entre les critères de performance ;
- Estimer la polyvalence du modèle ;
- Estimer les pertes de performance liées aux compétitions ;
- Identifier des groupes de critères de performances incompatibles entre eux ;
- Expliquer ces incompatibilités à partir des paramètres et des flux du modèle obtenus en optimisant chacun des groupes des critères.

Avant de détailler les étapes et concepts de la méthode, il convient d'indiquer que dans la mesure où la méthode n'est pas destinée à comparer des modèles entre eux mais à guider le diagnostic d'un modèle seul, elle est fondée sur une comparaison relative des performances entre elles, plutôt que sur une échelle absolue de limites d'acceptabilité. Ainsi, la capacité d'un modèle à satisfaire simultanément plusieurs critères de performance a été évaluée au regard des meilleures performances atteignables par le modèle pour chacun de ces critères.
Enfin, l'objectif de cette thèse étant l'évaluation de la robustesse des modèles face à des conditions climatiques variables, les critères évalués ont été choisis pour s'inscrire dans cette problématique. Bien que cette méthode puisse en théorie être appliquée à n'importe quel modèle et à n'importe quel ensemble de critères de performance, les conditions de son application seront discutées à la fin de ce chapitre.

7.1.2. Description de la méthode

7.1.2.a. Critères de performance considérés

Les critères utilisés pour le diagnostic de la structure diffèrent de ceux utilisés pour discriminer les performances des modèles aux Chapitres 5 et 6. L'objectif est à présent moins de donner une vision d'ensemble des performances des modèles dans des conditions de simulations variées que de permettre de comprendre les raisons pour lesquelles il semble si difficile d'aplatir la courbe des biais glissants d'un modèle hydrologique. Nous rappelons qu'en addition aux travaux de Coron *et al.* (2014), nous avons constaté, à travers l'étude du proxy de la robustesse des modèles (PMR) aux Chapitres 4 et 5, que l'aire sous la courbe des biais glissants était relativement peu sensible au

choix du jeu de paramètres.

Pour sélectionner les critères d'application de la méthode, nous avons dans un premier temps considéré que la possibilité d'aplatir la courbe des biais glissants devait être évaluée à une condition : celle de ne pas sacrifier le biais du modèle au profit de la diminution des erreurs relatives d'une période à une autre. En d'autres termes, il convient d'évaluer la capacité du modèle à reproduire les variations interannuelles des débits tout en conservant une estimation satisfaisante des débits moyens.
Dans un second temps, il est apparu qu'une autre contrainte importante était la nécessité d'éviter que l'amélioration des performances à reproduire certaines catégories d'écoulement (par exemple en crue) ne se fasse au détriment des performances dans d'autres catégories d'écoulement (par exemple en étiage). Nous avons donc pris le parti d'évaluer la capacité du modèle hydrologique à représenter simultanément la moyenne et les variations interannuelles du débit pour trois gammes de débit : les basses eaux, les hautes eaux, et les eaux intermédiaires entre hautes et basses eaux.

Ces considérations ont conduit à impliquer six critères d'évaluation dans l'analyse des limites structurelles du modèle hydrologique : trois critères de biais sur la période totale disponible des bassins versants de l'échantillon caractérisant les erreurs volumiques du modèle sur les trois gammes de débit, et trois critères correspondant au PMR appliqué aux trois gammes de débit. De même que dans les chapitres précédents, les biais ont été comptabilisés de manière absolue. Les six critères d'évaluation sont réunis dans le Tableau 7.1.

TABLEAU 7.1 : Critères de performance utilisés dans l'analyse des limites structurelles du modèle hydrologique.

Symbole	Définition	Optimum
$Biais_B$	Erreur volumique moyenne sur les basses eaux (s.u.)	0
PMR_B	Aire sous la courbe des biais glissants sur les basses eaux (s.u.)	0
$Biais_I$	Erreur volumique moyenne sur les eaux intermédiaires (s.u.)	0
PMR_I	Aire sous la courbe des biais glissants sur les eaux intermédiaires (s.u.)	0
$Biais_H$	Erreur volumique moyenne sur les hautes eaux (s.u.)	0
PMR_H	Aire sous la courbe des biais glissants sur les hautes eaux (s.u.)	0

La définition de gammes de débit pour chacun des bassins de l'échantillon a été réalisée à partir de considérations sur le cumul des débits classés de chaque bassin.
La gamme des basses eaux a été définie comme l'ensemble des débits observés les plus bas dont le cumul est égal à 10% du total des débits cumulés (Figure 7.1.a). L'équivalent de la gamme des basses eaux dans le débit simulé a été défini comme l'ensemble des débits simulés les plus bas tel que le nombre de ces valeurs de débit simulé soit égal au nombre de valeurs de débit observé (Figure 7.1.b).
La gamme des hautes eaux a été définie de la même manière, en considérant cette fois l'ensemble des débits les plus forts dont le cumul est égal à 50% du total des débits cumulés (Figure 7.1.a). De même, la gamme des hautes eaux du débit simulé a été définie de sorte à sélectionner un nombre de valeurs égal à celui de la gamme des hautes eaux du débit observé (Figure 7.1.b). Les débits observés et simulés constituant la gamme des eaux intermédiaires sont les débits n'appartenant ni à la gamme des basses eaux, ni à la gamme des hautes eaux.

Cette définition a été motivée par plusieurs raisons. D'une part, la définition de gammes de débit à partir d'une limite fixe et uniforme pour tous les bassins de la part des débits inclus dans les trois gammes de débit (par exemple 50% des débits les plus faibles, 10% des débits les plus forts, et les 40% restant respectivement pour les gammes des basses eaux, des hautes eaux, et des eaux

intermédiaires) n'est pas bien adaptée à la diversité des régimes hydrologiques rencontrée dans les bassins de l'échantillon. D'autre part, le choix d'une valeur fixe du total des débits cumulés pour la définition des gammes permet de garantir que les erreurs du modèle sur chacune des gammes considérées ont une influence quantifiable sur l'erreur totale du modèle si les écoulements n'avaient pas été divisés en plusieurs gammes. Le choix d'une part de 10% des débits cumulés totaux pour la gamme des basses eaux assure notamment que l'erreur considérée n'est pas négligeable dans le bilan du modèle. Celui des 50% des débits cumulés pour constituer la gamme des hautes eaux est motivé par la participation importante qu'ont les hautes eaux dans le cumul des débits des bassins et permet d'inclure un nombre suffisant de valeurs dans la gamme.
Ces choix ont fait l'objet d'une justification plus détaillée et d'une discussion approfondie, reportée en Annexe B, dans la mesure où ce sujet ne s'inscrit pas directement dans la problématique de ce chapitre.

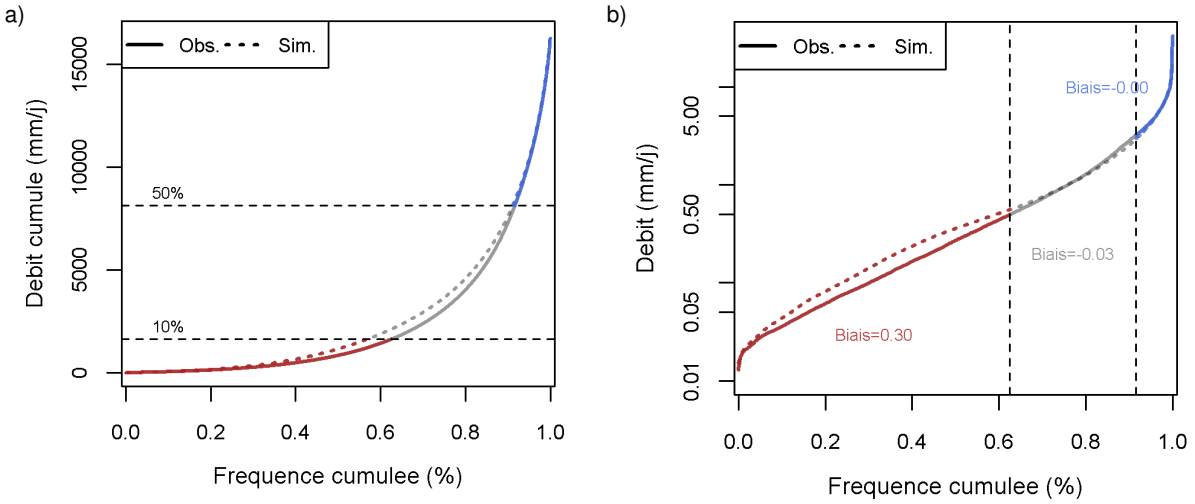


FIGURE 7.1 : Définition des gammes de basses eaux et de hautes eaux, exemple pour la Meuse à Goncourt avec le modèle GR4J. **a)** Courbe des débits classés cumulés, pour les débits observés (trait plein) et les débits simulés (trait pointillé). La gamme des basses eaux (rouge) est définie par les débits les plus faibles dont la somme équivaut à 10% des débits cumulés observés totaux. La gamme des hautes eaux (bleu) est définie par les débits les plus forts dont la somme équivaut à 50% des débits cumulés observés totaux. La gamme des débits intermédiaires (gris) correspond aux débits situés entre les basses eaux et les hautes eaux. **b)** Courbe des débits classés, pour les débits observés (trait plein) et les débits simulés (trait pointillé). Les gammes des basses eaux, des eaux intermédiaires, et des hautes eaux sont illustrées respectivement en rouge, gris et bleu. Le biais du modèle pour chacune des gammes de débit est reporté aux couleurs correspondantes.

Pour le calcul du PMR sur les trois gammes de débit, les gammes ont été redéfinies à partir de la distribution des débits observés sur chacune des sous-périodes de cinq ans du calcul, afin de garantir la cohérence des gammes sur chaque sous-période (Figure 7.2.a). L'exemple de la Meuse à Goncourt montre que l'évolution de l'erreur du modèle au cours du temps ne présente pas les mêmes motifs d'une gamme à une autre, et corrobore la pertinence de la distinction des écoulements en gammes de débit pour caractériser plus finement le comportement du modèle.

7.1.2.b. Principes de l'évaluation de la polyvalence du modèle

La méthode développée dans ce chapitre repose sur le tracé des fronts de Pareto des performances du modèle dans l'espace des critères de performance. De ce fait, il est nécessaire d'utiliser un algorithme spécialisé pour l'optimisation multi-critères, ou bien de tirer un très grand nombre de jeux de paramètres. Pour rendre l'exercice indépendant de la méthode d'optimisation appliquée,

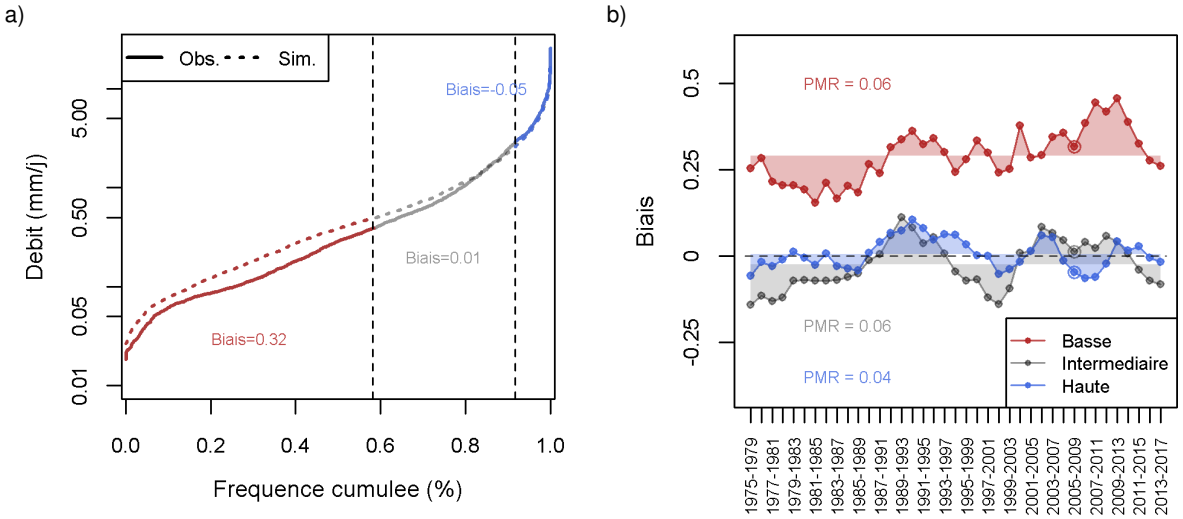


FIGURE 7.2 : a) Courbe des débits classés de la Meuse à Goncourt sur la période 10/2004-09/2009, pour les débits observés (trait plein) et les débits simulés par GR4J (trait pointillé). Les gammes des basses eaux, des eaux intermédiaires, et des hautes eaux sont illustrées respectivement en rouge, gris et bleu. Le biais du modèle sur chacune des gammes de débit est reporté aux couleurs correspondantes. **b)** Courbe des biais glissants du modèle pour les trois gammes de débit sur les sous-périodes glissantes de 5 ans. Les zones colorées correspondent aux aires sous les courbes des biais glissants servant au calcul du PMR pour les trois gammes de débit. Les valeurs de PMR du modèle pour chacune des gammes de débit est reporté aux couleurs correspondantes.

nous avons choisi de mettre en œuvre la deuxième option en explorant l'espace sur une grille de jeux de paramètres. Pour des modèles possédant un nombre de paramètres plus élevé que GR4J, ce procédé serait moins envisageable dans la mesure où les coûts de calcul augmentent exponentiellement avec le nombre de paramètres libres.

La première étape de la méthode consiste dans un premier temps à donner une estimation générale de la sévérité des compétitions entre les critères de performance. L'enjeu sous-jacent à ces compétitions est l'impossibilité pour le modèle d'atteindre avec un unique jeu de paramètres les meilleures performances qu'il est capable d'atteindre vis-à-vis de chacun des critères pris séparément. Dans de telles conditions, il est nécessaire d'employer des jeux de paramètres différents pour optimiser l'un ou l'autre des critères considérés, ce qui n'est a priori pas pertinent pour conserver la cohérence des processus modélisés.

En considérant que l'espace des paramètres a été exploré avec N jeux de paramètres candidats, on peut définir la polyvalence d'un jeu de paramètres comme la moyenne des rangs de ce jeu parmi les autres candidats pour chacun des critères de performance considérés.
Supposons qu'il existe un jeu de paramètres qui serait le meilleur quel que soit le critère évalué. Dans l'hypothèse où les N jeux de paramètres ont été suffisamment bien choisis ou que N est suffisamment grand pour qu'on puisse considérer que les régions de l'espace des paramètres où l'optimum « hydrologique » (cf. Chapitre 2) se trouve ont été explorées, alors identifier un tel jeu de paramètres indiquerait qu'il n'existe aucune compétition entre les différents critères de performance. Dans ce cas, le modèle est parfaitement polyvalent.
Dans le cas général néanmoins, un tel jeu de paramètres n'existe pas. On définit la polyvalence du modèle comme le meilleur rang moyen atteint parmi les N jeux de paramètres candidats. Pour plus de simplicité et de clarté, les rangs sont comptabilisés de 0 % à 100%, 0 % étant la pire valeur et 100 % la meilleure. Ainsi, plus la polyvalence d'un modèle sur un bassin versant est proche de 100%, plus il est facile d'optimiser simultanément les critères de performance considérés. Le jeu de

paramètres ayant le meilleur rang moyen est appelé jeu de paramètres « le plus polyvalent ».

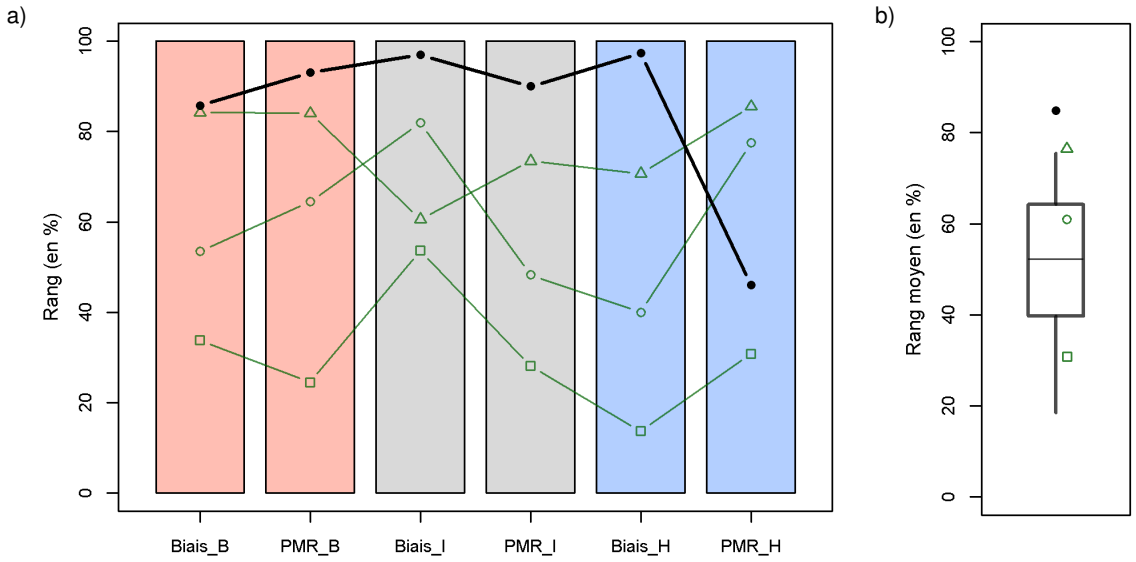


FIGURE 7.3 : a) Rangs pour les différents critères de performances de plusieurs jeux de paramètres échantillonés, dont le jeu de paramètre le plus polyvalent au regard des six critères (en noir) et des jeux de paramètres tirés au hasard (en vert), sur le bassin de la Meuse à Goncourt. Théoriquement, un jeu de paramètres parfait aurait un rang proche de 100% dans chaque critère. **b)** Distribution des rangs moyens des jeux de paramètres échantillonnés. Les jeux de paramètres précédents sont montrés respectivement par les points verts (carré, rond, triangle), le jeu de paramètres le plus polyvalent, qui maximise le rang moyen, est représenté par le point noir plein.

La Figure 7.3 illustre l'application de ces concepts au cas de GR4J sur le bassins de la Meuse à Goncourt. Chaque ligne de la Figure 7.3.a représente les rangs d'un jeu de paramètres donné pour chacun des six critères. Le jeu de paramètres le plus polyvalent, i.e. dont la moyenne des rangs est maximale, est matérialisé par la ligne noire. Un jeu de paramètres parfait serait associé à une ligne droite collée au bord supérieur du graphe.
Sur cet exemple, la moyenne des rangs des jeux de paramètres testés s'échelonne entre 9,9% et 84,9% (Figure 7.3.b). La polyvalence du modèle est donc 84,9% et est atteinte par un jeu de paramètres dont les performances sont quasiment les meilleures pour tous les critères considérés excepté pour le PMR du modèle sur la gamme des hautes eaux. D'autres jeux de paramètres atteignent des performances plus satisfaisantes pour ce critère, mais au détriment de performances vis-à-vis des cinq autres critères.
On notera que le jeu de paramètres le plus polyvalent est nécessairement optimal au sens de Pareto, puisqu'il n'existe par définition pas de jeu de paramètres plus performant pour tous les critères considérés à la fois. De ce fait, le jeu de paramètres le plus polyvalent est situé sur le front de Pareto dans l'espace des critères de performance.

Pour que le classement des jeux de paramètres ait un sens, il est nécessaire d'arrondir les valeurs des performances de sorte à considérer comme équiperformants des jeux de paramètres dont les écarts de performance sont faibles. Pour cette raison, il est plus avantageux d'impliquer des critères d'erreur de même nature dans l'analyse pour arrondir tous les critères de manière identique. Dans notre cas, les six critères correspondent à une erreur volumique relative par rapport au débit observé, et ont été arrondis au millième près, afin de filtrer les variations de l'erreur du modèle en deçà de 0,1% du débit moyen des gammes considérées.

L'intérêt majeur du raisonnement par classement des jeux de paramètres est de donner une estimation agrégée de la difficulté pour le modèle d'optimiser plusieurs critères simultanément, sans tenir compte des hétérogénéités entre eux. Cette méthode permet par exemple de mettre sur un même plan un critère variant dans de très larges gammes de valeurs et un critère variant dans une gamme plus restreinte. Néanmoins, pour que ce procédé soit pertinent, il faut que le cadre d'analyse satisfasse deux conditions :

- Une exploration suffisamment bien menée pour que les zones proches du front de Pareto des performances soient couvertes ;
- Arrondir judicieusement les performances du modèle calculées avec chacun des critères pour ne pas donner de poids aux variations très faibles des performances entre les jeux de paramètres.

7.1.2.c. Principes de l'évaluation de la compatibilité entre les performances

La polyvalence du modèle donne une estimation générale de la sévérité des compétitions entre les performances. Pour affiner l'analyse et identifier les critères les plus difficiles à concilier pour un modèle, la polyvalence du modèle peut également être calculée pour un sous-ensemble des critères.
La Figure 7.4 illustre ce procédé avec les critères de chaque gamme de débit pris deux à deux. Les graphes montrent bien que le jeu de paramètres le plus polyvalent pour chaque couple de critères est bien situé sur le front de Pareto qui leur est associé. La polyvalence du modèle calculée avec chacun de ces couples atteint respectivement 83,2%, 83,8% et 85,6% pour le couple de critères de basses eaux, celui des critères des eaux intermédiaires, et celui des critères des hautes eaux.

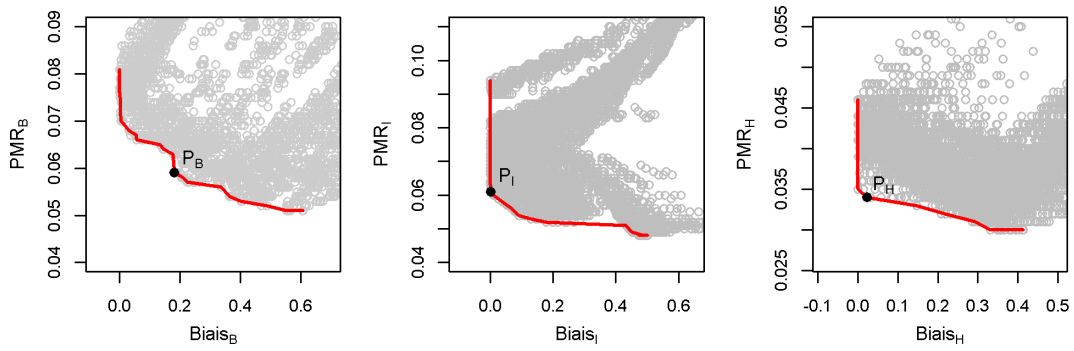


FIGURE 7.4 : Fronts de Pareto des couplets de performances du modèle sur le bassin de la Meuse à Goncourt pour l'ensemble des jeux de paramètres échantillonnés, pour les couples de critères correspondant à chaque gamme de débit (trois graphes pour trois gammes de débit). Les points P_B, P_I et P_H sont les jeux de paramètres les plus polyvalents respectifs pour chacun des couples de critères de performances considérés.

Pour clarifier la notion de polyvalence du modèle, nous avons choisi de nommer « polyvalence du modèle » uniquement le résultat de la prise en compte des six critères de performance. Lorsque son calcul est décliné sur un sous-ensemble des six critères, le résultat du calcul a été désigné comme « compatibilité entre les performances » du sous-ensemble. De ce fait, les chiffres précédents (83,2%, 83,8% et 85,6%) correspondent aux compatibilités entre les couples des critères de chaque gamme de débit.
La compatibilité entre plusieurs critères de performances n'a été calculée qu'entre des couples ou des triplets de performance. Lorsque cela s'avérait pertinent, nous avons comparé la polyvalence du modèle et la compatibilité entre certains critères de performance.

7.2. Évaluation de la polyvalence du modèle GR4J

Les résultats présentés dans cette section concernent dans un premier temps l'évaluation de la polyvalence du modèle GR4J sur les deux sous-échantillons de bassins français et australien, en la mettant en perspective avec les performances obtenues en optimisant tous les critères simultanément ou par groupes.

7.2.1. Échantillonnage des paramètres du modèle

La grille d'exploration des paramètres a été déterminée à partir des quantiles de la distribution des jeux de paramètres résultants du calage de GR4J avec le KGE racine sur chaque bassin versant de l'échantillon d'étude. Pour les paramètres X_1, X_2 et X_3, nous avons sélectionné parmi les valeurs ainsi obtenues pour chacun de ces paramètres les quantiles 0,05, 0,15, 0,25 jusqu'à 0,95, ainsi que la médiane (quantile 0,50). Les 11 valeurs de chacun des paramètres sont répertoriées dans le Tableau 7.2. Seules sept valeurs du paramètre X_4 réparties sur une grille régulière ont été testées puisque ce paramètre n'est relié qu'aux hydrogrammes unitaires et a une influence négligeable sur le bilan hydrique du modèle à l'échelle annuelle.

La grille ainsi définie permet de générer un ensemble identique de 9317 jeux de paramètres candidats pour chacun des bassins versants de l'échantillon. Cet ensemble relativement petit reste suffisamment dense en comparaison d'échantillonnages Monte Carlo employés dans d'autres études de modélisation hydrologique (e.g. Liu *et al.*, 2009; Coxon *et al.*, 2019). Les valeurs des paramètres de la grille ayant été sélectionnées parmi les plus fréquemment rencontrées en calage, nous avons supposé que l'échantillon de 9317 jeux de paramètres contenait toujours des jeux de paramètres situés au voisinage des fronts de Pareto.

TABLEAU 7.2 : Échantillonnage des paramètres de GR4J servant de base à l'exploration de l'espace des paramètres du modèle.

X_1 (mm)	X_2 (mm j^{-1})	X_3 (mm)	X_4 (j)
128	–5,92	8,6	0,5
167	–3,48	14,4	1,0
193	–2,56	23,5	1,5
219	–2,35	30,9	2,0
242	–1,62	40,0	2,5
262	–1,33	45,6	3,0
289	–1,02	50,6	3,5
379	–0,70	62,5	
493	–0,47	79,5	
695	–0,20	110,2	
1946	0,46	215,3	

7.2.2. Sensibilité des performances au choix des paramètres

Avant l'analyse de la polyvalence du modèle, nous avons dans un premier temps cherché à caractériser les comportements des critères de biais et de robustesse sur les trois gammes de

débit.

Nous avons d'abord analysé les distributions de leurs écarts-types sur les deux sous-échantillons de bassins versants (Figure 7.5).
Il apparaît clairement que les biais du modèle sont nettement plus sensibles au choix du jeu de paramètres que les PMR, quelle que soit la gamme de débit considérée. Malgré nos précautions pour ne pas définir une gamme des basses eaux représentant des volumes écoulés très faibles en comparaison du reste de l'hydrogramme, les biais sur cette gamme sont de loin les plus sensibles. Cette sensibilité accrue se retrouve d'ailleurs également entre le PMR calculé sur les basses eaux et les PMR des gammes supérieures. Il est intéressant de constater qu'en Australie, les PMR calculés sur les basses eaux sont presque aussi sensibles aux jeux de paramètres que les biais calculés sur la gamme des eaux intermédiaires.

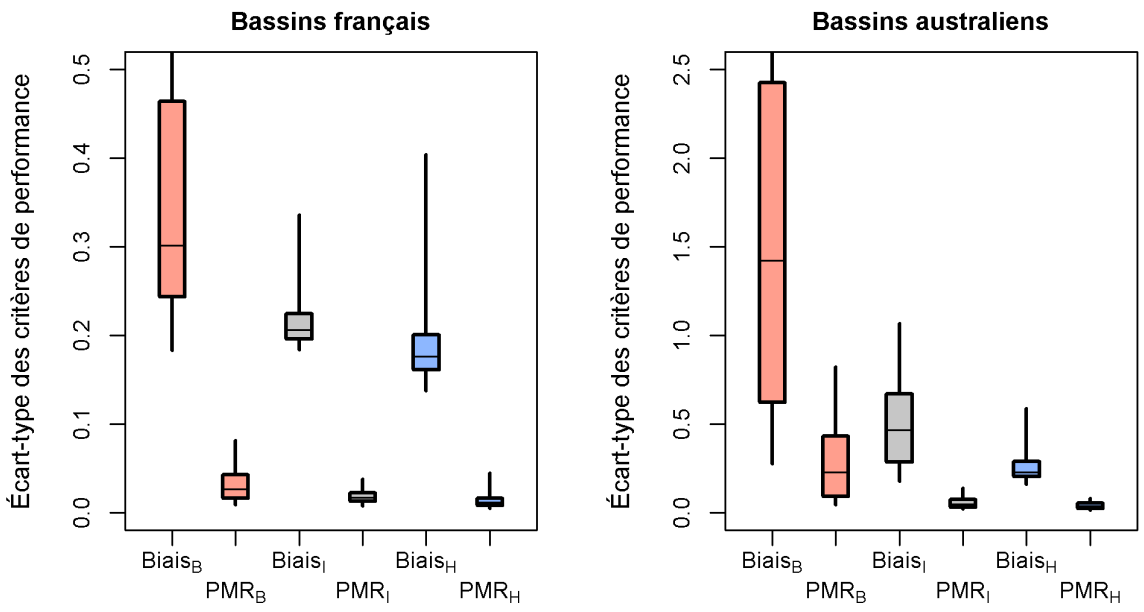


FIGURE 7.5 : Distribution des écarts-types des performances de GR4J pour chaque critère d'évaluation sur les sous-échantillons de bassins français et australien. Les couleurs des boîtes à moustache indiquent la gamme de débit des critères (basses eaux en rouge, eaux intermédiaires en gris, hautes eaux en bleu).

La distribution des meilleures performances atteignables, sans tenir compte de l'effet des compétitions, pour chacun des critères sur les bassins de l'échantillon est montrée en Figure 7.6. Dans la quasi-totalité des bassins de l'étude, il semble exister un jeu de paramètres permettant au modèle d'être parfaitement non biaisé sur la période totale de simulation pour chacune des gammes de débit considérées séparément. Comme cela a été mentionné au chapitre précédent, cette capacité est vraisemblablement permise grâce à la flexibilité du modèle pour ajuster son bilan par l'évaporation et les échanges souterrains. Le fait que cette capacité soit observable à partir de l'évaluation d'une grille fixe de 9317 candidats confirme que la grille couvre correctement l'espace des paramètres.
En revanche, les valeurs de PMR du modèle atteignent un palier infranchissable propre à chaque bassin et à chaque gamme de débit. En général, les meilleurs jeux de paramètres permettent d'atteindre une erreur interannuelle moyenne entre 5% et 15% du débit observé. Ainsi, dans l'hypothèse où le modèle ne souffrirait pas de problèmes de compétitions entre ces critères de robustesse sur les trois gammes de débit, sa robustesse serait malgré tout limitée. Ce constat démontre l'importance de l'étude de la structure du modèle à la fois pour limiter l'effet de ces

compétitions et pour repousser les limites des performances atteignables.

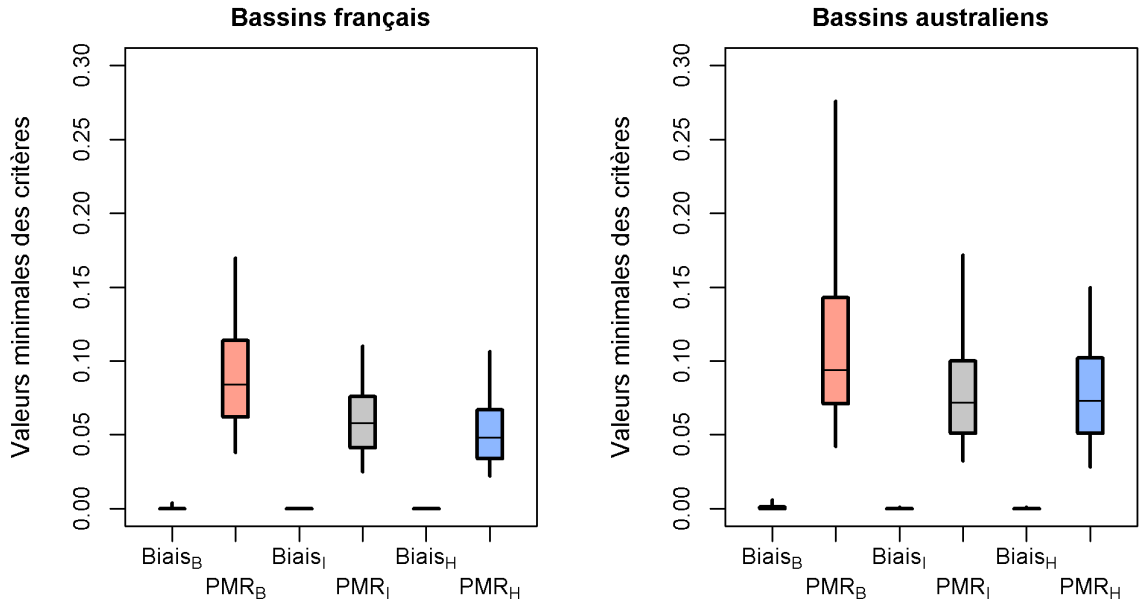


FIGURE 7.6 : Distribution des meilleures performances atteintes par GR4J pour chaque critère d'évaluation, sur les sous-échantillons de bassins français et australien. Les couleurs des boîtes à moustache indiquent la gamme de débit des critères (basses eaux en rouge, eaux intermédiaires en gris, hautes eaux en bleu).

Pour comprendre mieux les résultats de la procédure de sélection du jeu de paramètres le plus polyvalent, nous avons enfin indiqué dans le Tableau 7.3 la moyenne sur les bassins de son rang au regard de chaque critère de performance. Nos résultats montrent qu'en général le jeu de paramètres obtenu est associé à des biais mieux positionnés que les valeurs de PMR qui lui sont associées. Pour rappel, cela ne se traduit pas nécessairement par des valeurs de biais plus faibles que les valeurs de PMR puisque le rang moyen classe les jeux de paramètres relativement entre eux. Les valeurs de PMR sont en outre très nettement moins variables que les biais du modèle sur les bassins de l'échantillon (Figure 7.5).

TABLEAU 7.3 : Moyenne pour chaque critère sur les sous-échantillons de bassins français et australien du rang du jeu de paramètres le plus polyvalent au regard des six critères.

	Rang moyen par critère (en %)					
	$Biais_B$	PMR_B	$Biais_I$	PMR_I	$Biais_H$	PMR_H
Bassins français	98,9	84,7	92,5	86,3	93,6	85,9
Bassins australiens	92,0	91,3	90,8	81,5	95,4	87,5

Il est possible que le meilleur classement du jeu de paramètres le plus polyvalent en termes de biais soit causé par une compétition moins sévère entre les biais du modèle sur les trois gammes de débit qu'entre les PMR sur ces trois gammes. Pour vérifier cette hypothèse, nous avons dans la suite étudié les performances du modèle avec les jeux de paramètres les plus polyvalents vis-à-vis des trois critères de biais simultanément ou vis-à-vis des trois PMR simultanément.

7.2.3. Polyvalence du modèle au regard du biais et du PMR

La distribution de la polyvalence du modèle sur les sous-échantillons de bassins français et australien est montrée en Figure 7.7 et comparée à la compatibilité des biais sur les trois gammes de débit entre eux et à la compatibilité des PMR sur les trois gammes entre eux. Les résultats obtenus sur les bassins français et australiens montrent des motifs similaires.
Il apparaît généralement que la compatibilité entre les biais des trois régimes de débit est meilleure que celle entre les PMR. Dans plus de 95% des bassins de chaque sous-échantillon, la compatibilité entre les biais est supérieure à 90%. Cela signifie qu'il existe dans ces bassins un jeu de paramètres qui permet d'être parmi les 10% des meilleurs jeux de paramètres au regard des trois biais simultanément. En comparaison, la compatibilité entre les PMR des trois gammes de débit est supérieure à 90% dans moins de trois quarts des bassins.
Certains bassins français montrent une compatibilité des trois biais très faible ($<$15%). Ces bassins sont remarquables par la présence de substrats poreux : craie, sables ou calcaires. C'est le cas de la Petite Seille à Château-Salins (A7642010), de la Soudé à Soudron (H5213310), de la Brenne à Villedômer (K4873110), de l'Olivet à Beaumont-Village (K7424010) et du Couêtron à Souday (M1214010). Bien que certains de ces bassins ne soient pas caractérisés par un indice de débit de base particulièrement élevé, on peut supposer que les processus modélisés par GR4J ne lui permettent pas de reproduire convenablement l'influence des aquifères sur les écoulements saisonniers, notamment en étiage.

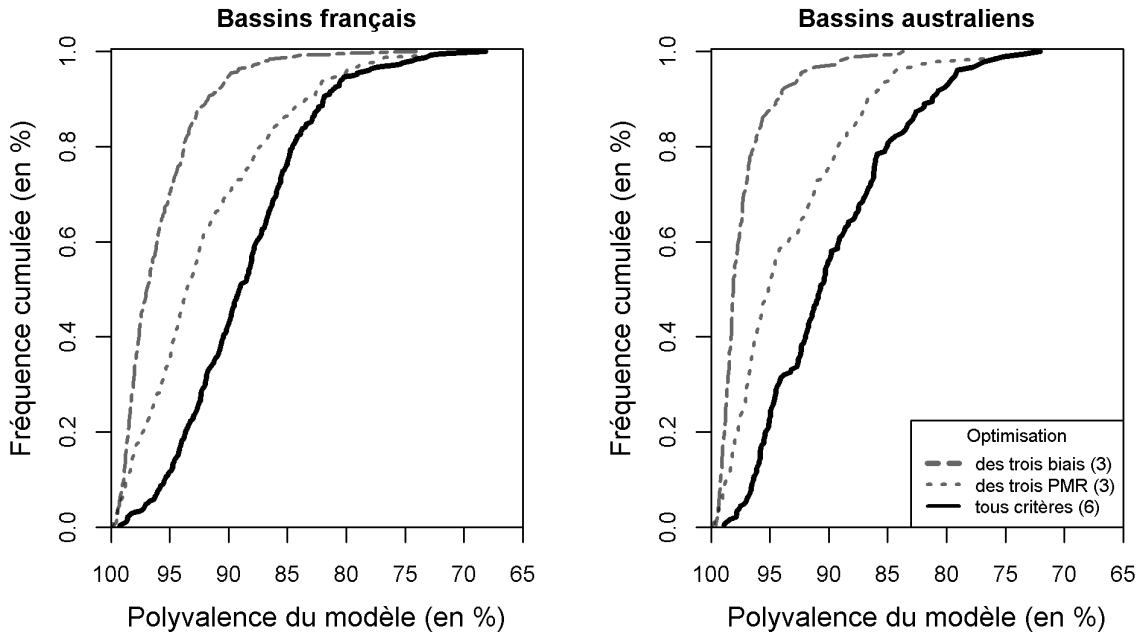


FIGURE 7.7 : Distribution de la polyvalence du modèle au regard des trois critères de biais (pointillés irréguliers), des trois critères de robustesse (pointillés réguliers) et des six critères simultanément (trait plein), sur les sous-échantillons de bassins français et australien.

La polyvalence du modèle varie enfin dans une gamme située entre 70% et 100%, inférieure à celle des compatibilités entre les biais ou entre les PMR. S'il n'existait aucune compétition entre l'optimisation des biais et des PMR, la polyvalence devrait pour chaque bassin être égale à la moyenne entre la compatibilité entre les biais et la compatibilité entre les PMR. Le fait que ça ne soit pas le cas signifie que des compétitions entre les PMR et les biais limitent d'autant plus les performances du modèle.

Les performances de GR4J associées aux jeux de paramètres les plus polyvalents selon les groupes de critères impliqués dans le calcul sont présentées en Figure 7.8 pour les bassins français et en Figure 7.9 pour les bassins australiens. Les deux échantillons montrent des motifs similaires.
La sélection des paramètres en ne considérant que les biais sur les trois gammes de débit débouche sur des biais non nuls sur la plupart des bassins de l'échantillon. De ce fait, malgré la bonne compatibilité des critères de biais constatée précédemment, les compétitions remettent en cause l'existence d'un jeu de paramètres permettant au modèle d'être parfaitement non biaisé sur les trois gammes de débit simultanément. En comparaison des biais constatés avec le jeu de paramètres le plus polyvalent au sens des PMR, ces biais restent cependant relativement restreints. Les biais du modèle avec le jeu de paramètres le plus polyvalent au sens des trois PMR sont en effet conséquents, même dans les gammes de débit supérieures où le débit observé élevé devrait atténuer ces erreurs.
Finalement, le jeu de paramètres le plus polyvalent identifié en impliquant les six critères dans le calcul produit des biais intermédiaires par rapport aux cas précédents. Dans la gamme des hautes eaux, ces biais sont assez proches des biais résultant de l'optimisation des trois critères de biais. Dans la gamme des eaux intermédiaires, ces biais sont plus importants pour les bassins australiens que pour les bassins français.

En ce qui concerne les performances de GR4J en termes de PMR, les distributions sont relativement stables selon que le jeu de paramètres le plus polyvalent est sélectionné avec les trois critères de biais seulement, les trois critères de PMR seulement, ou les six critères simultanément. L'amélioration la plus notable apportée par l'optimisation des trois PMR seuls concerne le PMR du modèle sur les eaux intermédiaires australiennes.
Dans tous les cas, les valeurs obtenues se démarquent des meilleures performances théoriquement atteignables si le modèle ne souffrait d'aucune compétition pour l'optimisation de ses performances. Ainsi, la prise en considération des trois PMR pour optimiser les performances n'améliore que très légèrement la robustesse du modèle dans les trois gammes du débit. Ce résultat est cohérent avec le constat fait précédemment sur la compatibilité des trois PMR entre eux, qui était généralement plus faible que la compatibilité des biais sur les trois gammes de biais entre eux (Figure 7.7). La robustesse du modèle sur les trois gammes de débit semble donc assez fortement limitée par l'existence de compétitions entre la capacité du modèle à simuler à la fois les variations interannuelles des basses eaux, des hautes eaux et des eaux intermédiaires.
Finalement, au contraire des critères de biais, il ne paraît pas impératif d'impliquer les critères de PMR dans le processus de sélection des jeux de paramètres tant les améliorations sont faibles entre la situation où seuls les biais sont considérés et la situation où les six critères le sont simultanément.

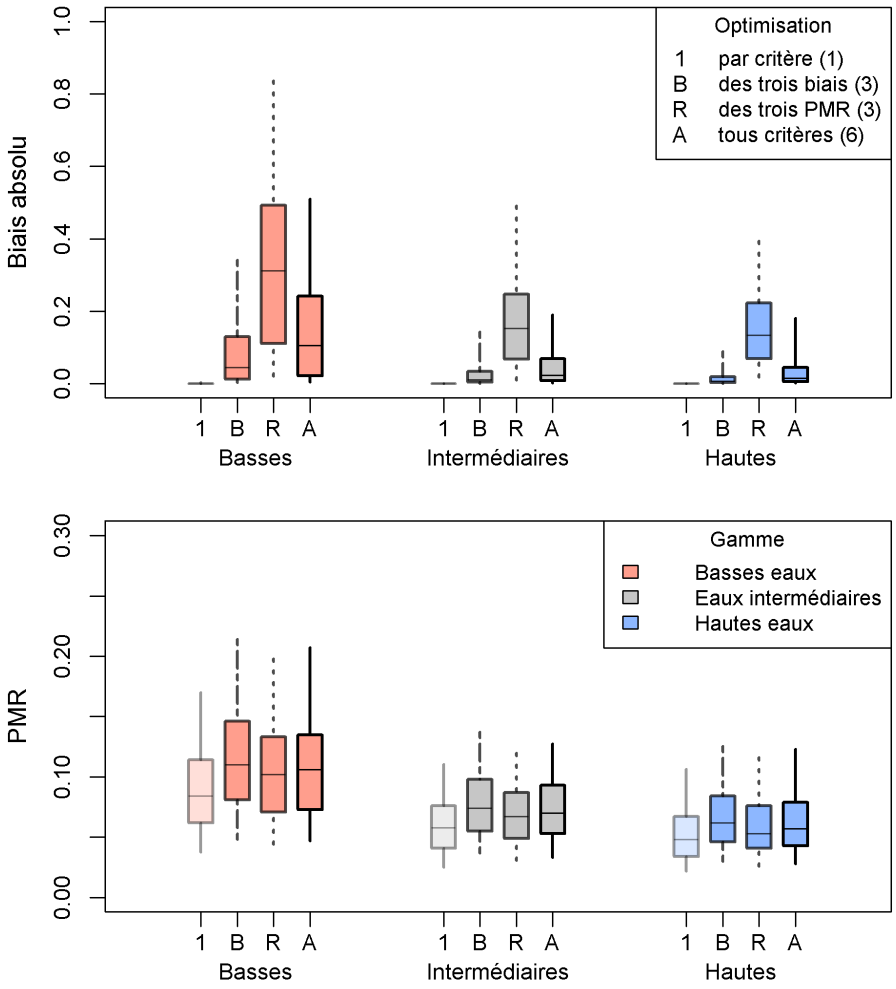


FIGURE 7.8 : Distribution des performances atteintes par le modèle sur le sous-échantillon de bassins français avec les jeux des paramètres les plus polyvalents au regard de chaque critère indépendamment (moustaches transparentes), des trois critères de biais (moustaches pointillées irrégulières), des trois critères de robustesse (moustaches pointillées régulières), ou des six critères simultanément (moustaches pleines). Les couleurs des boîtes à moustache indiquent la gamme de débit des critères (basses eaux en rouge, eaux intermédiaires en gris, hautes eaux en bleu).

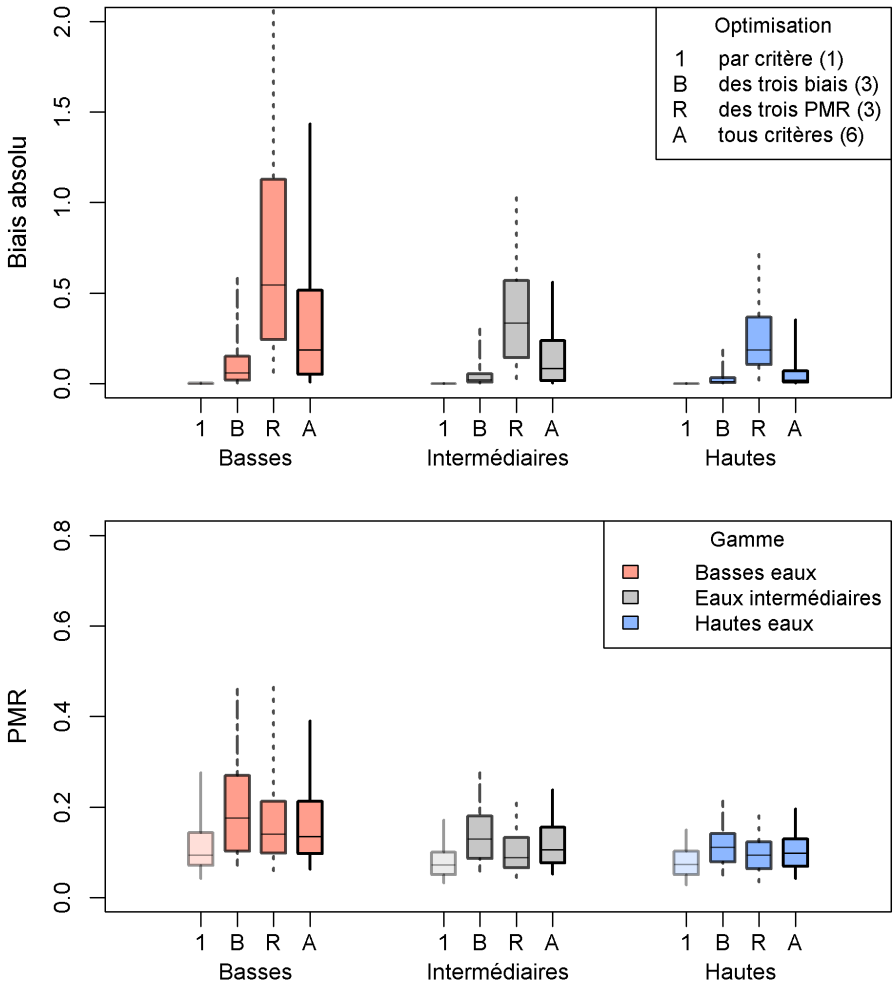


FIGURE 7.9 : Distribution des performances atteintes par le modèle sur le sous-échantillon de bassins australien avec les jeux des paramètres les plus polyvalents au regard de chaque critère indépendamment (moustaches transparentes), des trois critères de biais (moustaches pointillées irrégulières), des trois critères de robustesse (moustaches pointillées régulières), ou des six critères simultanément (moustaches pleines). Les couleurs des boîtes à moustache indiquent la gamme de débit des critères (basses eaux en rouge, eaux intermédiaires en gris, hautes eaux en bleu).

7.2.4. Polyvalence du modèle au regard des gammes de débit

Pour analyser l'influence des compétitions entre les critères de biais et les PMR, nous avons comparé les compatibilités entre le biais et le PMR de chaque gamme de débit à la polyvalence du modèle (Figure 7.10). À l'exception des eaux intermédiaires des bassins australiens, les compatibilités des deux critères de chaque gamme de débit montrent des distributions très similaires. De manière générale, le modèle éprouve peu de difficulté à reproduire à la fois le débit moyen d'une gamme de débit et ses variations interannuelles. Néanmoins, la compatibilité entre les couples de critères s'avère moins bonne dans environ un dixième de chaque échantillon. En comparant les bassins du dernier décile pour chacun des couples de critères, nous avons constaté qu'aucun bassin n'était concerné par une forte incompatibilité entre les couples de critères des trois gammes de débit à la fois. Seule une très faible partie d'entre eux est marquée par une incompatibilité des biais et des PMR de deux gammes de débit à la fois. De ce fait, puisque les bassins en queue des distributions ne sont pas les mêmes pour chaque distribution, le nombre de bassins concernés par au moins une incompatibilité entre biais et PMR est finalement assez important.
Au contraire des compatibilités des couples de critères des gammes des basses et hautes eaux, la compatibilité entre le biais et le PMR du modèle sur la gamme des eaux intermédiaires atteint des valeurs assez faibles sur une grande partie du sous-échantillon australien.

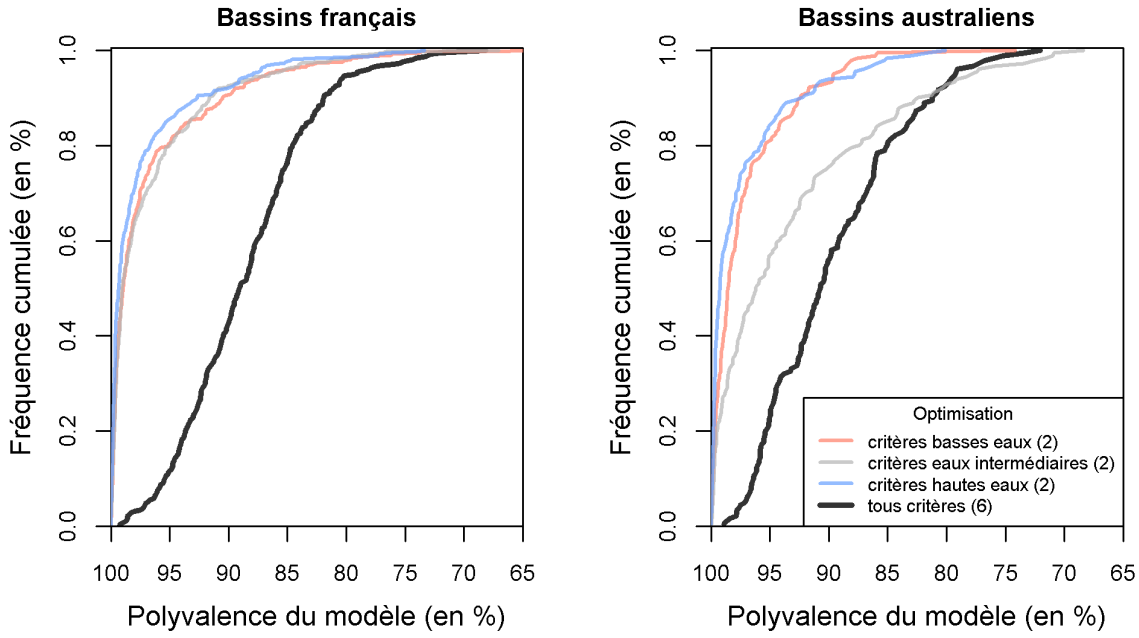


FIGURE 7.10 : Distribution de la polyvalence du modèle au regard des deux critères associés aux basses eaux (rouge), des deux critères associés aux eaux intermédiaires (gris), des deux critères associés aux hautes eaux (bleu), et des six critères simultanément (noir), sur les sous-échantillons de bassins français et australiens.

Pour étayer ce constat, les performances associées aux jeux de paramètres les plus polyvalents au regard des couples de critères de chaque gamme de débit sont montrées sur la Figure 7.11 pour les bassins français et sur la Figure 7.12 pour les bassins australiens.
Dans l'ensemble, les biais obtenus dans chaque gamme avec les jeux de paramètres les plus polyvalents associés à chacune de ces gammes sont faibles. Néanmoins, en comparaison des biais constatés précédemment avec les jeux de paramètres optimisant les trois biais pris simultanément (Figure 7.9), ceux-ci ne sont réellement plus proches de l'optimum que dans le cas des hautes eaux. Pour les eaux intermédiaires, ce résultat est cohérent avec le constat précédent de faible

compatibilité du biais et du PMR sur la gamme des eaux intermédiaires. Il est intéressant de noter que les limitations de performance du modèle liées à l'incompatibilité entre deux critères peuvent de fait dépasser les limites liées à l'incompatibilité entre trois critères.
On constate dans les deux sous-échantillons de bassins que les biais sur les eaux intermédiaires sont plus faibles avec les jeux de paramètres les plus polyvalents au regard des critères de basses eaux qu'avec les jeux de paramètres les plus polyvalents au regard des critères de hautes eaux. Bien que les gammes définies soient difficilement interprétables en termes physiques, il est possible que les processus impliqués dans le comportement des eaux intermédiaires soient proches de ceux impliqués dans celui des basses eaux.

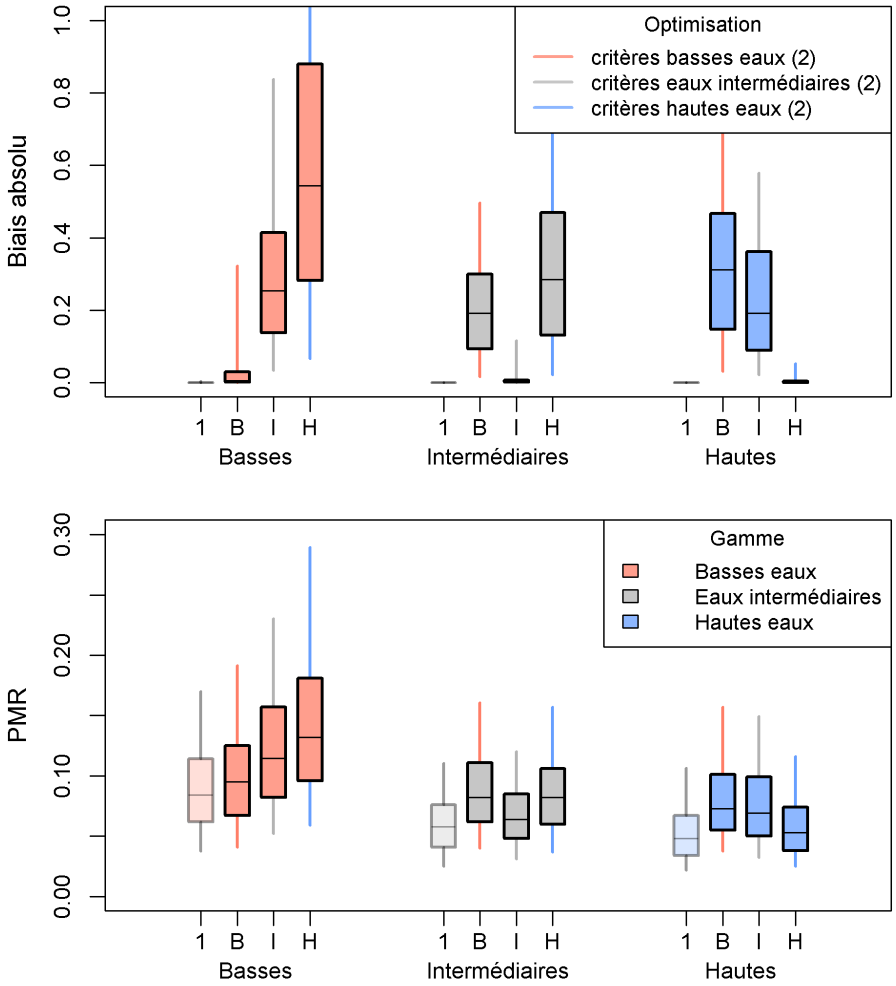


FIGURE 7.11 : Distribution des performances atteintes par le modèle sur le sous-échantillon de bassins français avec les jeux des paramètres les plus polyvalents au regard de chaque critère indépendamment (moustaches transparentes), des deux critères associés aux basses eaux (moustaches rouges), des deux critères associés aux eaux intermédiaires (moustaches grises), ou des deux critères associés aux hautes eaux (moustaches bleues). Les couleurs des boîtes à moustache indiquent la gamme de débit des critères (basses eaux en rouge, eaux intermédiaires en gris, hautes eaux en bleu).

Les distributions des valeurs de PMR de chaque gamme obtenues en optimisant les critères de la gamme associée sont assez proches des meilleures performances atteignables du modèle. Sans surprise, plus la gamme de débits associée aux jeux de paramètres considérés est éloignée de la gamme de débit pour laquelle le PMR est calculé, moins les performances du modèle sont satisfaisantes.

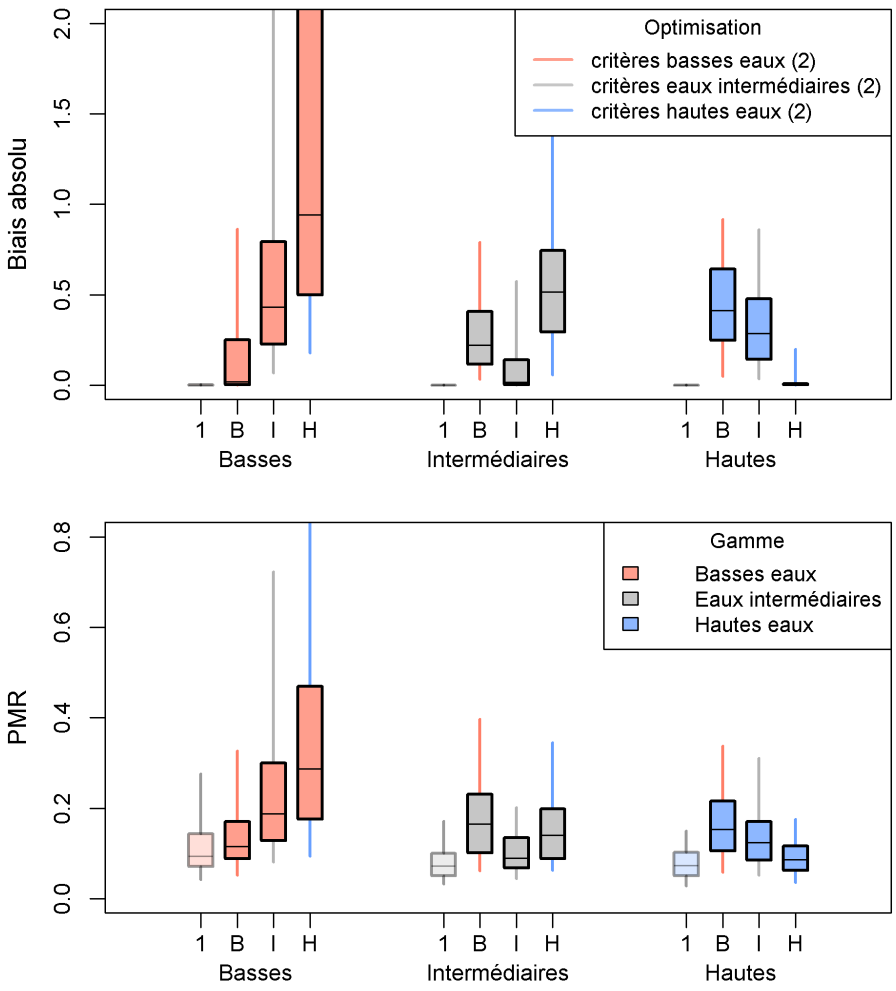


FIGURE 7.12 : Distribution des performances atteintes par le modèle sur le sous-échantillon de bassins australien avec les jeux des paramètres les plus polyvalents au regard de chaque critère indépendamment (moustaches transparentes), des deux critères associés aux basses eaux (moustaches rouges), des deux critères associés aux eaux intermédiaires (moustaches grises), ou des deux critères associés aux hautes eaux (moustaches bleues). Les couleurs des boîtes à moustache indiquent la gamme de débit des critères (basses eaux en rouge, eaux intermédiaires en gris, hautes eaux en bleu).

En conclusion, nos résultats confirment que des compétitions entre les critères limitent bel et bien l'optimisation des performances du modèle. Alors que le modèle possède pour chaque gamme d'écoulement un jeu de paramètres qui pourrait lui permettre de produire des simulations non biaisées, le jeu de paramètres optimisant les biais dans les trois gammes à la fois produit des débits simulés biaisés sur chacune de ces gammes. En outre, bien que les valeurs de PMR du modèle soient peu sensibles au jeu de paramètres choisi et qu'elles atteignent un palier pour chaque bassin versant, des compétitions sévères entre les PMR des trois gammes de débit semblent également réduire la robustesse du modèle.
Les compétitions n'existent pas seulement entre les performances du modèle pour chaque gamme de débit, mais également entre la capacité du modèle à reproduire simultanément le débit moyen et ses variations à l'intérieur d'une même gamme. Toutefois, la baisse des performances associée à ces compétitions intra-gammes semble plus faible que celle liée aux compétitions inter-gammes.

7.3. Analyse du comportement du modèle GR4J

Dans la suite, nous avons cherché à identifier les critères de performance les moins compatibles, puis à comprendre les raisons de ces incompatibilités à travers les paramètres et les états internes de GR4J.

7.3.1. Compatibilité entre les critères de performances

7.3.1.a. Compatibilité des critères en couplets

Les compatibilités entre les couples de critères sont répertoriées dans les Tableaux 7.4 et 7.5 pour les sous-échantillons de bassins versants français et australien respectivement. Les résultats obtenus dans les deux régions étudiées sont dans l'ensemble assez proches.
Dans les deux régions étudiées, les biais sont généralement très compatibles deux à deux (environ à 99%). Cela indique que le modèle n'éprouve pas de difficulté particulière à reproduire simultanément les débits moyens de deux catégories de débit différentes, même lorsque ces gammes caractérisent des écoulements opposés dans l'hydrogramme (basses et hautes eaux).
Les critères de PMR sont légèrement moins compatibles avec les critères de biais, quelles que soit les gammes de débit concernées. Toutefois, à l'exception du cas déjà observé du biais et de la robustesse du modèle sur les eaux intermédiaires australiennes, les compatibilités entre un critère de biais et un critère de PMR sont meilleures que celles entre couples de PMR. De ce fait, il semble que GR4J ait suffisamment de flexibilité pour reproduire correctement à la fois le débit moyen et ses variations interannuelles, même sur des gammes de débit différentes, mais que la simulation des variations interannuelles de deux gammes de débit soit un exercice assez difficile.

TABLEAU 7.4 : Moyennes des compatibilités sur le sous-échantillon de bassins français des critères d'évaluation en couplets (en %).

	$Biais_B$	PMR_B	$Biais_I$	PMR_I	$Biais_H$	PMR_H
$Biais_B$		96,9	98,9	96,7	99,0	98,0
PMR_B			97,7	95,9	97,9	96,0
$Biais_I$				96,9	99,1	98,1
PMR_I					98,6	95,9
$Biais_H$						97,5
PMR_H						

TABLEAU 7.5 : Moyennes des compatibilités sur le sous-échantillon de bassins australiens des critères d'évaluation en couplets (en %).

	$Biais_B$	PMR_B	$Biais_I$	PMR_I	$Biais_H$	PMR_H
$Biais_B$		97,2	99,0	95,7	99,3	97,7
PMR_B			98,9	96,4	98,4	96,3
$Biais_I$				93,3	99,2	97,5
PMR_I					98,2	96,7
$Biais_H$						97,5
PMR_H						

Pour déterminer les critères de performance les plus complexes à optimiser conjointement avec d'autres critères, nous avons calculé pour chaque critère la moyenne des compatibilités en couple avec les autres critères. Les résultats sont représentés en Figures 7.13 et 7.14 pour les sous-échantillons français et australiens respectivement.
Dans les bassins français comme australiens, la simulation de l'écoulement moyen en hautes eaux semble être l'aspect du régime hydrologique le plus simple à reproduire couplé avec un autre objectif de modélisation. À l'inverse, les performances au regard des critères de robustesse sont plus difficiles à concilier à d'autres objectifs de performances. Ce résultat est logiquement lié au constat précédent que les critères de PMR sont assez incompatibles entre eux et modérément compatibles avec les critères de biais.
Cependant, il est intéressant d'observer que dans les bassins australiens les compatibilités moyennes du biais et du PMR sur les basses eaux en couple avec les autres critères sont du même ordre de grandeur. Une interprétation possible de ce résultat pourrait être que pour représenter le débit moyen en étiage sur les nombreux bassins touchés par la *Millenium Drought*, le modèle doit être en mesure de simuler la baisse des débits d'étiages liée à cette sécheresse. De ce fait, les processus impliqués dans l'optimisation des performances en biais et en PMR sur les basses eaux étant les mêmes, il est tout aussi difficile pour le modèle de concilier cet objectif avec un l'optimisation d'un autre critère.
Enfin, dans les bassins australiens, la capacité du modèle à reproduire les variations interannuelles des eaux intermédiaires est clairement la plus limitée par les compétitions avec les autres critères de performance.

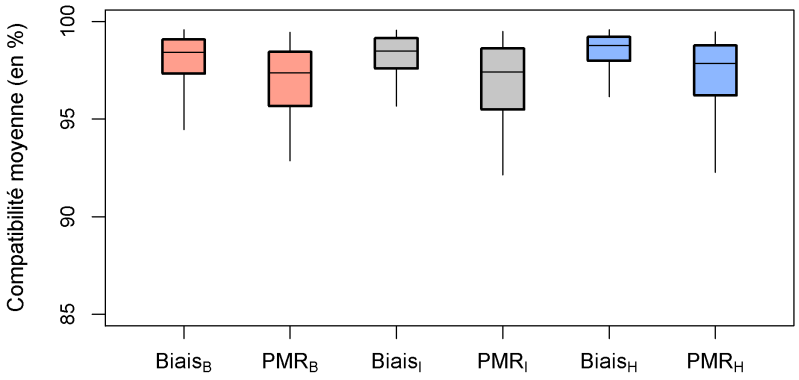


FIGURE 7.13 : Distribution des compatibilités en couplets moyennées de chaque critère de performance pour le modèle sur le sous-échantillon français. Les couleurs des boîtes à moustache indiquent la gamme de débit des critères (basses eaux en rouge, eaux intermédiaires en gris, hautes eaux en bleu).

7.3.1.b. Compatibilité des critères en triplets

L'évaluation des compatibilités entre les critères pris trois à trois a été réalisée afin d'étayer le constat fait précédemment sur les combinaisons des performances du modèle sur les trois gammes de débits simultanément. Nous ne présentons ces résultats que brièvement dans la mesure où l'interprétation de certaines combinaisons est assez compliquée à formuler (comme par exemple la combinaison d'un critère de biais sur une gamme de débit et des PMR sur les deux autres gammes).
Sur les deux sous-échantillons, il apparaît nettement que la combinaison des biais des trois gammes des débits est la plus facile à optimiser pour GR4J. La compatibilité entre ces trois critères est en moyenne la meilleure des toutes les compatibilités entre triplets de performance, avec respectivement 95,9% et 97,2% en France et en Australie (non représenté dans un tableau,

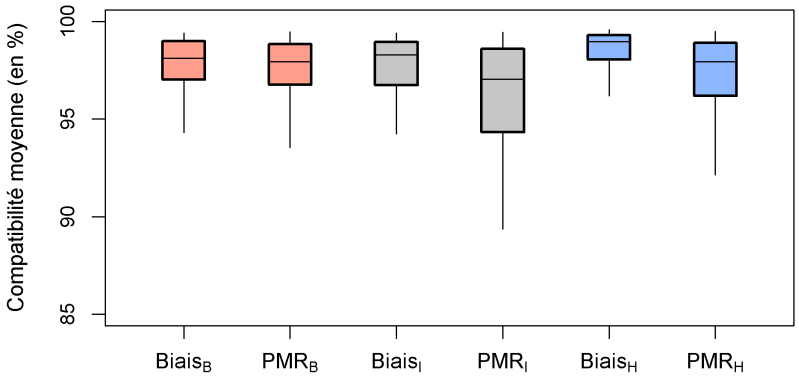


FIGURE 7.14 : Distribution des compatibilités en couplets moyennées de chaque critère de performance pour le modèle sur le sous-échantillon australiens. Les couleurs des boîtes à moustache indiquent la gamme de débit des critères (basses eaux en rouge, eaux intermédiaires en gris, hautes eaux en bleu).

pour des raisons de concision). À l'opposé, la compatibilité entre les trois critères de PMR est la pire en moyenne sur les bassins français (92,1%). Sur les bassins australiens, la situation est moins contrastée, mais la compatibilité entre ces critères fait également partie des moins bonnes (93,4%). Sur ce sous-échantillon, les pires scores sont atteints de manière générale pour les triplets incluant le PMR sur les eaux intermédiaires (pire compatibilité de 91,0% entre biais sur les eaux intermédiaires et PMR sur les eaux intermédiaires et hautes).

Cette situation est visible sur les compatibilités moyennes de chacun des critères en triplets avec les autres critères (Figures 7.15 et 7.16). En effet, ici encore les performances qui s'accomodent le moins aisément aux autres objectifs de modélisation pour les bassins australiens sont les variations interannuelles du débit sur les eaux intermédiaires.

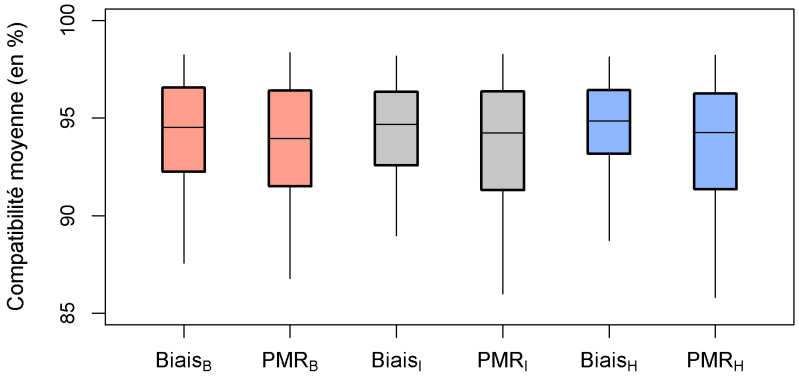


FIGURE 7.15 : Distribution des compatibilités en triplets moyennées de chaque critère des performance pour le modèle sur le sous-échantillon français. Les couleurs des boîtes à moustache indiquent la gamme de débit des critères (basses eaux en rouge, eaux intermédiaires en gris, hautes eaux en bleu).

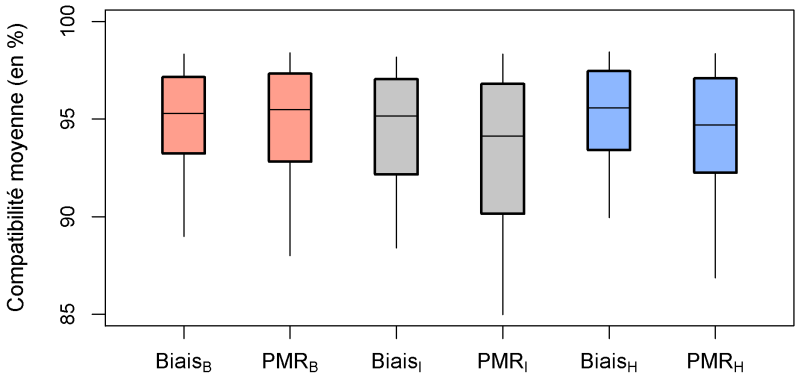


FIGURE 7.16 : Distribution des compatibilités en triplets moyennées de chaque critère des performance pour le modèle sur le sous-échantillon australiens. Les couleurs des boîtes à moustache indiquent la gamme de débit des critères (basses eaux en rouge, eaux intermédiaires en gris, hautes eaux en bleu).

7.3.2. Comparaison des paramètres associés aux gammes de débit

Il convient à ce stade de l'analyse de chercher à identifier les éléments de la structure de GR4J qui sont à l'origine des compétitions constatées entre les critères de performance. En première approche, l'analyse des paramètres peut renseigner sur les configurations préférentiellement adoptées par le modèle pour maximiser ses performances dans un domaine ou dans un autre.
Considérant que les compétitions entre performances culminent lors de la prise en compte simultanée des trois PMR, nous avons considéré qu'il était pertinent de comparer les distributions des paramètres liés à chacune des gammes de débit. Dans la mesure où dans la majorité des situations observées, le biais et le PMR du modèle d'une même gamme se sont révélés être relativement compatibles l'un avec l'autre, nous avons choisi pour chaque gamme les jeux de paramètres les plus polyvalents au regard du biais et du PMR de cette gamme.

Les distributions des paramètres résultant de l'optimisation des performances pour chaque gamme de débit sont présentées en Figures 7.17 et 7.18.
Sur les bassins du sous-échantillon français, l'optimisation des critères de basses eaux conduit le réservoir de production du modèle à adopter une taille X_1 très nettement plus grande que dans les autres cas, traduisant vraisemblablement le besoin d'une mémoire hydrologique plus importante pour simuler les basses eaux. Ce phénomène est visible dans une moindre mesure sur le sous-échantillon australien. À apport pluviométrique et demande évaporatoire donnés, les petits réservoirs de production sont plus réactifs : ils se remplissent et se vident plus rapidement que les grands réservoirs. La petite taille des réservoirs de production associée aux jeux de paramètres résultant de l'optimisation des critères des eaux intermédiaires et des hautes eaux suggère que le modèle bénéficie de cette réactivité accrue pour la simulation de ces écoulements. Dans le cadre de la simulation des crues par exemple, il est possible que le modèle ait besoin de retrouver un état simulé d'humidité du sol quasi-saturé dans une période de temps courte pour que la majeure partie des précipitations participe à l'écoulement.
Les distributions du paramètre d'échanges souterrains X_2 se distinguent par un croissance des valeurs du paramètre à mesure que la gamme représentée correspond à des écoulements plus forts. De ce fait, alors que le modèle recourt à un facteur de pertes par le souterrain important pour la simulation des basses eaux, ce facteur de pertes tend à s'atténuer à mesure que l'objectif de simulation se déporte vers les hautes eaux. Dans une part non négligeable des bassins français, les échanges souterrains se produisent d'ailleurs pour la simulation des hautes eaux dans le sens

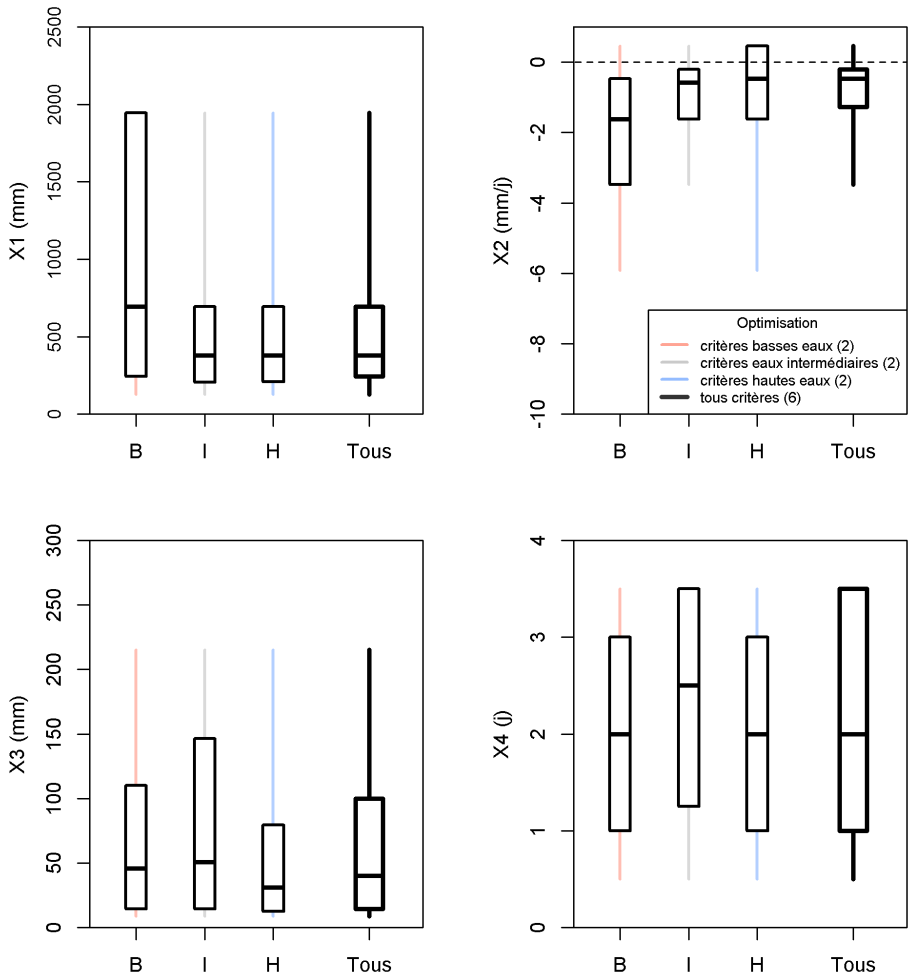


FIGURE 7.17 : Distribution des paramètres du modèle sur le sous-échantillon de bassins français pour les jeux de paramètres les plus polyvalents au regard des deux critères associés aux basses eaux (moustaches rouges), des deux critères associés aux eaux intermédiaires (moustaches grises), des deux critères associés aux hautes eaux (moustaches bleues), et des six critères simultanément (moustaches noires).

d'un gain en eau (X_2 positifs). On constate en outre que dans ce sous-échantillon les valeurs de X_2 des jeux de paramètres les plus polyvalents au regard des six critères simultanément sont pour la plupart faiblement négatives. Comme suggéré dans la Section 6.1.4, cette caractéristique s'accorde avec le positionnement amont des bassins versants étudiés par rapport aux bassins plus larges qu'ils alimentent. En Australie néanmoins, le facteur de pertes par échanges souterrains est clairement négatif, et suggère que le modèle repose fortement sur ces pertes pour boucler son bilan hydrique.

Les motifs observés dans les bassins des deux pays diffèrent également à travers la répartition des valeurs du paramètre X_3 selon les gammes de débit considérées. Ce paramètre correspond à la capacité à un jour du réservoir de routage. Dans les bassins français, les capacités du réservoir de routage couvrent un large spectre de valeurs pour la simulation des eaux intermédiaires. En comparaison, les capacités correspondant aux hautes eaux sont généralement plus faibles. À apports en eau donnés, et à facteur d'échanges souterrains constant, un réservoir de routage de plus grande capacité tend à réduire le volume des pertes par le souterrain, ainsi qu'à allonger le temps de récession du débit simulé. Il est possible que la gamme des eaux intermédiaires recoupe en partie des périodes de récession du débit précédant les étiages et que leur simulation

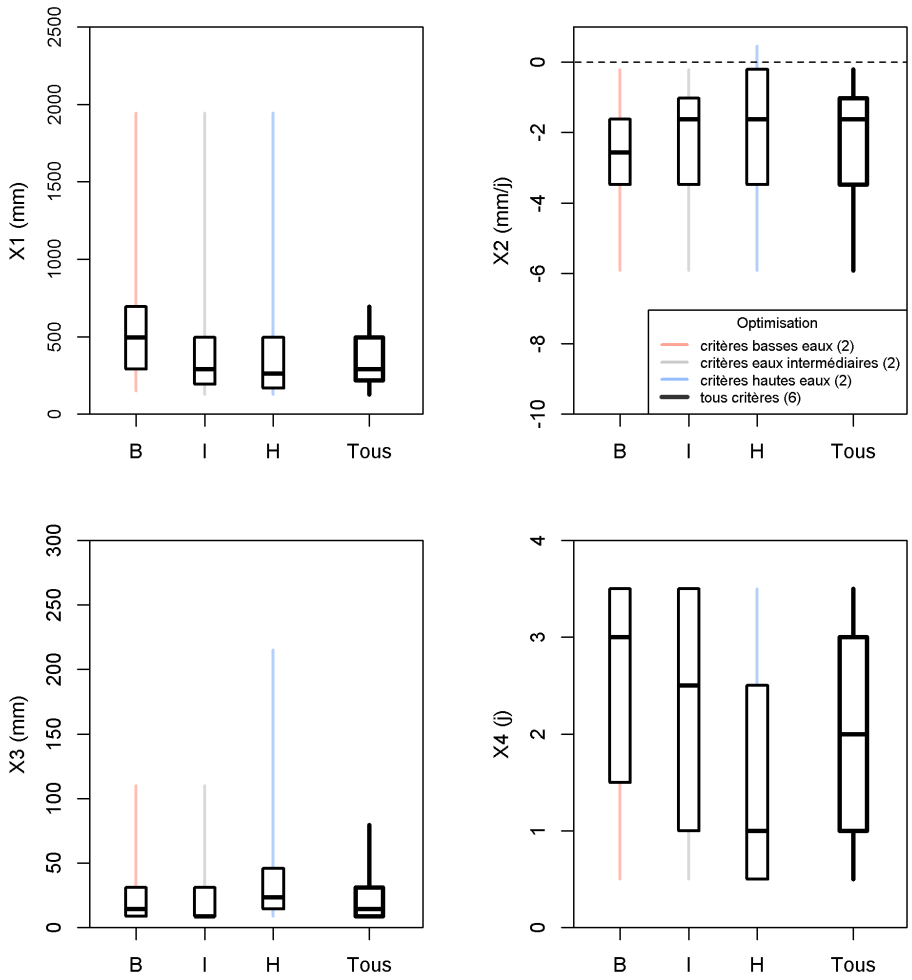


FIGURE 7.18 : Distribution des paramètres du modèle sur le sous-échantillon de bassins australiens pour les jeux de paramètres les plus polyvalents au regard des deux critères associés aux basses eaux (moustaches rouges), des deux critères associés aux eaux intermédiaires (moustaches grises), des deux critères associés aux hautes eaux (moustaches bleues), et des six critères simultanément (moustaches noires).

nécessite un comportement du modèle plus lisse dans le temps. On peut également supposer que les capacités du réservoir sont plus faibles pour la simulation des basses eaux afin de permettre au modèle de perdre suffisamment d'eau par les échanges souterrains pour reproduire les débits d'étiage, même si cela implique de perdre en partie la douceur du comportement du réservoir de routage.
Pour les bassins australiens en revanche, les distributions du paramètre X_3 montrent plutôt une augmentation légère de la capacité à un jour du réservoir de routage à mesure que l'objectif de modélisation concerne des écoulements forts. Ce phénomène est difficile à interpréter simplement à partir de la répartition des paramètres, et nécessite une analyse du comportement du modèle à travers ses états internes.

Il convient de noter que la répartition des valeurs du paramètre X_4, qui correspond au temps de base des hydrogrammes unitaires, est d'un intérêt secondaire dans le cadre de notre problématique, dans la mesure où les performances étudiées concernent des échelles de temps pluri-annuelles. Nous suggérons que les distributions obtenues relèvent davantage d'une indétermination du paramètre dans le processus d'optimisation que de la manifestation d'un processus hydrologique.
L'effet de l'indétermination devrait en outre être pris en considération lors de l'analyse des autres

paramètres, dans la mesure où les jeux de paramètres ont été sélectionnés au regard de couples de critères impliquant les erreurs du modèle sur une unique gamme de débit. Ces gammes de débit représentant des parts variables de l'hydrogramme, selon la gamme et le bassin considéré, il n'est pas exclu que certains de ces trois paramètres aient été dans certains cas insuffisamment contraints dans le processus de sélection des jeux de paramètres les plus polyvalents. Cependant, nous suggérons que l'étude d'un grand échantillon de bassins versants permet malgré tout d'observer les tendances générales sur la façon dont la sélection des paramètres est contrainte par les groupes de critères considérés. Dans cette perspective, la pertinence de la valeur des paramètres pour chaque bassin pris au cas par cas nous semble secondaire. En outre, il n'est pas vraiment possible d'écarter le risque d'indétermination dans une analyse des configurations adoptées par le modèle pour répondre à différents objectifs de modélisation considérés séparément.

D'autres regroupements de critères de performance ont également été évalués mais leurs résultats, peu probants, ne sont pas montrés dans la thèse. Parmi les essais réalisés, la comparaison des jeux de paramètres les plus polyvalents au regard des biais sur les trois gammes de débit et des PMR sur les trois gammes de débit a été menée pour tenter de distinguer des différences entre les configurations du modèle répondant à un objectif de reproduction du bilan de long terme et les configurations du modèle répondant à l'objectif de simulation de la variabilité interannuelle. Aucune conclusion robuste n'a pu être dégagée de cette analyse.

L'analyse des paramètres de GR4J a donc finalement montré que la simulation des différentes gammes de débit fait appel à des configurations du modèle différentes. Ces différences sont, pour la plupart, cohérentes sur les bassins des deux sous-échantillons français et australien. Les différents comportements du modèle que ces configurations supposent sont généralement interprétables physiquement à partir de la physionomie des gammes de débit considérées.
Toutefois, les paramètres du modèle sont d'une part fortement interconnectés et d'autre part leur influences respectives sur le comportement du modèle sont dépendantes des caractéristiques des forçages d'entrée. De ce fait, aucune des interprétations suggérées précédemment ne peut être considérée comme définitive tant qu'elles n'ont pas été étayées par une analyse des états et des flux internes du modèle.

7.3.3. Étude des états internes du modèle GR4J associés aux gammes de débit

Dans la suite, nous avons étudié les flux internes du modèle associés aux jeux de paramètres les plus polyvalents au regard de performances sur chacune des gammes de débit en séparant deux aspects du comportement du modèle.
Dans un premier temps, nous avons cherché à étayer les hypothèses formulées dans l'analyse des paramètres au sujet des variations temporelles de la réponse du modèle. Ces variations incluent les suggestions selon lesquelles le modèle aurait besoin d'être plus réactif pour la simulation des hautes eaux, notamment de moduler rapidement son rendement, ainsi que celle impliquant un comportement plus lisse du réservoir de routage pour la simulation des eaux intermédiaires.
Dans un second temps, notre analyse s'est focalisée sur les moyens employés par le modèle pour moduler son rendement moyen en fonction de la gamme de débit à représenter. Nous avons notamment cherché à confirmer l'hypothèse selon laquelle l'augmentation de la taille du réservoir de production contribue à moduler le rendement moyen du réservoir de production de GR4J.

Avant la démonstration de ces résultats, il convient de clarifier la terminologie employée pour désigner les différentes parties du modèle impliquées dans les discussions. La participation des

échanges souterrains, calculés à partir du remplissage du réservoir de routage, au bilan hydrique de GR4J contribue à brouiller la distinction entre module de production et module de routage du modèle (Section 1.1.3.b).
Pour cette raison, nous désignerons par *fonction* de production l'ensemble des processus du modèle situés en amont des hydrogrammes unitaires, et par *fonction* de routage l'ensemble des hydrogrammes unitaires et des processus situés en aval (Figure 7.19). Le rendement de la fonction de production du modèle correspond à la part des précipitations transmises vers la fonction de routage (précipitations « efficaces », notées P_r), et le rendement de la fonction de routage correspond à la part des précipitations efficaces effectivement écoulée à l'exutoire du bassin simulé. Le rendement du réservoir de production est la part des précipitations nettes (après la prise en compte de l'interception, notées P_n) effectivement transmise vers la fonction de routage.

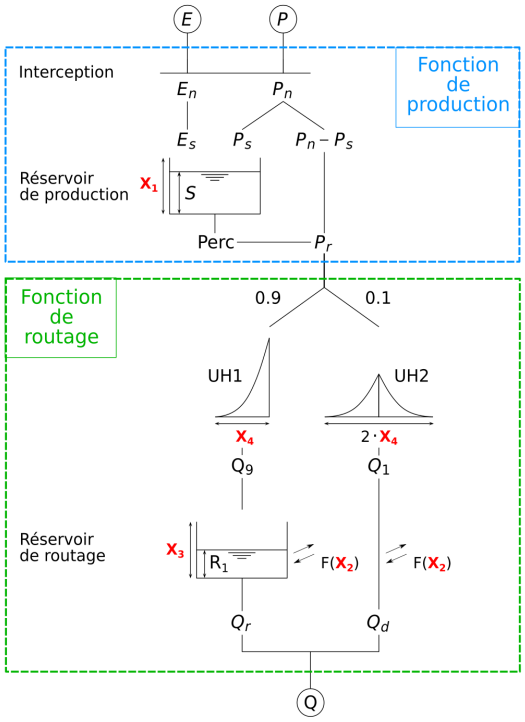


FIGURE 7.19 : Schéma des fonctions de GR4J.

7.3.3.a. Variabilité temporelle des états internes

Pour mieux comprendre l'effet des valeurs des paramètres observé pour les différentes gammes de débit sur la réactivité des réservoirs de GR4J, la distribution des écarts-types du taux de remplissage des réservoirs de production (S/X_1) et la distribution du coefficient de variation de la pluie efficace P_R en sortie de la fonction de production sont présentées en Figure 7.20 pour les bassins français et en Figure 7.21 pour les bassins australiens.
Il apparaît clairement que le remplissage du réservoir de production est moins variable pour la simulation des basses eaux, dans les deux régions étudiées. En considérant les valeurs en moyenne plus élevées du paramètre X_1, ce constat confirme le lien existant entre taille du réservoir et variabilité de son remplissage. Comme attendu, la moindre variabilité du remplissage du réservoir de production pour la simulation des basses eaux est associée à une moindre variabilité des précipitations efficaces en sortie de la fonction de production, par rapport à ce qui est constaté avec les jeux de paramètres des autres gammes de débit.

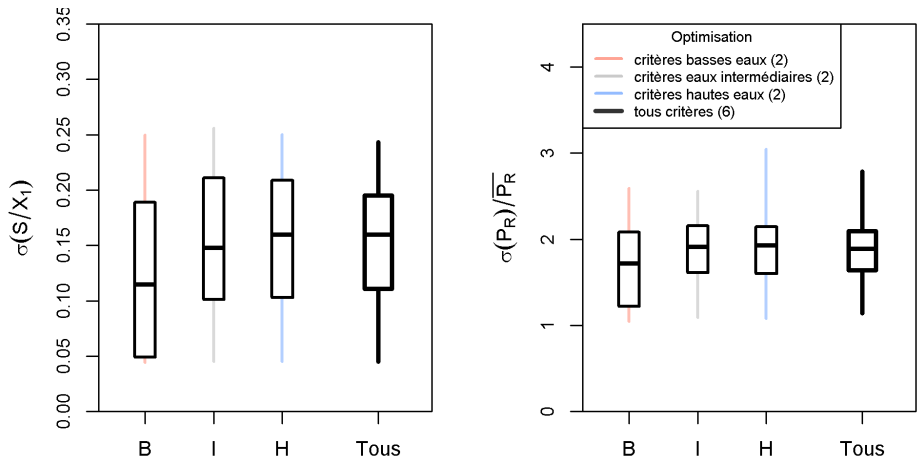


FIGURE 7.20 : Distribution sur le sous-échantillon de bassins français des écarts-types du taux de remplissage du réservoir de production et du coefficient de variation des précipitations efficaces de GR4J, pour les jeux de paramètres les plus polyvalents au regard des deux critères associés aux basses eaux (moustaches rouges), des deux critères associés aux eaux intermédiaires (moustaches grises), des deux critères associés aux hautes eaux (moustaches bleues), et des six critères simultanément (moustaches noires).

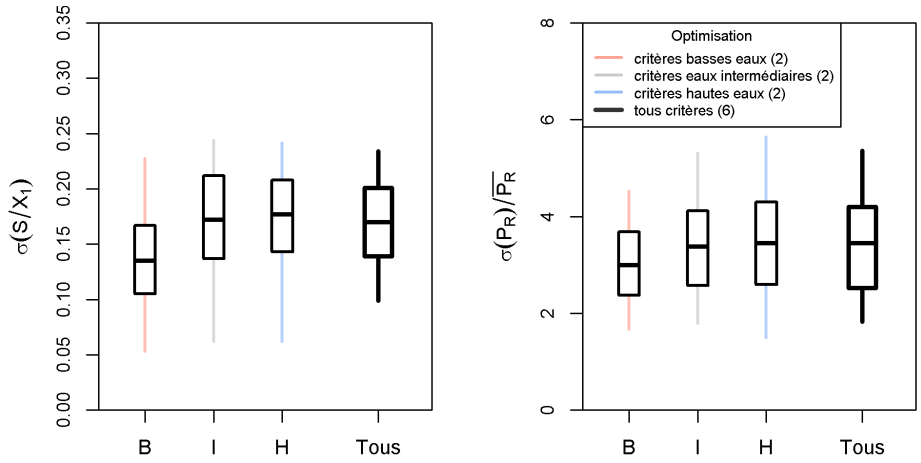


FIGURE 7.21 : Distribution sur le sous-échantillon de bassins australien des écarts-types du taux de remplissage du réservoir de production et du coefficient de variation des précipitations efficaces de GR4J, pour les jeux de paramètres les plus polyvalents au regard des deux critères associés aux basses eaux (moustaches rouges), des deux critères associés aux eaux intermédiaires (moustaches grises), des deux critères associés aux hautes eaux (moustaches bleues), et des six critères simultanément (moustaches noires).

De fait, la fonction de production du modèle agit d'autant plus comme un filtre passe-bas des précipitations que le modèle cherche à simuler les basses eaux. Ce comportement amène à stabiliser les apports en eau alimentant la fonction de routage, et donc contribue vraisemblablement à stabiliser également le comportement du réservoir de routage. Celui-ci est décrit en Figures 7.22 et 7.23, à travers les distributions des écarts-types du taux de remplissage du réservoir de routage, et celles de la variabilité de l'écoulement en sortie du réservoir rapportée à la pluie efficace P_R moyenne issue de la fonction de production.
Il n'est pas aisé de dégager une tendance concernant le rapport qu'entretiennent la variabilité du taux de remplissage du réservoir de routage et les gammes de débit considérées pour la sélection des paramètres. Une caractéristique commune aux deux sous-échantillons de bassins versants est la variation de la dispersion spatiale de l'écart-type du taux de remplissage du réservoir, selon la gamme de débit considérée. Ces différentes dispersions sont cohérentes avec celles des valeurs du paramètre X_3 dans le cas des bassins français, mais sont difficilement explicables en ce qui concerne les bassins australiens (cf. Figures 7.17 et 7.18).

Au contraire de la faible interprétabilité de la variabilité du taux de remplissage du réservoir de routage, on constate une tendance marquée de l'écoulement produit par le réservoir de routage à être davantage variable lorsque la gamme de débits à simuler croît. Pour la simulation des basses eaux, le modèle simule un écoulement Q_R très stable. En France, l'amplitude des variations de l'écoulement en sortie du réservoir de routage est généralement inférieure à la moyenne des entrées de pluie efficace P_R, suggérant un fort lissage des forçages climatiques. En comparaison de l'instabilité de l'écoulement produit par le modèle avec les paramètres correspondant aux hautes eaux, l'écoulement Q_R associé à la modélisation des eaux intermédiaires est également assez stable.

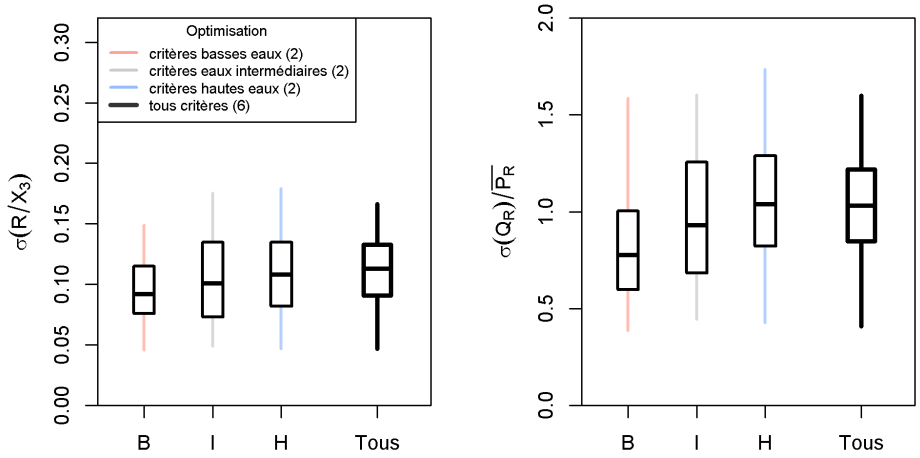


FIGURE 7.22 : Distribution sur le sous-échantillon de bassins français des écarts-types du taux de remplissage du réservoir de routage et de l'écart-type de l'écoulement en sortie du réservoir de routage rapporté aux précipitations efficaces moyennes de GR4J, pour les jeux de paramètres les plus polyvalents au regard des deux critères associés aux basses eaux (moustaches rouges), des deux critères associés aux eaux intermédiaires (moustaches grises), des deux critères associés aux hautes eaux (moustaches bleues), et des six critères simultanément (moustaches noires).

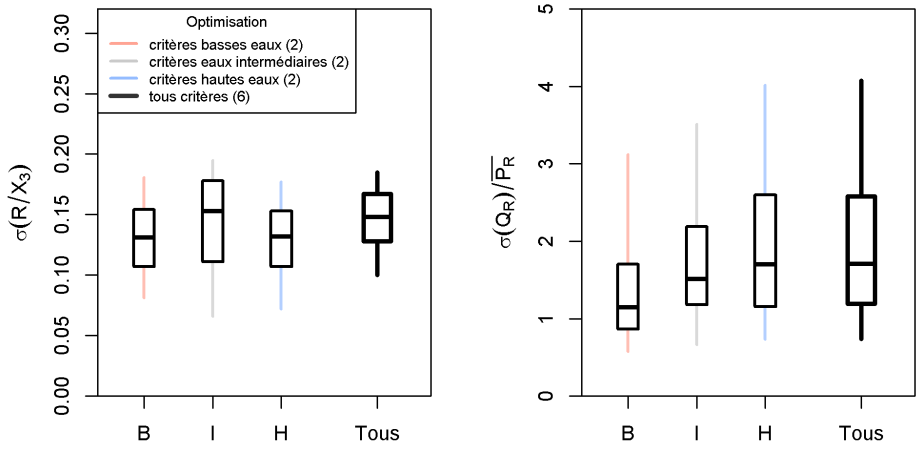


FIGURE 7.23 : Distribution sur le sous-échantillon de bassins australien des écarts-types du taux de remplissage du réservoir de routage et de l'écart-type de l'écoulement en sortie du réservoir de routage rapporté aux précipitations efficaces moyennes de GR4J, pour les jeux de paramètres les plus polyvalents au regard des deux critères associés aux basses eaux (moustaches rouges), des deux critères associés aux eaux intermédiaires (moustaches grises), des deux critères associés aux hautes eaux (moustaches bleues), et des six critères simultanément (moustaches noires).

7.3.3.b. Rendements internes des fonctions de production et de routage

L'influence des paramètres sur le rendement du modèle est analysée à travers les distributions du rendement de la fonction de production et celles du rendement de la fonction de routage en Figures 7.24 et 7.25.

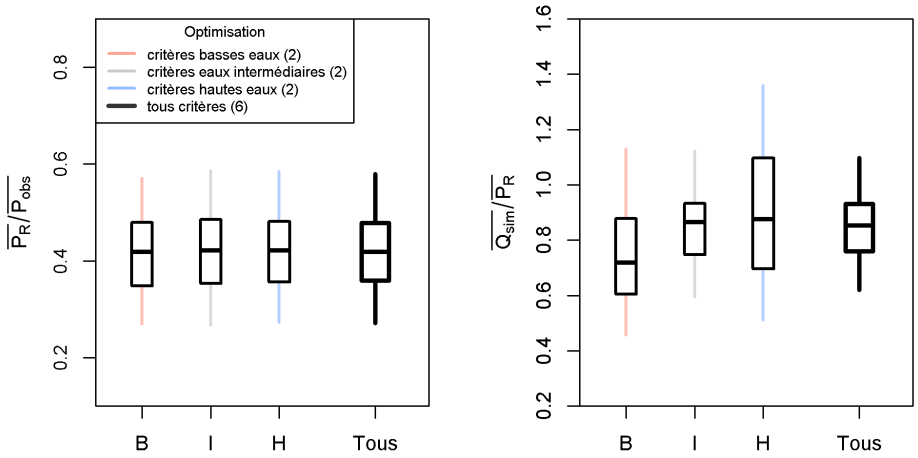


FIGURE 7.24 : Distribution sur le sous-échantillon de bassins français des rendements de la fonction de production et de la fonction de routage de GR4J, pour les jeux de paramètres les plus polyvalents au regard des deux critères associés aux basses eaux (moustaches rouges), des deux critères associés aux eaux intermédiaires (moustaches grises), des deux critères associés aux hautes eaux (moustaches bleues), et des six critères simultanément (moustaches noires).

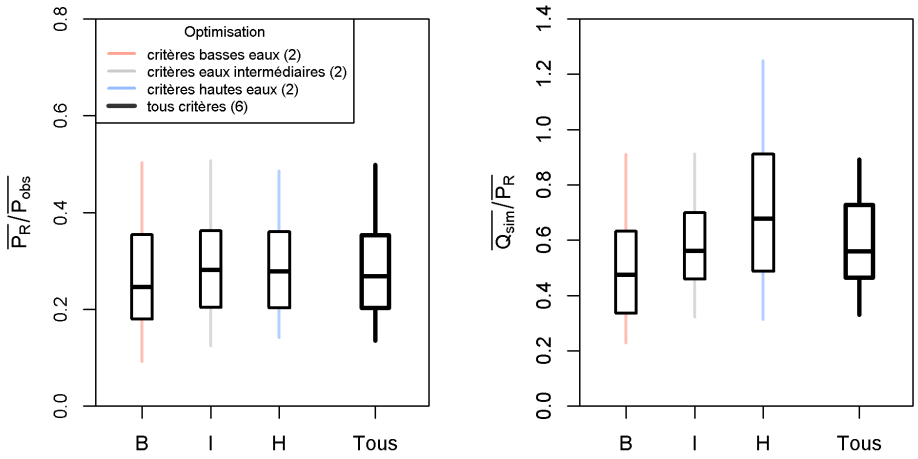


FIGURE 7.25 : Distribution sur le sous-échantillon de bassins australien des rendements de la fonction de production et de la fonction de routage de GR4J, pour les jeux de paramètres les plus polyvalents au regard des deux critères associés aux basses eaux (moustaches rouges), des deux critères associés aux eaux intermédiaires (moustaches grises), des deux critères associés aux hautes eaux (moustaches bleues), et des six critères simultanément (moustaches noires).

De façon surprenante, les rendements de la fonction de production du modèle ne semblent pas dépendre des différences entre les jeux de paramètres correspondant aux trois gammes de débit. Surtout, alors que les tailles des réservoirs de production sont nettement plus grandes pour la simulation des basses eaux, le rendement n'en est pas affecté. Seule une légère baisse est observée dans les cas des bassins australiens. Ce résultat est d'autant plus inattendu que la fonction de production est la seule partie du modèle décrivant un processus d'évaporation sur lequel l'influence de la taille du réservoir de production paraît aller de soi.

En comparaison, le rendement de la fonction de routage suit des distributions très contrastées selon que les simulations sont réalisées avec les jeux de paramètres optimisant une gamme de débit ou une autre. Les écarts majeurs se situent entre le rendement de la fonction de routage avec les jeux de paramètres associés aux basses eaux et ceux associés aux hautes eaux : l'écart entre les médianes des deux distributions atteint plus de 20% sur les deux sous-échantillons.

Ces résultats indiquent clairement que la fonction d'échanges souterrains gère presqu'à elle seule l'ajustement du bilan hydrique de GR4J. Pour comprendre les raisons de l'insensibilité du rendement de la fonction de production du modèle au paramètre X_1, nous avons étudié le comportement des deux termes contribuant au calcul des précipitations efficaces P_R. Pour rappel, ces précipitations efficaces sont le résultat de la somme de la partie des précipitations non interceptées P_n qui n'est pas captée par le réservoir de production et d'un terme de « percolation », qui correspond à la vidange lente du réservoir de production :

$$P_R = (P_n - P_S) + Perc$$

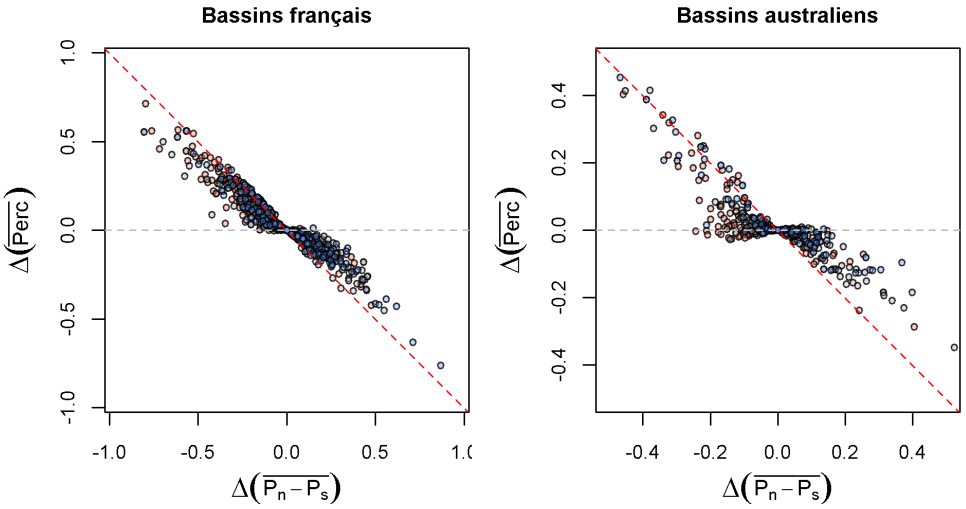


FIGURE 7.26 : Écarts sur les deux sous-échantillons de bassins français et australiens des percolations moyennes en fonction des écarts de précipitations non captées. Les écarts sont calculés pour chaque bassin entre les jeux de paramètres les plus polyvalents au regard des critères de basses eaux (rouge), des deux critères des eaux intermédiaires (gris), des critères de hautes eaux (bleu), d'une part et le jeu de paramètres le plus polyvalent au regard des six critères simultanément.

La Figure 7.26 montre les variations de percolation moyenne en fonction des variations de précipitations non captée ($P_n - P_S$) moyennes. Ces variations sont calculées entre les grandeurs simulées avec les jeux de paramètres associés aux trois gammes de débit et les jeux de paramètres les plus polyvalents au regard des six critères simultanément.
Il apparaît que les variations dans les précipitations non captées moyennes entre les jeux de paramètres sont compensées partiellement voire complètement par des variations opposées de la percolation. Ce phénomène est particulièrement flagrant sur l'échantillon de bassins français. En outre, même dans les bassins australiens où ce phénomène est moins visible, les variations de la percolation moyenne ne sont jamais de même signe que celles des précipitations non captées moyennes. Cet effet stabilise très fortement la réponse moyenne de la production du modèle aux entrées de précipitations.

Pour comprendre les raisons de ces compensations, nous avons dans un premier temps cherché

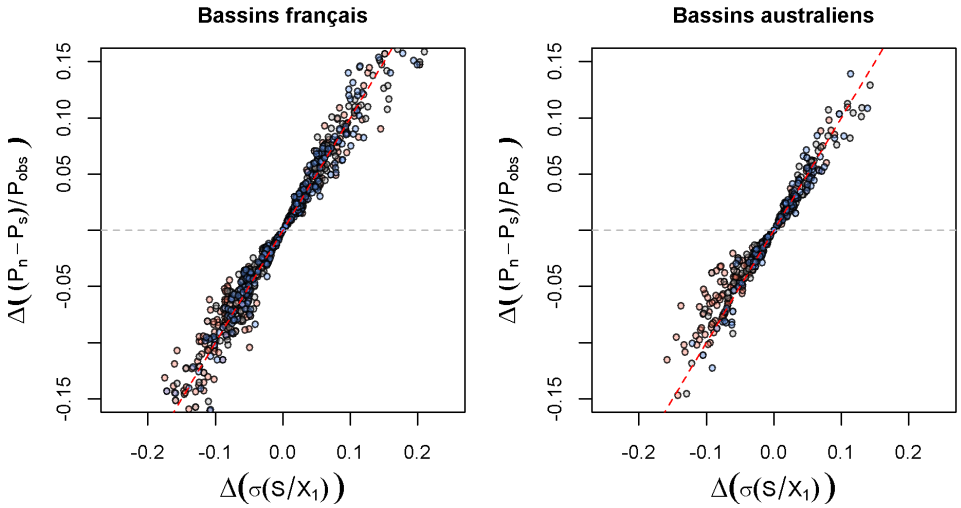


FIGURE 7.27 : Écarts sur les deux sous-échantillons de bassins français et australiens des rendements moyens du réservoir de production liés aux précipitations non captées en fonction des écarts d'écarts-types du taux de remplissage du réservoir. Les écarts sont calculés pour chaque bassin entre les jeux de paramètres les plus polyvalents au regard des critères de basses eaux (rouge), des deux critères des eaux intermédiaires (gris), des critères de hautes eaux (bleu), d'une part et le jeu de paramètres le plus polyvalent au regard des six critères simultanément.

à étudier le comportement du terme de précipitations non captées $P_n - P_S$ dans le rendement du réservoir de production du modèle. Il est apparu que le rendement du réservoir lié aux précipitations non captées est dépendant quasiment linéairement de la variabilité du taux de remplissage du réservoir de production (Figure 7.27). Cette dépendance n'a rien d'évident au regard de l'équation de P_S :

$$P_S = \frac{P_n \left(1 - \left(\frac{S}{X_1}\right)^2\right)}{1 + \frac{P_n}{X_1}\left(1 + \frac{S}{X_1}\right)}$$

La contribution des précipitations non captées au rendement est dépendante de la variabilité du remplissage du réservoir, dont on a constaté précédemment qu'il était influencé par la taille du réservoir X_1, de sorte que les petits réservoirs étaient généralement les plus variables.

Pour expliquer le comportement de la percolation, nous rappelons son équation :

$$Perc = S\left(1 - \frac{1}{\left(1 + \left(\frac{4}{9}\frac{S}{X_1}\right)^4\right)^{\frac{1}{4}}}\right)$$

Pour un taux de remplissage donné du réservoir de production, la percolation du réservoir de production dépend linéairement de la quantité d'eau dans le réservoir. C'est pourquoi la percolation augmente en moyenne lorsque la taille du réservoir augmente. Toutefois, la structure en puissance de l'équation de la percolation la fait décroître très rapidement lorsque le réservoir est peu rempli. Étant donné que le taux de remplissage moyen du réservoir dans les bassins australiens est plus faible que le taux de remplissage dans les bassins français (0,43 contre 0,54, quelle que soit la gamme de débit considérée pour la sélection des jeux de paramètres), l'effet de compensation de la percolation sur le rendement du réservoir de production est atténué (Figure 7.26.b).

7.3.4. Synthèse et discussion de l'analyse de la structure de GR4J

7.3.4.a. Problèmes relevés dans la structure du modèle

Les résultats de notre analyse indiquent clairement que la fonction de production n'a pas une influence directe sur le rendement du modèle. La quantité d'eau moyenne filtrée par le réservoir de production ne dépend pas du paramètre X_1, à cause d'une compensation entre les contributions de la percolation, qui augmente avec la taille du réservoir, et des précipitations non captées par le réservoir, qui diminuent avec la taille du réservoir.
De ce fait, la fonction d'échanges souterrains de la fonction de routage gère presque seule l'ajustement du bilan hydrique de GR4J. Il faut néanmoins noter que la fonction de production du modèle joue un rôle de régulateur dans le temps des entrées de précipitations efficaces dans la partie routage du modèle, et donc qu'elle est susceptible d'avoir un effet indirect sur le bilan hydrique du modèle. On peut supposer que les réservoirs de production plus petits, qui atteignent vite la saturation, laissent ruisseler une partie plus importante des précipitations lors des phases où le réservoir de routage est rempli, ce qui contribue à augmenter le volume des échanges souterrains.

La dépendance de GR4J au processus d'échanges souterrains pour moduler son rendement peut apparaître discutable. Pour mieux comprendre l'implication des hypothèses de modélisation de GR4J, considérons deux bassins dont les précipitations et la demande évaporatoire moyennes seraient identiques mais dont d'autres caractéristiques telles que le couvert végétal, la topographie, la saisonnalité des forçages ou le type de sol, seraient différentes. Dans cette situation, il serait vraisemblable que ces bassins n'évaporent pas la même quantité d'eau, et que ces réponses différenciées aux forçages climatiques induisent deux bilans hydriques différents. Sur de tels bassins, le modèle serait pourtant incapable d'expliquer les différences de bilan hydrique par des rendements de l'évaporation potentielle différents, puisque son bilan n'est géré directement que par la fonction d'échanges souterrains.
Ceci ne pose pas uniquement un problème évident de réalisme physique des processus modélisés, mais également un problème de robustesse en contexte de changement climatique. En effet, il n'est pas certain que la fonction d'échanges souterrains puisse fournir une réponse adaptée à des modifications très importantes de la demande évaporatoire, ou du volume des précipitations. Cette analyse est corroborée par les travaux de Fowler *et al.* (2020), qui ont notamment montré que GR4J n'était pas capable de reproduire les effets sur le bilan de bassins australiens d'une baisse drastique et prolongée des précipitations pendant la *Millenium Drought*.
Cependant, il convient de nuancer le propos en rappelant que le modèle est assez apte à garantir des biais faibles dans les simulations de débits de plusieurs gammes simultanément, malgré une fonction unique de gestion de son bilan. Ce résultat indique une excellente coordination entre la gestion temporelle des apports en eau réalisée par la fonction de production et le facteur d'intensité des échanges souterrains, et témoigne en un sens d'une certaine harmonie dans la structure du modèle. En réalité, le problème principal posé par la structure actuelle du modèle n'est donc pas la possibilité pour le modèle de simuler un bilan hydrique moyen et une réponse saisonnière adaptée, mais sa relative incapacité à nuancer cette réponse en fonction des conditions climatiques.

7.3.4.b. Voies d'amélioration du modèle

Au vu de ces conclusions, plusieurs voies se dessinent pour amender GR4J et en améliorer la robustesse.

Premièrement, il peut être envisagé de partager la responsabilité de la gestion du bilan hydrique entre la fonction de production et la fonction de routage. Pour y parvenir, il est nécessaire de modifier

la production du modèle. Le rendement moyen actuel de la production du modèle est caractérisé par (i) une quasi insensibilité de sa moyenne à la taille du réservoir à cause de compensations entre percolation et précipitations non captées et (ii) une forte dépendance de sa variabilité à la taille du réservoir qui a un effet indirect sur le rendement de la fonction de routage. On peut donc suggérer que la modification de la production du rendement doive permettre la modulation de son rendement moyen, tout en respectant autant que possible la flexibilité de son comportement en filtre passe-bas.
Quelques travaux ont déjà été menés sur le paramétrage de la production du modèle. Ceux de Westra *et al.* (2014) impliquent par exemple l'implémentation d'une dépendance temporelle de la taille du réservoir, par exemple aux conditions climatiques des mois précédents, ou l'ajout d'un paramètre de forme à l'équation des précipitations captées par le réservoir du production. Les modifications proposées par Hughes *et al.* (2013) consistent également dans l'incorporation de plusieurs paramètres de forme aux équations des précipitations captées et à l'évaporation réelle du réservoir de production, ainsi que l'inclusion de mesures de l'évolution de l'indice foliaire des bassins modélisés pour moduler son comportement. Ces versions adaptées de GR4J n'ont été cependant expérimentées par leurs auteurs que pour une poignée de bassins versants. Toutefois, les résultats de Fowler *et al.* (2016) suggèrent que la version du modèle proposée par Hughes *et al.* (2013) pourrait être plus adaptée que le modèle original sur les bassins australiens ayant subi la *Millenium Drought*.

Deuxièmement, un axe d'amélioration consiste à réduire les difficultés éprouvées par le modèle à représenter parallèlement les caractéristiques des écoulements de basses eaux et des eaux intermédiaires d'une part, plutôt stables, et celles des hautes eaux d'autre part, plus instables, qui demandent au modèle d'être plus réactif. Pour y parvenir, l'ajout d'un réservoir de routage supplémentaire ayant un fonctionnement différent du réservoir actuel est une option envisageable. En ce sens, le modèle GR6J proposé par Pushpalatha *et al.* (2011), qui possède un réservoir exponentiel en parallèle du réservoir de routage de GR4J, s'inscrit dans cette démarche avec un certain succès pour la simulation des récessions et des étiages.

7.4. Discussion de la méthode mise en œuvre

7.4.1. Possibilités d'application de la méthode à d'autres modèles

La méthode mise en œuvre a conduit à dégager des résultats exploitables pour la modification de GR4J. Appliquée à d'autres modèles, la démarche pourrait peut-être aussi s'avérer bénéfique. Beaucoup de modèles sont en effet vraisemblablement affectés par des compétitions entre leurs performances, rendant l'optimisation de leurs paramètres plus difficile et limitant leur capacité générale à simuler le régime hydrologique des bassins versants. Toutefois, la plupart des modèles hydrologiques sont plus complexes que GR4J, ce qui soulève un certain nombre de questions quant à la possibilité de l'application de la méthode développée dans ce chapitre.

La méthode tire avant tout son intérêt des possibilités qu'elle offre de travailler sur un grand échantillon de bassins versants, dans la mesure où elle permet en partie de se passer d'une analyse visuelle des performances vis-à-vis de multiples critères du modèle sur chaque bassin versant. Néanmoins, le coût en calcul important que suppose l'exploration de l'espace de paramètres du modèle limite les possibilités d'appliquer la méthode sur de nombreux bassins si le modèle est trop complexe. Il est possible que des modèles communs, qui ont généralement plus de 6

ou 7 paramètres libres, soient trop complexes pour envisager d'appliquer la méthode à plus de quelques dizaines de bassins versants. Dans cette situation, une analyse multi-critères visuelle serait peut-être plus informative que notre approche, qui ne donne qu'un aperçu des compétitions entre les critères de performance.
En outre, les paramètres des modèles plus complexes sont généralement à la fois davantage interconnectés, ce qui réduit les possibilités d'interpréter l'effet de leurs valeurs sur le comportement du modèle, et davantage sujets à l'indétermination, ce qui augmente le risque d'interprétation hasardeuse des valeurs obtenues en optimisant un sous-ensemble de critères de performance. Ce deuxième aspect est d'ailleurs important à prendre en considération même pour un modèle simple comme GR4J.
Enfin, les modules de production et de routage sont dans certains modèles encore plus difficiles à dissocier que dans la structure de GR4J. Dans la structure du modèle TOPMO par exemple, l'évaporation affecte un réservoir exponentiel qui contribue directement au débit à l'exutoire, dont l'appartenance au routage ou à la production n'est pas claire. De manière générale, l'analyse des états internes et des flux d'un modèle nécessite l'étude simultanée de tous les réservoirs, ce qui dans le cas de GR4J était finalement assez simple.

Pour ces raisons, nous suggérons que la méthode est plus pertinente pour les modèles parcimonieux, qui gagneraient à être complexifiés. Pour les modèles complexes possédant de nombreux paramètres, qui bénéficieraient plutôt d'une simplification de leur structure et de leur paramétrage, d'autres approches seraient plus avantageuses. Parmi ces approches, l'analyse de sensibilité permet notamment d'identifier les paramètres superflus et les groupes de paramètres les plus influents sur certains aspects des performances d'un modèle.
Malgré tout, nous pensons que notre approche invite à étudier les rapports qu'entretiennent les critères de performance non seulement avec les paramètres du modèle mais également avec ses états internes. La compréhension d'un modèle qu'une telle approche peut permettre est dans tous les cas essentielle pour éviter d'une part à tout modélisateur d'accorder une confiance démesurée au(x) modèle(s) qu'il utilise, et notamment au réalisme des processus simulés, et d'autre part pour aiguiller les tentatives de correction d'une structure de modèle vers des solutions pertinentes.

7.4.2. Discussion sur le calcul de la polyvalence du modèle

La possibilité d'application de la méthode à d'autres modèles que GR4J suppose en partie de réduire les temps de calcul liés à l'exploration de l'espace des paramètres, et donc de réduire le nombre de jeux de paramètres finalement évalués. Notre approche repose sur le classement de ces jeux de paramètres entre eux pour estimer la polyvalence du modèle et la compatibilité entre les différents critères de performance. De ce fait, un même jeu de paramètres peut être positionné différemment selon la méthode d'échantillonnage retenue. Un autre aspect important est le besoin d'échantillonner des jeux de paramètres proches des fronts de Pareto des performances. Pour ces raisons, il semble légitime de questionner la robustesse de la méthode développée dans ce chapitre à la technique d'échantillonnage choisie pour explorer l'espace des paramètres.

Pour vérifier la dépendance des résultats obtenus à la densité de l'exploration de l'espace des paramètres, nous avons conduit à nouveau l'analyse à partir d'une liste de 558 jeux de paramètres issus du calage du modèle sur chacun des bassins de l'échantillon avec le KGE racine.
Estimée ainsi, la polyvalence moyenne du modèle est légèrement inférieure d'environ 1% à celle calculée précédemment. Un échantillonage réduit conduit donc logiquement à sous-estimer les capacités du modèle. Toutefois, les répartitions de la polyvalence du modèle et des compatibilités entre les critères de performance restent assez proches de celles obtenues avec les 9317 jeux de paramètres, comme en attestent les fortes corrélations répertoriées dans les Tableaux 7.6 et 7.7. Il

semblerait que les estimations de la polyvalence du modèle soient particulièrement peu affectées par le choix de l'échantillonnage de l'espace des paramètres.
Même s'il faut malgré tout reconnaître que certains résultats de la méthode peuvent être dépendants de l'échantillonnage réalisé (par exemple la polyvalence du modèle sur les bassins australiens), les principales conclusions de notre analyse émergent malgré tout de l'analyse effectuée avec un échantillon réduit de jeux de paramètres. Ainsi, la compatibilité moyenne sur l'échantillon australien des critères de biais entre eux atteint 95,2%, contre 91,5% pour la compatibilité des PMR entre eux (résultat non montré). Le constat d'une difficulté particulière de GR4J à simuler simultanément les variations interannuelles du débit sur plusieurs gammes à la fois reste donc le même.
Ces résultats tendent à montrer que notre approche est plutôt robuste au nombre de jeux de paramètres impliqués.

TABLEAU 7.6 : Corrélations de Spearman entre les compatibilités des critères de biais ou de robustesse estimées avec l'échantillon de 9317 jeux de paramètres et celles estimées avec l'échantillon de 558 jeux de paramètres.

	Corrélation entre les compatibilités estimées		
	Tous critères	$Biais_{B,I,H}$	$PMR_{B,I,H}$
Bassins français	0,71	0,90	0,92
Bassins australiens	0,53	0,88	0,92

TABLEAU 7.7 : Corrélations de Spearman entre les compatibilités des critères de chaque gamme de débit estimées avec l'échantillon de 9317 jeux de paramètres et celles estimées avec l'échantillon de 558 jeux de paramètres.

	Corrélation entre les compatibilités estimées		
	$Biais_B, PMR_B$	$Biais_I, PMR_I$	$Biais_H, PMR_H$
Bassins français	0,69	0,74	0,72
Bassins australiens	0,87	0,91	0,86

Il convient de noter qu'une rationalisation de l'estimation de la sévérité des compétitions entre les performances a également été tentée à partir de considérations géométriques sur les fronts de Pareto, en s'appuyant notamment sur le calcul de distances euclidiennes dans l'espace des performances. L'objectif de cette approche était de réduire la dépendance des résultats à un classement relatif des jeux de paramètres entre eux, et donc la dépendance à la technique d'échantillonnage des jeux de paramètres. Les résultats obtenus se sont cependant révélés être très influencés par les différences entre les gammes de variations des critères, notamment entre biais et PMR d'une même gamme. Cette méthode a donc été rejetée.

7.5. Synthèse

Objectifs

Les travaux menés dans ce chapitre étaient consacrés au diagnostic des lacunes structurelles du modèle GR4J, et en particulier à l'identification des composantes du modèle impliquées dans des compétitions entre plusieurs critères de performances dans le cadre de l'optimisation des paramètres du modèle. Ce diagnostic a été réalisé afin de préparer une tentative de modification de la structure du modèle pour en améliorer la robustesse.
Une méthode de diagnostic fondée sur une rationalisation des analyses multi-critères a été proposée, permettant de fournir une estimation de l'ampleur des limitations des performances du modèle à cause des compétitions entre critères. La méthode permet de se passer d'analyses visuelles des fronts de Pareto entre performances du modèle. Les critères choisis pour l'application à GR4J concernait la capacité du modèle à simuler la moyenne (biais) et les variations interannuelles (PMR) du débit sur trois gammes d'écoulement, allant des basses eaux aux hautes eaux.

Méthode mise en œuvre

La méthode de diagnostic proposée repose sur l'exploration de l'espace des paramètres et sur le classement des jeux de paramètres en fonction de leurs performances au regard de plusieurs critères considérés simultanément. L'identification des critères de performance les plus difficiles à optimiser conjointement pour le modèle est réalisée en comparant entre elles les valeurs d'une métrique *ad hoc* de compatibilité entre ces critères, calculée à partir de la moyenne des rangs des jeux de paramètres pour les critères considérés. Les performances réelles du modèle obtenues avec les jeux de paramètres sélectionnés pour optimiser différents groupes de critères incompatibles entre eux sont également comparées afin de quantifier les effets des compétitions entre ces groupes sur les performances du modèle.
L'identification des composantes du modèle impliquées dans ces compétitions est menée à partir de la comparaison des distributions des jeux de paramètres associés à chaque groupe de critères, puis à partir de l'analyse des états internes et des flux du modèle.

Diagnostic de la polyvalence de GR4J

Nos résultats confirment que des compétitions entre les critères limitent bel et bien l'optimisation des performances de GR4J. Le modèle se montre généralement capable de simuler les trois gammes de débit avec des erreurs modérées sur les volumes moyens écoulés, bien qu'il ait la possibilité d'être non biaisé dans chaque gamme prise isolément. Les compétitions les plus sévères entre performances ont davantage lieu entre les critères de PMR qu'entre les critères de biais. Ces compétitions se traduisent par une baisse de la qualité des variations interannuelles simulées sur chacune des gammes de débit.
Des compétitions n'existent pas seulement entre les performances du modèle pour chaque gamme de débit, mais également entre la capacité du modèle à reproduire le débit moyen et ses variations à l'intérieur d'une même gamme. Toutefois, la baisse des performances associée à ces compétitions intra-gammes semble plus faible que celle liée aux compétitions inter-gammes.
Le diagnostic du modèle montre donc que le modèle éprouve des difficultés particulières à représenter simultanément les variations interannuelles des débits de plusieurs gammes à la fois.

Lacunes structurelles identifiées de GR4J

L'analyse des paramètres et des états internes de GR4J indique que la capacité limitée de GR4J à être simultanément robuste sur plusieurs gammes de débit pourrait être liée à plusieurs raisons.
La première raison est une impossibilité pour le modèle de simuler les comportements lents et stables caractérisant l'écoulement des basses eaux et dans une moindre mesure des eaux intermédiaires, tout en montrant suffisamment de réactivité pour la représentation des hautes eaux. Ce constat se traduit par un comportement de la fonction de production en filtre passe-bas dont la dynamique est d'autant plus lente que les débits simulés sont bas. Cette lenteur contribue à stabiliser les apports alimentant la fonction de routage du modèle, elle-même faiblement réactive. Ce comportement est incompatible avec celui exigé par la simulation adéquate des hautes eaux. En définitive, les différentes configurations nécessitées par la modélisation des différentes gammes de débit supposent des valeurs de jeux de paramètres différentes et de ce fait coexistent difficilement. Une possibilité pour remédier à ce problème pourrait être l'ajout d'un réservoir supplémentaire au routage pour permettre au modèle de produire des écoulements de temps caractéristiques différents.
La deuxième raison est la dépendance du modèle à sa fonction d'échanges souterrains comme unique levier d'action direct pour ajuster son bilan hydrique. Des compensations entre les termes contribuant au rendement du réservoir de production le rendent en effet quasiment insensible aux variations des paramètres du modèle. Cette dépendance du modèle aux échanges souterrains pose d'une part un problème de réalisme des processus modélisés, notamment en ce qui concerne l'évaporation, et d'autre part est susceptible de rigidifier la réponse du modèle aux variations des conditions climatiques. Une possibilité pour amender la structure du modèle serait d'inclure un ou plusieurs paramètres à la fonction de production ayant un effet sur son rendement moyen, sans altérer trop fortement son rôle de filtre passe-bas des précipitations.
Au chapitre suivant, une modification de la production de GR4J est proposée et évaluée.

Chapitre 8.

Modifications de la puissance de la production du modèle GR4J

Introduction

Dans ce chapitre, nous proposons une modification de la structure de GR4J et comparons les performances du modèle modifié au modèle GR4J original. La modification opérée concerne un facteur de forme de la partition des précipitations et de l'évaporation dans le réservoir de production. Les performances des deux modèles sont évaluées sur plusieurs gammes de débit, au moyen notamment des critères définis au chapitre précédent. Le comportement du modèle modifié est ensuite discuté et critiqué, notamment afin de déterminer les types de bassins qui bénéficient de ce degré de liberté supplémentaire.

8.1. Principes et méthodologie

8.1.1. Motivations et présentation de la modification

8.1.1.a. Motivations

Les modifications du modèle GR4J opérées dans ce chapitre sont fondées sur les résultats de l'analyse des lacunes relevées dans le chapitre précédent. Ces lacunes expliquent notamment les difficultés du modèle à simuler de manière simultanément robuste les variations interannuelles du débit sur plusieurs gammes d'écoulement. L'une de ces lacunes semble être la faible sensibilité du rendement moyen de la fonction de production à la taille de son réservoir, et donc sa faible sensibilité aux paramètres du modèle. Le rôle de régulateur direct du bilan hydrique du modèle incombe de ce fait à la fonction d'échange souterrain. Outre son manque possible de réalisme, ce fonctionnement soulève des interrogations concernant sa robustesse en climat changeant.

Pour toutes ces raisons, il peut être judicieux de considérer l'ajout d'un degré de liberté supplémentaire à la fonction de production du modèle, afin d'ajuster au mieux la réponse hydrologique du bassin modélisé aux conditions climatiques et ainsi moduler son rendement moyen. Le choix du paramétrage s'est porté sur un paramètre « caché » de GR4J, identifié dans les travaux de Santos *et al.* (2018a). Celui-ci concerne la puissance dans l'équation qui lie le rendement du réservoir à son remplissage. Dans GR4J, ce réservoir de production capte en effet une partie des précipitations et de l'évaporation en fonction de son remplissage et filtre ainsi les précipitations qui participent à l'écoulement. A l'échelle temporelle continue, une loi non-linéaire (quadratique) fixe cette relation. Ce paramétrage quadratique garantit l'intégrabilité de la relation sur un pas de temps discret, suggère qu'un réservoir bien rempli laisse une quantité d'eau importante accessible à l'évaporation et qu'au contraire un réservoir vide capte une plus grande partie des précipitations. En modifiant ces relations, il est possible d'ajuster la non-linéarité de la réponse du réservoir de production aux conditions météorologiques, au prix de l'intégrabilité de l'équation. En d'autres termes, l'ajout d'un paramètre libre au modèle peut permettre de moduler le rendement du réservoir de production.

8.1.1.b. Équations du modèle

Nous proposons une version de GR4J où les deux relations du rendement du réservoir de production dépendent d'un paramètre α à un pas de temps infinitésimal δt. Pour rappel, les équations de GR4J présentées au Chapitre 3 sont le résultat d'une intégration à un pas de temps journalier d'équation continues. Les équations continues des précipitations captées et de l'évaporation

réservoir de production du modèle s'écrivent ainsi :

$$\frac{\delta S}{\delta t} = P_S = P_n \left(1 - \left(\frac{S}{X_1}\right)^{\alpha}\right) \tag{8.1}$$

$$\frac{\delta S}{\delta t} = E_S = E_n \left(2\frac{S}{X_1} - \left(\frac{S}{X_1}\right)^{\alpha}\right) \tag{8.2}$$

$$\text{où } \alpha \geq 1$$

Dans GR4J, le paramètre α est fixé à 2, afin de garantir l'intégrabilité de ces équations au pas de temps journalier, et ne constitue donc pas un paramètre libre. Faire varier α revient à moduler le rendement de la pluie nette et de l'évaporation dans le réservoir. Dans un premier temps, on laissera arbitrairement α varier dans l'intervalle $[1\,;3]$. Si $\alpha = 1$, le modèle répartit la pluie nette et l'évaporation en fonction de son remplissage selon une loi linéaire. Si $\alpha = 3$, le modèle est très non-linéaire et produit abondamment de l'évaporation réelle tout en se retenant de produire de l'écoulement.
Dans la suite, l'équation (8.1) sera appelée « partition des précipitations nettes », et l'équation (8.2) sera appelée « partition de l'évaporation nette ». Il convient de remarquer que contrairement à la partition des précipitations nettes, la partition de l'évaporation nette ne divise pas l'évaporation nette en deux termes utiles dans la mesure où le reliquat d'évaporation nette après calcul de E_S n'est pas réutilisé. Le choix du terme « partition » plutôt que « rendement » répond au besoin de distinguer chacun des processus à l'œuvre dans le réservoir de production et le rendement du réservoir de production, qui est le résultat de tous ces processus.

Graphiquement (Figure 8.1), l'identité du paramètre α pour le terme de pluie et pour le terme d'évaporation implique que les deux courbes de partition du réservoir en fonction de son remplissage sont symétriques par une symétrie centrale de centre $(0.5\,;0.5)$. Cette symétrie repose sur l'hypothèse qu'un bassin peu enclin à la production d'écoulement ($\alpha > 2$, retenant une grande partie de l'eau précipitée) laisse l'eau retenue accessible à l'évaporation (à taux quasi potentiel), et qu'inversement un bassin très enclin à la production d'écoulement ($\alpha < 2$, dont une grande partie de l'eau précipitée ruisselle) oppose une plus grande résistance à l'évaporation.

Nous avons nommé GR4J_sym, pour « symétrique », cette version de GR4J à 5 paramètres. Le modèle ainsi défini n'est pas intégrable analytiquement, même en résolvant chaque terme séquentiellement comme cela est fait dans GR4J. La résolution employée est donc celle développée par Santos *et al.* (2018a) en suivant la recommandation de Clark et Kavetski (2010), à savoir un schéma d'Euler implicite, combiné à un algorithme de division adaptative des pas-de-temps en sous-pas de temps (Press *et al.*, 1992). Selon Santos *et al.* (2018a), l'approche d'Euler implicite permet de résoudre les équations différentielles du modèle simultanément, et la division en sous-pas de temps limite les problèmes d'instabilité numérique (division limitée à 100 sous-pas de temps). Il convient de noter que pour des raisons de temps de calcul, seules les équations différentielles de la fonction de production du modèle (i.e. la partie à l'amont des hydrogrammes unitaires) font l'objet d'une résolution numérique, tandis que la partie routage est résolue analytiquement comme dans le modèle GR4J original. Dans la suite, on nommera GR4J « intégré » le modèle original (i.e. dont les équations différentielles sont résolues analytiquement), GR4J « numérique » le modèle GR4J_sym dont la fonction de production est résolu numériquement pour $\alpha = 2$ et simplement GR4J_sym lorsque le paramètre α est libre ou fixé à une autre valeur que 2. La dénomination « GR4 state space » introduite par Santos *et al.* (2018a) n'a pas été retenue dans la mesure où la fonction de routage est ici résolue analytiquement quel que soit le modèle envisagé.

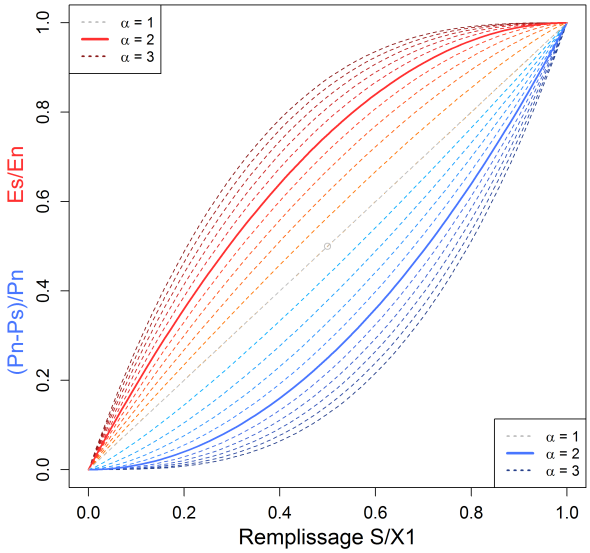


FIGURE 8.1 : Comportement des partitions des précipitations nettes et de l'évaporation nette du réservoir de production de GR4J en fonction du niveau du réservoir de production selon différentes valeurs de α (une courbe par incrément de 0,2). Plus α tend vers 1, plus le comportement du modèle se rapproche de la bissectrice. Le point $(0,5\,;0,5)$ est centre de symétrie pour les courbes de partition des précipitations et de l'évaporation.

8.1.2. Méthodologie d'évaluation

L'évaluation de GR4J_sym a été réalisée en deux temps. Nous avons en premier lieu choisi de comparer le modèle à GR4J en calage sur l'ensemble des données disponibles des bassins versants de l'échantillon. Cette première évaluation permet de placer les modèles dans une situation aussi favorable que possible pour exploiter au mieux les potentialités qu'offrent leur structures respectives, et ainsi constater les gains maximaux obtenus avec le nouveau paramétrage. Cette étape n'est néanmoins pas suffisante dans la mesure où elle ne discrimine pas la robustesse des structures respectives des modèles dans des conditions climatiques changeantes. Une deuxième évaluation consiste donc à tester les modèles à partir d'expériences de *Differential Split-Sample Test*.

8.1.2.a. Méthode d'évaluation du modèle en calage

Pour cette première étape d'analyse, les modèles ont été calés sur la totalité des chroniques disponibles des bassins de l'échantillon, et leurs performances ont été évaluées sur cette même période. Les résultats obtenus sont donc représentatifs des performances du modèle en simulation sur une longue période susceptible de contenir des conditions climatiques variables. Une fonction objectif adaptée du KGE racine, le RKGE racine, a été employée. Le RKGE est construit comme le KGE mais possède un terme supplémentaire, ajouté aux termes de corrélation, de biais, et de rapport des variabilités du simulé et observé. Le quatrième terme est le PMR_1, c'est-à-dire calculé sur des sous-périodes de un an, censé contraindre les paramètres à mieux représenter les variations inter-annuelles du débit. Le RKGE racine correspond au RKGE calculé sur la racine carrée des débits observés et simulés :

$$RKGE\left[\sqrt{Q}\right] = 1 - \sqrt{\left(r\left[\sqrt{Q}\right]-1\right)^2 + \left(\frac{\sigma_s\left[\sqrt{Q}\right]}{\sigma_o\left[\sqrt{Q}\right]}-1\right)^2 + \left(\frac{\mu_s\left[\sqrt{Q}\right]}{\mu_o\left[\sqrt{Q}\right]}-1\right)^2 + \left(PMR_1\left[\sqrt{Q}\right]\right)^2} \quad (8.3)$$

$$\text{où } PMR_1\left[\sqrt{Q}\right] = \frac{1}{N}\frac{1}{\mu_o\left[\sqrt{Q}\right]}\sum_{i=1}^{N}\left|\left(\mu_{s,i}\left[\sqrt{Q}\right]-\mu_{o,i}\left[\sqrt{Q}\right]\right)-\left(\mu_s\left[\sqrt{Q}\right]-\mu_o\left[\sqrt{Q}\right]\right)\right| \quad (8.4)$$

Pour rappel, $r\left[\sqrt{Q}\right]$ correspond à la corrélation entre les racines carrées des débits simulés et observés, $\sigma_s\left[\sqrt{Q}\right]$ et $\sigma_o\left[\sqrt{Q}\right]$ sont les écarts-type des racines carrées des débits simulés et observés, et $\mu_s\left[\sqrt{Q}\right]$ et $\mu_o\left[\sqrt{Q}\right]$ sont les moyennes des racines carrées des débits simulés et observés. Les termes $\mu_{s,i}\left[\sqrt{Q}\right]$ et $\mu_{o,i}\left[\sqrt{Q}\right]$ désignent les moyennes des racines carrées des débits simulés et observés sur la sous-période d'un an i. Le $\text{PMR}_1\left[\sqrt{Q}\right]$ correspond donc à l'aire sous la courbe des biais glissants de la racine carrée des débits simulés par rapport à la racine carrée des débits observés, pour N sous-périodes d'un an.

Le RKGE racine a été proposé à la suite d'essais de modifications du KGE racine pour inclure un terme sur les variations interannuelles du débit. Un test des performances obtenues avec GR4J est présenté en Annexe C. Son principe repose sur l'idée que la modification de GR4J vise explicitement l'amélioration de la robustesse du modèle, et qu'il est donc nécessaire de contraindre explicitement les paramètres pour qu'ils optimisent la capacité du modèle à représenter les variations interannuelles du débit. L'évaluation présentée en Annexe C montre que l'emploi du KGE racine ou du RKGE racine donne des résultats similaires, et que le RKGE racine peut être utilisé également avec les modèles simples comme GR4J en calage sur des chroniques longues. La comparaison du RKGE racine et du KGE racine n'a pas été réalisée en contrôle.

8.1.2.b. Méthode d'évaluation du modèle en contrôle

La deuxième étape de l'analyse s'appuie sur des expériences de calage-contrôle adaptées de celles conduites au Chapitre 4 pour l'implémentation du PMR. Pour chaque bassin, six périodes de calage ont été constituées à partir des six années les plus pluvieuses, des six années les moins pluvieuses, des six années les plus chaudes, des six années les plus froides, des six années les plus productives (au sens du rendement moyen des précipitations) et enfin des six années les moins productives. Pour chacune des expériences réalisées, les six années sélectionnées ne sont pas nécessairement consécutives. Les performances des modèles ont été systématiquement évaluées sur l'ensemble de la chronique disponible. L'évaluation a donc été donc réalisée de la même manière que lors de la première étape, à ceci près que les paramètres des modèles ont été optimisés dans des conditions climatiques spécifiques. Il convient de noter que les périodes de calage et la période de contrôle ne sont pas strictement indépendantes, les premières étant inclues dans la deuxième.

Le choix des précipitations et de la température pour la définition des périodes de calage permet d'inscrire ce travail dans la continuité des études de robustesse relevées dans la littérature scientifique. Celui du rendement moyen des précipitations répond à l'exigence de tester les modèles dans des conditions très défavorables, tel que cela a été démontré aux Chapitres 4 et 5.
Le choix de constituer des périodes de calage de six ans a été déterminé d'une part afin de limiter le chevauchement des périodes de calage et de la période de contrôle, et d'autre part pour éviter

que certains paramètres soient sous-contraints par une quantité insuffisante de données en calage.

Le calage des paramètres a été réalisé avec le KGE racine, dont nous avons démontré la pertinence au Chapitre 5 pour maximiser la robustesse des modèles hydrologiques.

8.1.2.c. Critères de performance utilisés

Outre le critère de calage, nous avons principalement utilisé comme critères d'évaluation le KGE sur les débits non transformés, qui est un bon critère d'évaluation pour donner un aperçu de la capacité du modèle à reproduire les principales caractéristiques statistiques du débit.
Nous avons également considéré le critère de Nash-Sutcliffe calculé sur les débits transformés avec la fonction inverse, la fonction logarithme, la racine carrée, la fonction identité (pas de transformation) et la fonction carrée. Ces transformations permettent de considérer la capacité des modèles sur des gammes de débit croissantes (Le Moine, 2008). Le Nash est préféré au KGE pour pouvoir utiliser le logarithme, ce qui n'est pas possible avec le KGE (Santos *et al.*, 2018b).
Enfin, le PMR du modèle calculé sur les trois gammes de débit, à savoir le PMR_B, le PMR_I et le PMR_H, a été employé pour évaluer les bénéfices du nouveau paramétrage sur la représentation des variations interannuelles du débit (voir Chapitre 7).

8.2. Évaluation du paramétrage symétrique du réservoir de production

8.2.1. Évaluation de GR4J_sym en calage

8.2.1.a. Validation de la résolution numérique du modèle continu

Nous avons vérifié dans un premier temps que pour une relation quadratique inchangée, les modèles GR4J « intégré » et GR4J « numérique » se comportaient de la même façon. Pour cela, nous avons étudié les distributions des paramètres libres des deux modèles sur l'échantillon de bassin, comme cela avait déjà été fait par Santos *et al.* (2018a) pour « GR4 state space ». Cette vérification s'avère en effet utile dans la mesure où des erreurs numériques pourraient exister et se propager dans le calcul des états successifs du réservoir de production, par exemple à cause d'une division insuffisante en sous-pas de temps. A cela s'ajoute le fait que le remplissage du réservoir (ou vidange par évaporation) et la percolation sont calculés de façon concomitante dans GR4J « numérique », alors qu'ils sont pris en compte séquentiellement dans GR4J « intégré ».

La Figure 8.2 montre que les paramètres obtenus en calage avec le RKGE racine des deux versions du modèle sont quasiment égaux sur la plupart des bassins. Sur une poignée de bassins néanmoins, les paramètres peuvent être assez différents. Ces différences sont particulièrement visibles en ce qui concerne X_4. Pour 24 bassins (sur 577), l'écart entre les deux valeurs de X_4 est supérieur à 0,5 jour. Dans la plupart des cas, les deux jeux de paramètres produisent des performances égales. Sur 9 de ces 24 bassins cependant, l'écart de valeur du KGE est supérieur à 0,01 à la faveur de GR4J numérique. L'exemple des trois bassins où les différences en KGE sont les plus grandes est montré dans le Tableau 8.1.

Les écarts de performance sont principalement portés par le terme de corrélation du KGE. Dans les trois cas, la valeur de X_4 obtenue avec GR4J intégré est plus grande que celle obtenue avec GR4J

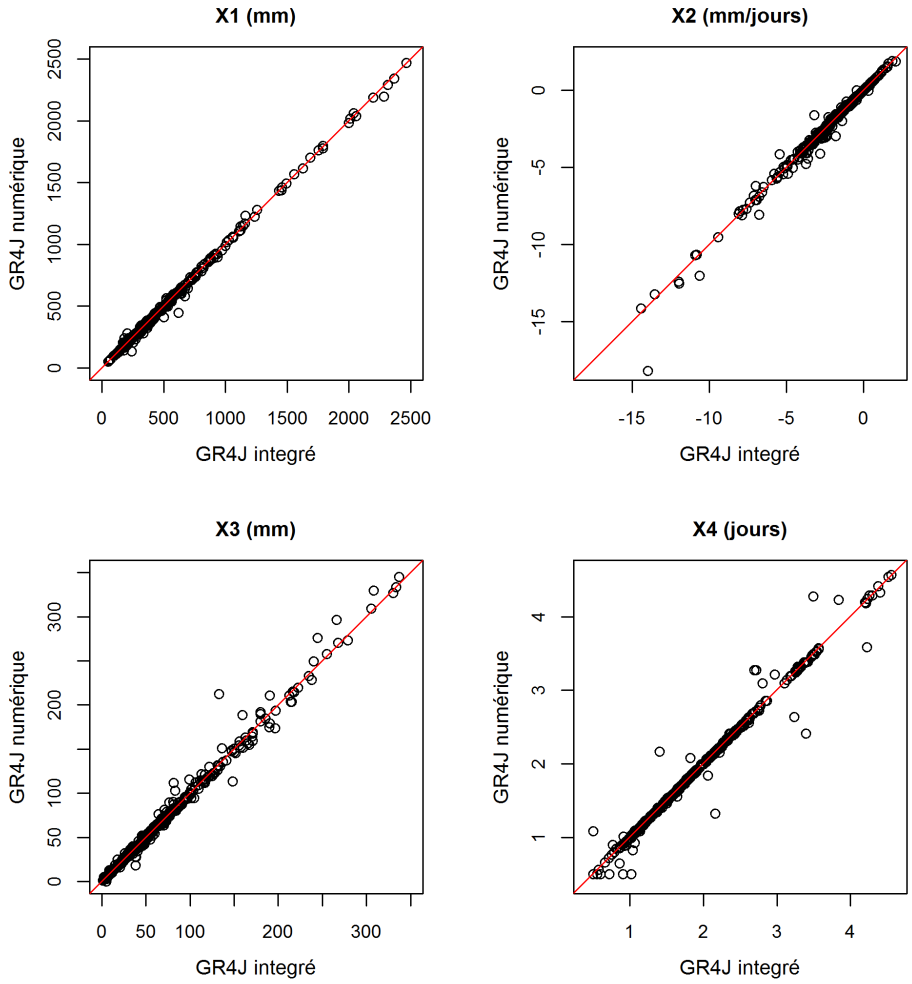


FIGURE 8.2 : Comparaison des paramètres de GR4J intégré (abscisse) et GR4J numérique (ordonnées) obtenus en calage sur toute la période des bassins de l'échantillon d'étude.

numérique. Une rapide analyse visuelle montre que bien que GR4J intégré et GR4J numérique aient globalement des comportements très similaires sur ces bassins, GR4J intégré est toujours en retard par rapport à GR4J numérique. Ce retard est détecté par le terme de corrélation du KGE, mais pas par celui du RKGE racine, comme en témoignent les très faibles écarts constatés pour ce critère. Cela suggère que cette situation est provoquée par une faible sensibilité du critère de calage RKGE racine au retard de GR4J intégré dans ces bassins.

Cette hypothèse a été confirmée en munissant GR4J numérique du jeu de paramètres obtenu pour GR4J intégré et en comparant les hydrogrammes obtenus. Les comportements des deux modèles sont presque identiques à quelques millièmes de millimètre d'eau près. Cette conclusion est également valable en munissant GR4J intégré du jeu de paramètres obtenu pour GR4J numérique. Ces différences très légères peuvent s'expliquer par des erreurs numériques liées à la discrétisation des équations en sous-pas de temps, ainsi que par le fait que dans GR4J numérique les équations de la fonction de production sont résolues de façon simultanées par le schéma implicite, tandis que la résolution est séquentielle dans GR4J intégré. Étant donné la très faible sensibilité du critère de calage aux différences entre les deux jeux de paramètres, il est possible que ces écarts suffisent pour justifier la convergence de l'algorithme de calage des paramètres de GR4J intégré vers un optimum secondaire.

TABLEAU 8.1 : Bassins où les écarts en KGE de GR4J « intégré » et GR4J « numérique » sont les plus grands. Le symbole Δ correspond à la différence entre les performances de GR4J_sym et celles de GR4J.

Bassin versant	ΔRKGE$_{[\sqrt{Q}]}$	ΔKGE	ΔX_4 (en jours)
Morass Creek River à Uplands (AU401215)	0,005	0,04	–0,87
Rose River à Matong North (AU403217)	0,012	0,07	–0,97
Le Ceou à Saint-Cybranet (P2484010)	0,004	0,03	–0,76

Malgré ces légers problèmes de calage sur la poignée de bassins évoqués, on peut supposer que dans la plupart des cas les deux versions de GR4J donnent des résultats presque identiques. Dans la suite, ce qu'on appellera GR4J est la version de GR4J résolue numériquement, mais dont les puissances sont inchangées par rapport à la version originale ($\alpha = 2$). Pour rappel, on appelle GR4J_sym le modèle à 5 paramètres dont le paramètre α est laissé libre en calage.

8.2.1.b. Comparaison vis-à-vis du critère de calage

Pour obtenir un aperçu des bénéfices apportés par le nouveau paramétrage libre du réservoir de production, la Figure 8.3 montre les différences observées en calage dans les valeurs du critère de calage.

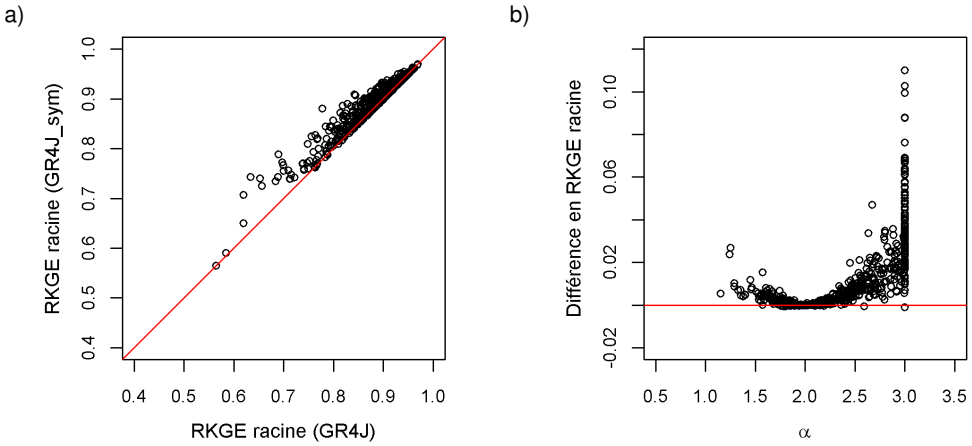


FIGURE 8.3 : a) Comparaison des valeurs de RKGE racine obtenus en calage sur les bassins français et australiens (mélangés) avec GR4J (abscisse) et GR4J_sym (ordonnées) en laissant le paramètre α libre. b) Écart entre RKGE racine de GR4J et GR4J_sym, en fonction de la valeur calée pour α.

Dans le cas général, il apparaît que le calage du cinquième paramètre n'apporte qu'une très légère hausse de performance au regard du RKGE racine, de l'ordre de quelques centièmes de points. Les calages conduisant le paramètre α vers des valeurs proches de 2 n'ont logiquement quasiment aucune influence sur les performances en calage. En revanche, certains bassins semblent bénéficier de plus de non-linéarité dans le fonctionnement du réservoir de production ($\alpha \simeq 3$). Beaucoup de calages aboutissent d'ailleurs à la valeur maximale autorisée pour α, suggérant qu'il aurait fallu définir un domaine plus large. On note que les bassins où l'augmentation du RKGE racine est la plus forte, autour de $\alpha \simeq 3$, sont situés en Australie (Figure 8.4).

8.2.1.c. Comparaison au regard de critères alternatifs

Afin de donner un aperçu général des simulations de GR4J_sym sur l'ensemble du régime hydrologique des bassins de l'échantillon, les performances de GR4J et GR4J_sym sont comparées au

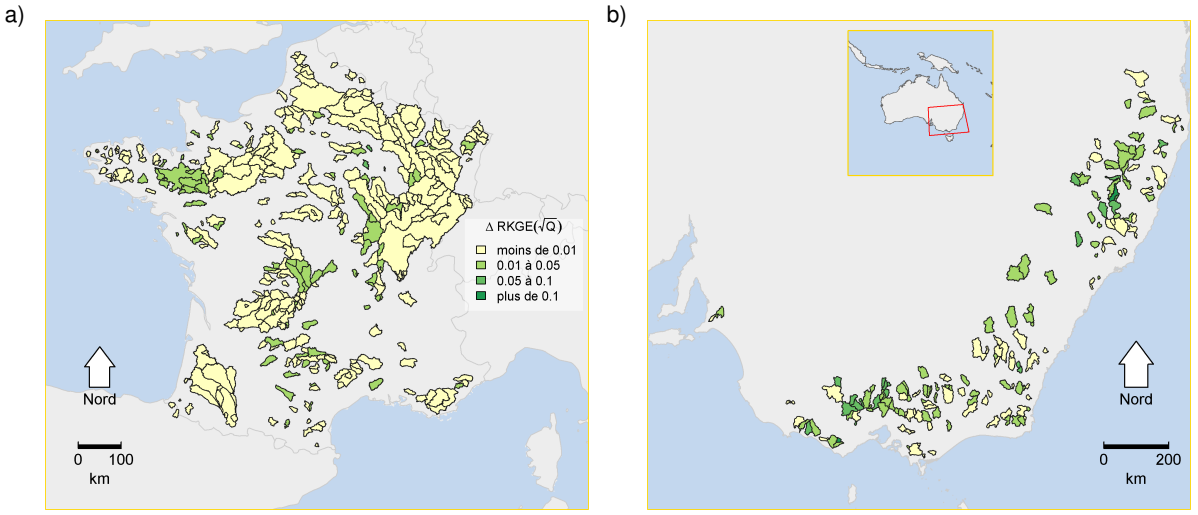


FIGURE 8.4 : Cartes des améliorations en termes de RKGE racine apportées en calage par le modèle GR4J_sym (5 paramètres) par rapport à GR4J (4 paramètres), sur les bassins français (a) et australiens (b).

regard de la déclinaison du critère de Nash-Sutcliffe sur les différentes transformations du débit en Figure 8.5.

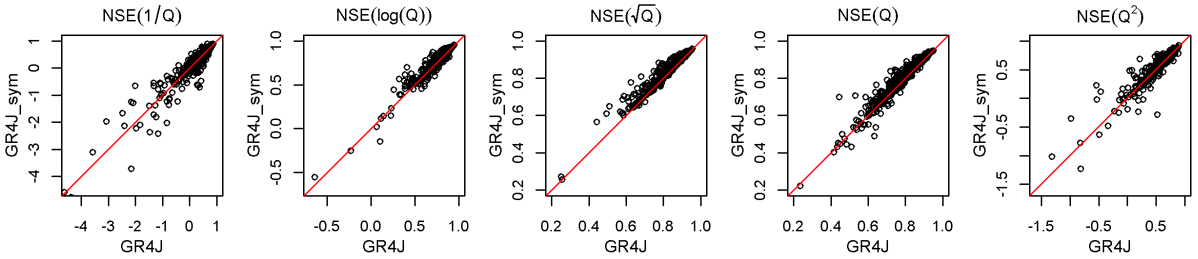


FIGURE 8.5 : Comparaison des valeurs de NSE calculé sur les débits transformés par cinq différentes transformations (fonctions inverse, logarithme, racine carrée, identité, et carrée) entre le modèle GR4J original et le modèle GR4J_sym avec paramètre α libre. Les résultats issus des deux sous-échantillons français et australien ont été mélangés.

La distribution des valeurs du critère de Nash-Sutcliffe des gammes intermédiaires (de NSE$[\ln Q]$ à NSE$[Q]$) confirme les bons résultats de GR4J_sym observés précédemment. En revanche, le modèle à 5 paramètres n'apporte pas d'amélioration générale des performances sur les Nash « extrêmes » (NSE$[1/Q]$ et NSE$[Q^2]$). Ce résultat suggère que moduler le rendement du réservoir de production affecte peu les erreurs de simulation des pics de crues et des étiages.
Pour vérifier l'apport de GR4J_sym sur la robustesse du modèle pour les gammes de débit extrêmes, les valeurs de PMR sur les gammes des basses eaux, des eaux intermédiaires et des hautes eaux sont comparées à GR4J en Figure 8.6.
Ces résultats montrent que la robustesse du modèle bénéficie du paramétrage sur les trois gammes de débits, en particulier les gammes des basses eaux et des eaux intermédiaires.
Pour mieux comprendre les raisons pour lesquelles le paramétrage libre des partitions des précipitations et de l'évaporation nettes du réservoir de production conduit à l'amélioration des performances du modèle, nous avons étudié spécifiquement le cas de la Canal Creek River à Leyburn (AU422338). Dans ce bassin, les écarts de performances entre GR4J_sym et GR4J sont les plus forts, tant au regard du KGE que celui des différents Nash. Ce bassin n'est bien sûr pas représentatif de ce qui se passe sur le reste de l'échantillon, mais il permet d'espérer voir des différences plus nettes entre

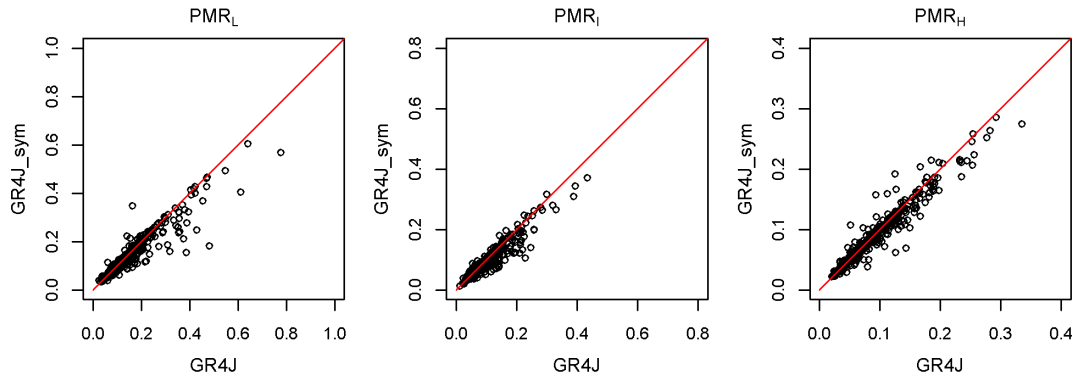


FIGURE 8.6 : Comparaison des valeurs du PMR calculé sur les trois gammes de débits (PMR_B : basses eaux, PMR_I : eaux intermédiaires, PMR_H : hautes eaux) entre le modèle GR4J original et le modèle GR4J_sym avec paramètre α libre. Les résultats issus des deux sous-échantillons français et australien ont été mélangés.

les deux modèles. Il illustre d'une certaine façon un cas « typique » de bassins où une petite modification de GR4J peut permettre de faire un progrès notable dans la description du comportement hydrologique.

8.2.1.d. Analyse de la Canal Creek River à Leyburn

Le bassin de la Canal Creek River, à Leyburn (AU422338) est situé dans l'État du Queensland, à l'est de l'Australie. Il fait 400 km^2 environ, et alimente la Condamine River (CSIRO, 2008)[1]. Il est situé dans une zone où les activités humaines sont faibles, comprenant de nombreuses réserves naturelles.
Sur le bassin, le KGE de GR4J_sym atteint 0,83 en calage, contre 0,72 pour GR4J. La nette augmentation du KGE entre GR4J et GR4J_sym est principalement portée par les sous-critères de corrélation et de variabilité du débit. Le paramètre α de GR4J_sym associé à ces performances est égal à 3, c'est-à-dire la valeur maximale autorisée a priori. Les débits observés sur le bassin sont comparés aux simulations de GR4J et GR4J_sym en Figure 8.7.

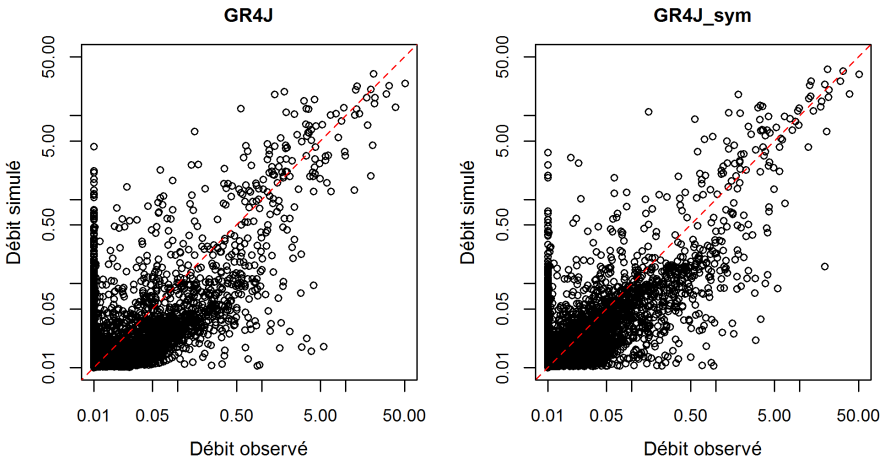


FIGURE 8.7 : Comparaison des débits journaliers simulés par GR4J et GR4J_sym aux débits journaliers observés sur le bassin de la Canal Creek River à Leyburn (AU422338).

[1] `http://www.clw.csiro.au/publications/waterforahealthycountry/mdbsy/pdf/condamine-report.pdf`

Bien que les écarts avec les observations soient assez similaires pour les deux modèles, les débits intermédiaires semblent mieux simulés par GR4J_sym. Pour mieux comprendre l'influence du paramétrage du réservoir de production sur le comportement de GR4J_sym, nous avons également comparé les flux simulés par les deux modèles (Figure 8.8).

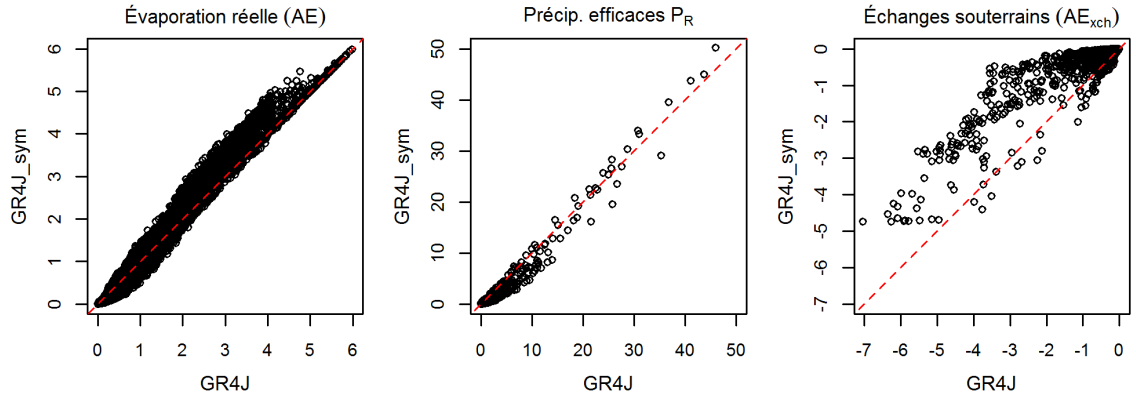


FIGURE 8.8 : Comparaison des flux journaliers des modèles GR4J (abscisse) et GR4J_sym (ordonnées) sur le bassin de la Canal Creek River à Leyburn (AU422338). L'évaporation réelle (notée AE) est calculée comme la somme des évaporations de l'interception et du réservoir de production. Les précipitations efficaces correspondent au flux d'eau P_R en sortie la fonction de production, alimentant dans les hydrogrammes unitaires. Les échanges souterrains (notés AE_xch) correspondent à la somme des composantes d'échange avec l'extérieur du système du réservoir de routage et du flot direct.

Alors que les deux modèles sont également non biaisés sur le bassin de la Canal Creek River, les termes du bilan des deux modèles se révèlent très différents. Le modèle GR4J_sym produit moins de pluie efficace à partir de son réservoir de production (30% de moins en moyenne que GR4J), et donc simule des fuites souterraines moins conséquentes (de 45% de moins en moyenne). L'évaporation réelle moyenne est également légèrement plus grande (5% en plus en moyenne).
Ce résultat n'est pas surprenant au regard des paramètres des deux modèles. À taux de remplissage donné, une valeur de α égale à 3 permet à la fois au modèle GR4J_sym de stocker plus de pluie dans le réservoir de production et d'évaporer à un taux plus proche du taux potentiel que le modèle GR4J. Un test avec l'évaporation potentielle de Morton, dont la moyenne est de 24% supérieure à celle d'Oudin sur le bassin, suggère que GR4J_sym compense peut-être une sous-estimation de la demande évaporatoire lorsque forcé avec le modèle d'Oudin. Ce constat confirme la plus grande indépendance de GR4J_sym vis-à-vis des échanges souterrains que GR4J pour l'ajustement de son bilan hydrique, conformément au but recherché à travers le paramétrage du réservoir de production du modèle. Sur ce bassin, cette indépendance améliore clairement la qualité des simulations en calage.

Nous avons donc constaté qu'un paramétrage libre du réservoir de production montrait un certain potentiel d'amélioration des performances de GR4J. Si le calage de la puissance du réservoir aboutit à des valeurs proches de la puissance originale du modèle dans une partie importante de l'échantillon, certains bassins bénéficient particulièrement d'un réservoir au fonctionnement encore plus non-linéaire. Il est possible que ces bassins soient notamment ceux où l'évaporation potentielle est sous-estimée. Il convient désormais de vérifier que l'ajout d'un paramètre ne pose pas plus de robustesse en dehors des conditions de calage.

8.2.2. Définition de l'intervalle de variations du paramètre α

8.2.2.a. Analyse des liens entre performances et paramètre α

Un premier aperçu des performances en contrôle de GR4J_sym est présenté par la Figure 8.9, qui montre la répartition géographique des gains en termes de KGE par rapport à GR4J.

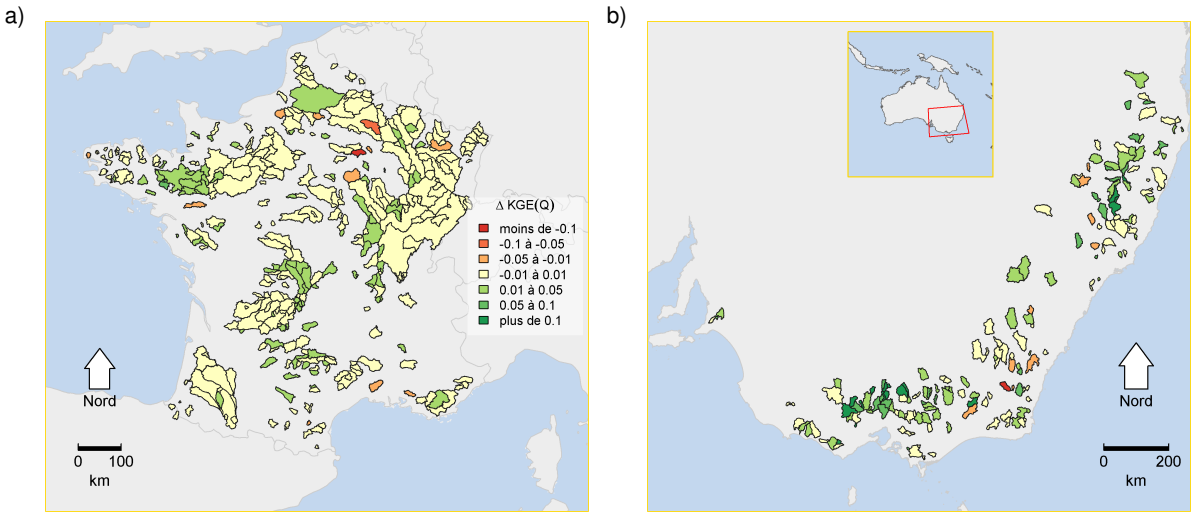


FIGURE 8.9 : Cartes des écarts moyens en termes de KGE sur les six expériences de DSST apportées en contrôle par le modèle GR4J_sym (5 paramètres) par rapport à GR4J (4 paramètres), sur les bassins français (a) et australiens (b).

Le gain de performance apporté par un paramétrage libre du réservoir de production n'est pas uniforme dans l'échantillon de bassins. En France, certains bassins bénéficient modérément du paramétrage. Ces bassins sont pour la plupart ceux où le modèle montrait déjà une amélioration du KGE en calage. Sur certains bassins, le modèle perd au contraire en robustesse. Sur ces bassins, il se trouve que le calage du paramètre α aboutit à des valeurs généralement inférieures à 2 (Figure 8.10). En Australie, GR4J_sym peut également être moins robuste que GR4J sur certains bassins, bien que sur une grande partie de l'échantillon ses performances soient sensiblement meilleures. Sur le bassin déjà étudié de la Canal Creek River, le modèle gagne par exemple 0,05 points de KGE en moyenne sur les expériences de DSST réalisées. Les gains dépassent ce seuil de 0,05 points sur une partie significative de l'échantillon australien (36 des 181 bassins).

La Figure 8.10 montre les écarts de KGE entre GR4J_sym et GR4J en fonction de la valeur calée du paramètre α. Conformément aux attentes, les écarts sont très faibles lorsque α est proche de 2. Dans la plupart des cas où α atteint des valeurs plus proches de 3, les performances de GR4J_sym sont meilleures que celles de GR4J. Les résultats sont en revanche moins univoques, lorsque α tend vers une valeur entre 1 et 2. Il convient de noter que les écarts de performance les plus marqués (dans les deux sens) sont obtenus en calage sur une sous-période plus sèche que la période totale utilisée pour le contrôle.

Si dans l'ensemble ces résultats sont corrects, un nombre assez conséquent de situations de calage-contrôle aboutit sur une détérioration moyenne des performances est d'au moins 0,01 point de KGE (19% des cas). L'analyse des valeurs du critère de calage montre pourtant que dans ces situations les performances en calage sont très généralement améliorées. Ces situations correspondent donc à des problèmes de sur-apprentissage des paramètres de GR4J_sym. Dans ces cas, les processus hydrologiques dimensionnés sont peut-être trop spécifiques à la période de calage et

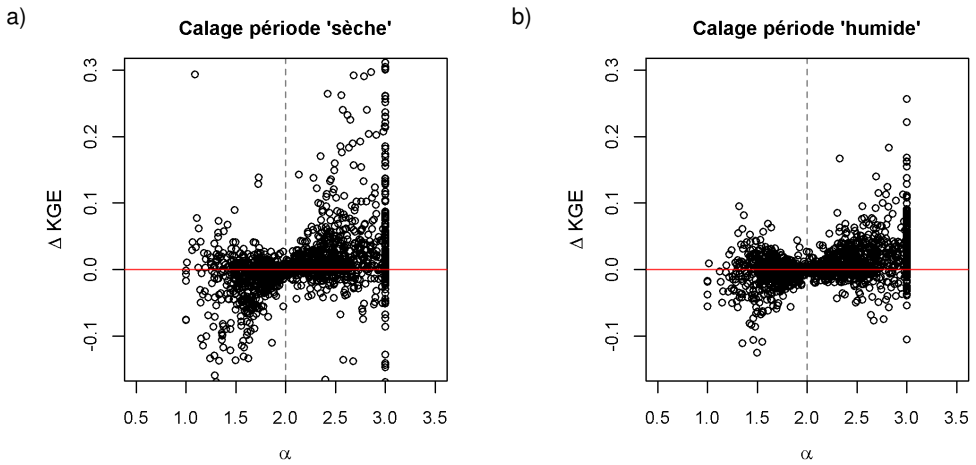


FIGURE 8.10 : a) Écart de KGE par bassin en contrôle entre GR4J_sym (5 paramètres) et GR4J (4 paramètres), en fonction de la valeur calée du paramètre α de GR4J_sym, en calant les modèles sur les trois sous-périodes « sèches » (période la moins pluvieuse, période la plus chaude, période la moins productive). b) Idem, en calant les modèles sur les trois sous-périodes « humides » (période la plus pluvieuse, période la moins chaude, période la plus productive).

ne représentent pas le cycle habituel du bassin versant. Le sur-apprentissage des paramètres d'un modèle révèle parfois un défaut structurel : le modèle est peut-être trop complexe ou au contraire trop simple, trop sensible aux erreurs locales dans les données, ou à la non-stationnarité des processus dominants (Andréassian *et al.*, 2012). Puisque dans notre cas, le modèle le plus complexe est également le moins robuste, le paramètre α est vraisemblablement inadéquatement contraint dans les situations où les performances sont détériorées. Ce phénomène semble lié à des valeurs du paramètre α inférieures à 2.

8.2.2.b. Analyse du comportement du modèle sur le bassin du Gapeau à Hyères

Pour analyser les manifestations d'un sur-apprentissage de GR4J_sym, on mène une étude approfondie du bassin français du Gapeau à Hyères (Y4624010), où les pertes moyennes de performance en contrôle s'élèvent à environ 0,05 points de KGE (seuls deux calages sur six donnent un résultat faiblement positif). Les pires pertes, qui concernent principalement les termes de variabilité et de bilan du KGE, sont obtenues en calage sur les périodes sèches non productives. Sur ce bassin, les calages de GR4J_sym amènent le paramètre α vers des valeurs entre 1,1 et 2,2, tant en calage sur les six sous-périodes ou sur la période complète.
Pour vérifier l'hypothèse d'un sur-apprentissage, la Figure 8.11 répertorie les moyennes et écarts-types de plusieurs flux simulés par les modèles GR4J et GR4J_sym, calculés pour chacun des 6 jeux de paramètres obtenus en DSST. À travers cette analyse, nous avons cherché à identifier des sensibilités différentes de ces variables au jeu de paramètres utilisé. Pour une même variable, un modèle montrant les valeurs les plus stables d'un jeu de paramètres à l'autre est moins sensible aux conditions de calage. A l'inverse, des flux modélisés moins stables sont susceptibles d'illustrer des problèmes de sur-apprentissage des paramètres.

Contrairement à ce qui était attendu, les distributions des caractéristiques de la pluie efficace (et des autres flux du réservoir de production non représentés) de GR4J et de GR4J_sym sont de largeur à peu près comparable. Si ce constat reste valable pour les échanges souterrains, on note néanmoins la présence d'un point aberrant parmi les réalisations de GR4J_sym, correspondant à une perte en eau forte et très saisonnière.
Malgré des distributions des flux internes d'écartement apparemment semblable, les caractéristiques des débits simulés par GR4J_sym suivent une distribution beaucoup plus large que celles

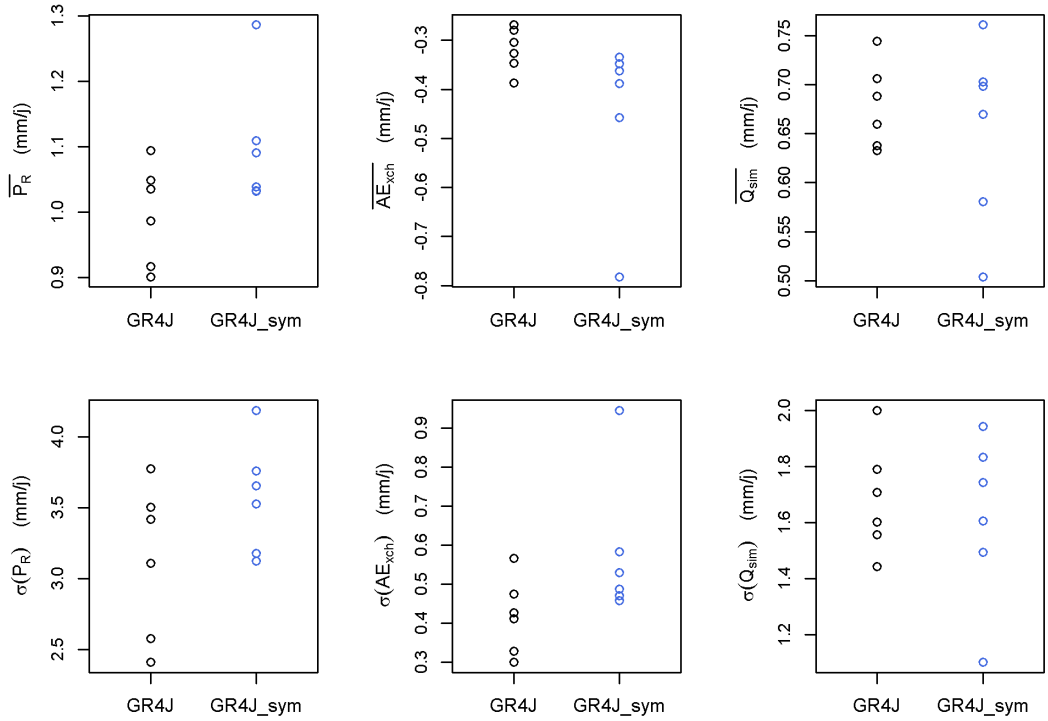


FIGURE 8.11 : Moyennes et écarts-types des précipitations efficaces (P_R), des échanges souterrains (AE_xch) et du débit (Q_sim) simulés par les modèles GR4J (noir) et GR4J_sym (bleu) sur le bassin du Gapeau à Hyères (4624010) en DSST (un point par expérience de calage-contrôle).

de GR4J. Cela peut s'expliquer en rappelant que l'addition de la pluie efficace moyenne et des échanges souterrains moyens est strictement égale au débit moyen modélisé. Un modèle presque parfait devrait ajuster ces deux termes pour aboutir à un débit simulé moyen fixe, égal au débit moyen observé. Ce n'est pas du tout le cas de GR4J_sym, ce qui signifie que le changement d'un des termes de l'équation (par exemple la pluie efficace moyenne) n'est pas compensé par celui de l'autre (par exemple les échanges souterrains moyens). En d'autres termes, la plus grande flexibilité de GR4J_sym dans l'ajustement du bilan n'est pas bien compensée par l'ajustement des fuites souterraines.
Dans ce cas précis, le paramètre libre supplémentaire de GR4J_sym est un défaut. Sa présence semble provoquer une sous-contrainte du dimensionnement des échanges souterrains, lors du calage des paramètres du modèle, particulièrement si celui-ci est fait sur une période sèche (au sens des précipitations, de la température et du rendement des précipitations).

8.2.2.c. Redéfinition de l'intervalle de variation du paramètre α

De manière générale, le comportement du modèle est plus sensible à la valeur de α lorsque celle-ci est entre 1 et 2, en comparaison de valeurs comprises entre 2 et 3. Les courbes des partitions des précipitations et de l'évaporation nettes du réservoir représentées en Figure 8.1 sont en effet moins serrées entre 1 et 2 qu'entre 2 et 3, pour un pas fixe d'évolution du paramètre α. Nous suggérons qu'en dessous de $\alpha = 2$, soit le comportement du réservoir de production est trop linéaire pour une majorité des bassins de l'échantillon, soit la valeur du paramètre est peu transférable d'une sous-période de calage à la période totale utilisée en contrôle. De fait, il est vraisemblable que les bassins aux comportements hydrologiques les plus linéaires soient déjà simulés convenablement par GR4J, dont le paramétrage quadratique de la fonction de production constituerait une sorte de limite inférieure à la non-linéarité de la réponse d'un bassin.

En réponse aux discussions précédentes, nous avons redéfini l'intervalle de variation α entre 2 et 8, plutôt qu'entre 1 et 3, puisque les situations où $\alpha \leq 2$ correspondent généralement à situations de détériorations de la transférabilité des paramètres. La limite supérieure initialement fixée à 3 a été étendue en constatant qu'elle était atteinte dans une part non négligeable des bassins de l'échantillon. La distribution des valeurs des écarts de KGE entre GR4J et GR4J_sym recalé avec le nouvel intervalle de variations du paramètre α est montrée en Figure 8.12.

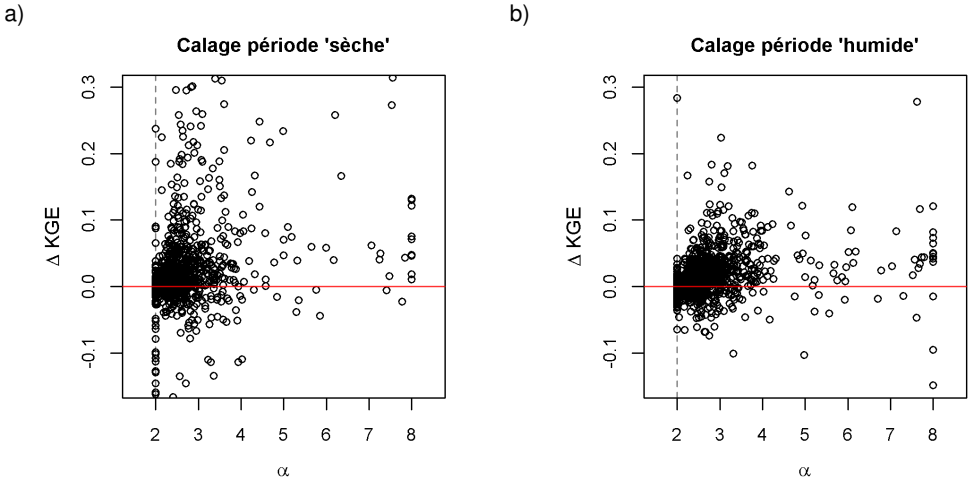


FIGURE 8.12 : a) Écart de KGE par bassin en contrôle entre GR4J_sym (5 paramètres) et GR4J (4 paramètres), en fonction de la valeur calée du paramètre α de GR4J_sym, en calant les modèles sur les trois sous-périodes « sèches » (période la moins pluvieuse, période la plus chaude, période la moins productive). b) Idem, en calant les modèles sur les trois sous-périodes « humides » (période la plus pluvieuse, période la moins chaude, période la plus productive).

Si le nombre de calages où α est supérieur à 3 reste limité, les performances sont néanmoins en moyenne améliorées par la redéfinition de l'intervalle de variation du paramètre. La restriction de la limite inférieure de l'intervalle de variation du paramètre permet en outre de réduire le nombre de situations où GR4J_sym est moins bon que GR4J d'au moins 0,01 point de KGE (9% des bassins contre 19% précédemment). Les calages sur les sous-périodes sèches restent ceux où les écarts de performance suivent une distribution plus large, tant en termes de pertes que de gains de performance.
Dans toute la suite du chapitre et pour toutes les modifications testées du modèle GR4J, l'intervalle de variation de α sera donc $[2\,;8]$.

8.2.3. Performances de GR4J_sym en contrôle

8.2.3.a. Comparaison au regard de critères alternatifs

De même que précédemment en calage, nous avons comparé plus en détail les écarts de performances entre GR4J_sym et GR4J au moyen de la déclinaison du critères de Nash-Sutcliffe avec les transformations de débit (Figure 8.13). Afin de vérifier que la valeur de α fixée à 2 dans GR4J est bien choisie, nous avons évalué en parallèle du modèle GR4J_sym à 5 paramètres libres des versions du modèle où α est fixé respectivement à 2,5 et 3.

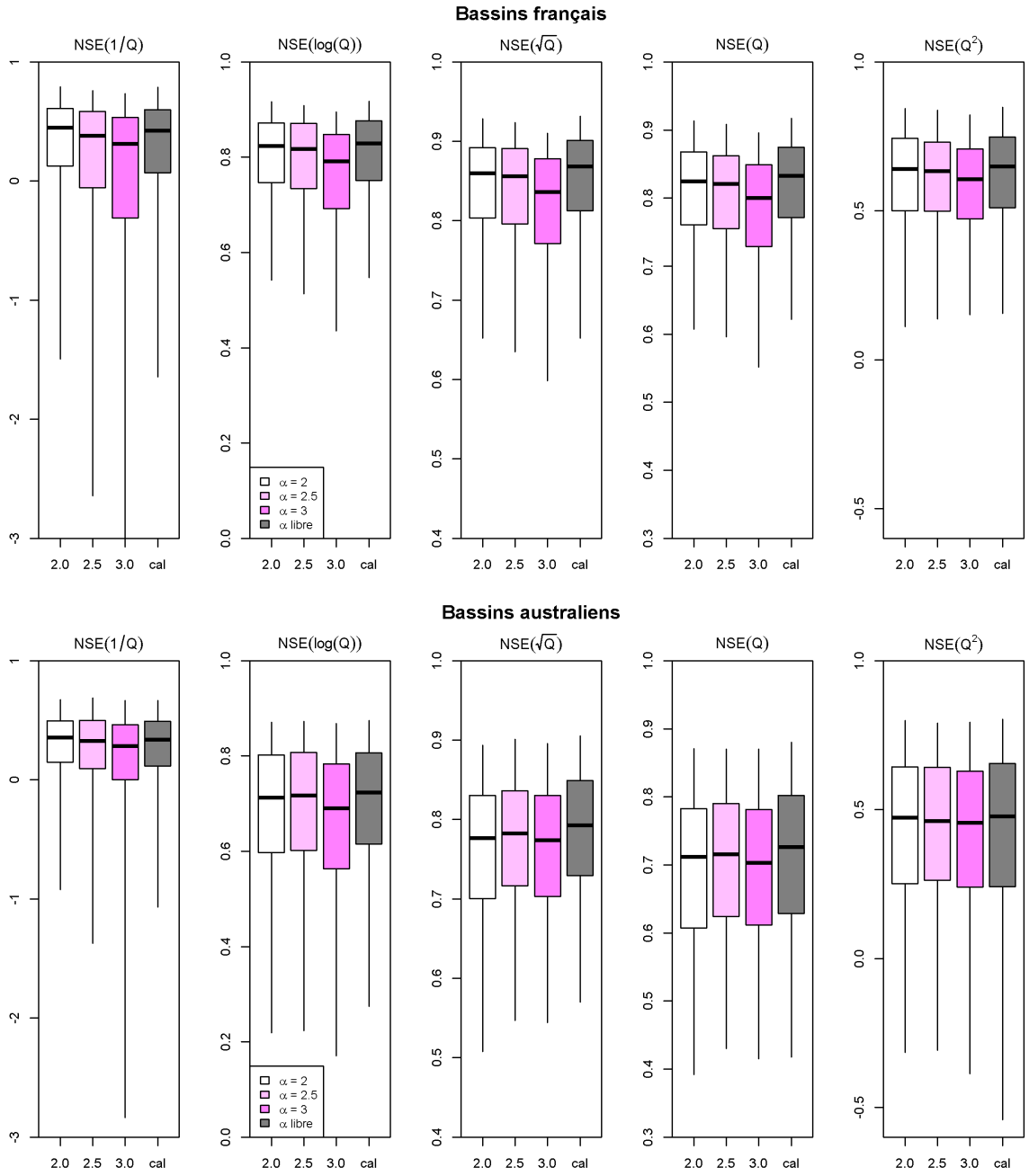


FIGURE 8.13 : Distribution des valeurs de NSE calculé sur les débits transformés par cinq différentes transformations (fonctions inverse, logarithme, racine carrée, identité, et carrée) sur les bassins du sous-échantillon français et du sous-échantillon australien obtenues pour le modèle GR4J original ($\alpha = 2$, en blanc), le modèle GR4J_sym avec paramètre α fixé ($\alpha = 2.5$ ou $\alpha = 3$, en rose), et le modèle GR4J_sym avec paramètre α libre (en gris).

Les résultats les moins satisfaisants sont obtenus avec la version de GR4J_sym à 4 paramètres libres la moins linéaire ($\alpha = 3$). Les performances de ce modèle sont particulièrement moins bonnes sur les bassins français, quelle que soit la gamme de débit considérée. En Australie, les simulations sont meilleures sur les gammes de débits intermédiaires lorsque le paramètre α est fixé à 2,5, confirmant ainsi les conclusions précédentes sur les bassins bénéficiant de plus de non-linéarité dans la fonction de production du modèle. De manière générale, les résultats les plus robustes avec un paramètre α fixe sont obtenus avec $\alpha = 2$, c'est-à-dire avec GR4J. Bien que sur certains bassins, un fonctionnement davantage non-linéaire du modèle améliore ses

performances, GR4J est malgré tout le modèle le plus générique, au sens où il permet d'atteindre des performances en moyenne meilleures sur un échantillon très large de bassins. Ce résultat permet d'apprécier la robustesse spatiale du paramétrage original de la fonction de production.

Les performances du modèle à 5 paramètres montrent toutefois qu'il est possible de complexifier légèrement le paramétrage de GR4J sans en altérer les performances en contrôle. Sur les bassins français, le paramétrage libre du réservoir de production améliore très modestement les performances générales du modèle sur les gammes intermédiaires (transformations racine carrée et identité). Les différences sont en revanche plus marquées en Australie. Le modèle GR4J_sym à 5 paramètres atteint des performances supérieures dans une majorité de bassins, sans perdre en robustesse sur les bassins les plus difficiles à modéliser. Les seules pertes générales sont observées dans les pires valeurs du NSE$[Q^2]$.

8.2.3.b. Caractéristiques des bassins où le paramétrage améliore les performances

Généralement, les bassins qui bénéficient le plus du paramétrage libre du réservoir de production sont ceux où α atteint des valeurs assez largement supérieures à 2, c'est-à-dire quand la réponse du modèle est très non-linéaire. Afin de déterminer les caractéristiques des bassins « gagnants », où GR4J_sym est plus performant que GR4J, nous avons effectué un test de corrélation de Spearman des écarts moyens en KGE avec les caractéristiques des bassins de l'échantillon.

Les résultats montrent que les bassins « gagnants » sont en premier lieu ceux dont les débits sont les plus variables, tant à l'échelle saisonnière (coefficient de variation du débit journalier, corrélation de 0,35) qu'à l'échelle interannuelle (coefficient de variation du débit annuel, 0,32). Ceux dont l'évaporation potentielle est la plus forte sont également mieux représentés par GR4J_sym (0,23). Ces résultats discriminent clairement les bassins de l'échantillon australien par rapport aux bassins français, moins arides et moins variables. Les raisons du succès de GR4J_sym sur les bassins australiens seront décryptées plus loin dans le chapitre. L'analyse de la corrélation de Spearman sur le seul échantillon français confirme néanmoins que la simulation des bassins dont le débit est le plus variable (échelles saisonnières et interannuelles) bénéficie le plus du paramétrage du réservoir de production. Les bassins avec les altitudes les plus élevés sont également mieux représentés par GR4J_sym.

En résumé, l'implémentation d'un paramétrage des partitions des précipitations nettes et de l'évaporation nette du réservoir de production conduit à une amélioration générale des performances du modèle. Le modèle GR4J_sym est plus robuste que GR4J pour la simulation d'une gamme assez large de débit. Ces résultats ont été obtenus après une redéfinition de l'ensemble des valeurs accessibles au paramètre α. Ce changement a posteriori souligne la difficulté à laquelle il est possible de faire face lors de la conception ou de la modification d'un modèle. Deux structures identiques peuvent être différemment robustes et sur-paramétrées, selon l'utilisation qui est faite des paramètres.

8.2.4. Discussion des résultats

L'ajout d'un nouveau processus ou de paramètres supplémentaires dans un modèle peut conduire à une modification du comportement des processus préexistants. L'objectif étant d'améliorer le modèle GR4J en le rendant plus robuste et plus performant dans la simulation de toutes les gammes de débit, nous avons cherché à éviter dans la mesure du possible que les aptitudes du modèle à simuler différentes signatures hydrologiques se compensent. Dans la suite, nous avons étudié l'effet de ce paramètre sur la distribution des autres paramètres et sur les flux du modèle pour mieux comprendre le fonctionnement du nouveau modèle.

8.2.4.a. Analyse du comportement de GR4J_sym

Les distributions des valeurs du paramètre α de modèle GR4J_sym obtenues en calage sur la période totale de chaque bassin est représentée en Figure 8.14 sur les deux sous-échantillons de bassins français et australien. Le positionnement des autres paramètres est représenté en Figure 8.15.

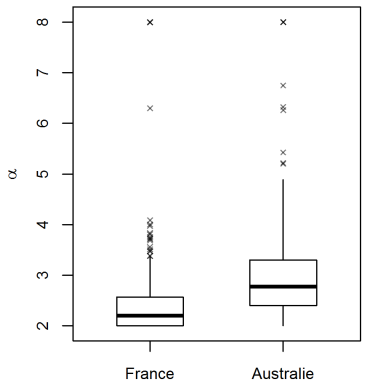


FIGURE 8.14 : Valeurs du paramètre α de GR4J_sym obtenues en calage sur les sous-échantillons de bassins australiens et français.

Les valeurs du paramètre α sont généralement supérieures dans les bassins australiens, où le modèle surperforme en comparaison de GR4J. Les variations du comportement du réservoir de production avec α donnent lieu à des compensations d'ampleurs variables de la part des autres paramètres. Pour certains des paramètres initiaux de GR4J, il n'est pas possible de distinguer des motifs clairs entre les valeurs des paramètres des deux modèles. Les tailles des réservoirs de production et de routage de GR4J et GR4J_sym sont par exemple généralement d'ordres de grandeur comparables. Bien qu'il ne pas soit pas évident d'interpréter ce résultat en ce qui concerne le routage, il aurait en revanche pu être attendu du réservoir de GR4J_sym qu'il soit plus profond que celui de GR4J, particulièrement dans les bassins australiens où son comportement est moins linéaire. En effet, plus grande est la valeur du paramètre α, plus grande est la vitesse de remplissage et de vidange du réservoir de production. De ce fait, la taille du réservoir doit augmenter pour que le modèle GR4J_sym retrouve le comportement de GR4J. La très légère augmentation de la taille du réservoir suggère donc que cette liberté supplémentaire octroyée au modèle n'est la plupart du temps pas compensée par ce moyen, au moins en ce qui concerne les bassins australiens. Il est en revanche possible qu'un autre mécanisme du modèle joue ce rôle de régulation.
Les variations de X_4 ne jouant presqu'aucun rôle sur le bilan du modèle, il est vraisemblable que ce rôle soit assumé par la fonction d'échanges souterrains du modèle. Le facteur d'échanges souterrains X_2 atteint en effet des valeurs supérieures pour GR4J_sym, c'est-à-dire moins négatives.

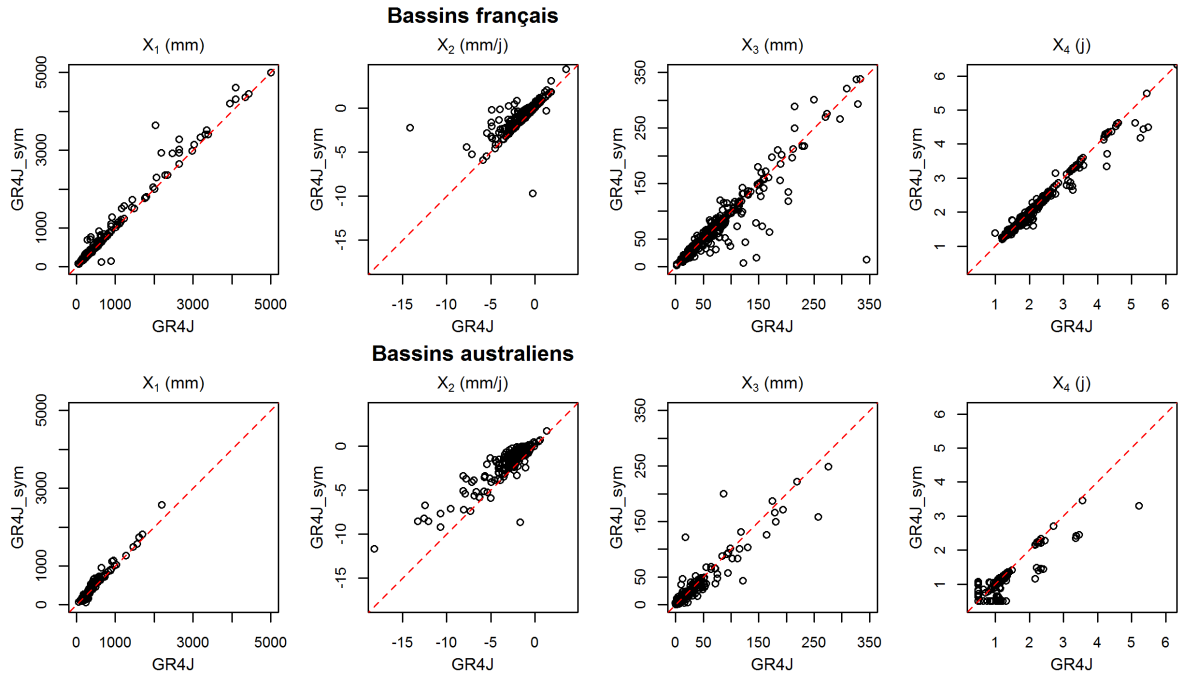


FIGURE 8.15 : Comparaison des paramètres de GR4J_sym et GR4J en calage sur les bassins des sous-échantillons français et australiens.

Pour un taux de remplissage du réservoir de routage égal, le modèle GR4J_sym simule donc des pertes souterraines moins importantes (Figure 8.16).

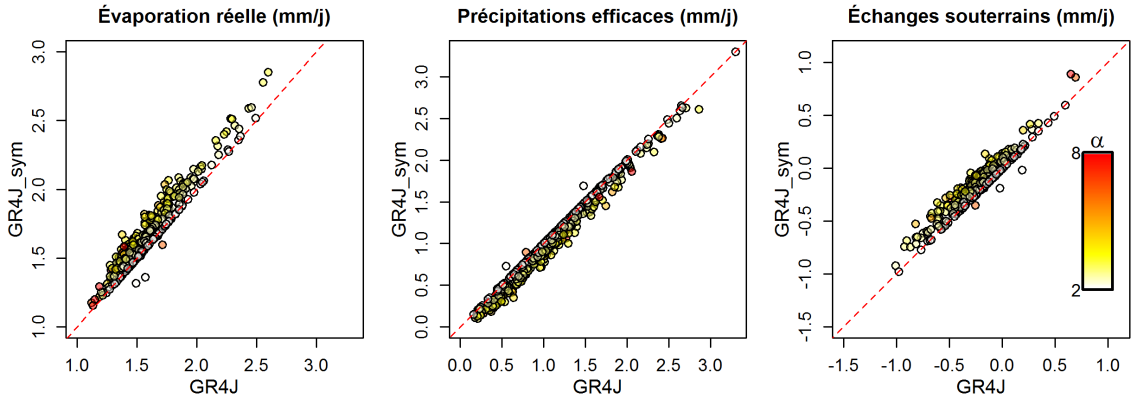


FIGURE 8.16 : Comparaison des moyennes des précipitations efficaces (P_R), des échanges souterrains (AE_xch) et du débit (Q_sim) simulés par les modèles GR4J (noir) et GR4J_sym sur les bassins de l'échantillon. Les bassins des sous-échantillons français et australiens ont été mélangés. La couleur de chaque point indique la valeur du paramètre α en calage sur le bassin (de blanc à rouge pour les α croissants).

Ces pertes souterraines réduites compensent la baisse du volume de pluie efficace en sortie de la fonction de production de GR4J_sym. La baisse du volume de pluie efficace P_R s'explique par le fait qu'une part plus importante de la pluie nette est captée par le réservoir de production sans que ce déficit soit compensé par une percolation accrue. Les rares cas où la pluie efficace moyenne de GR4J est supérieure correspondent à des situations d'égales performances des jeux de paramètres de GR4J et GR4J_sym, où $\alpha = 2$ et $X_{1,\mathrm{GR4J}} \gg X_{1,\mathrm{GR4J_sym}}$. Sur ces bassins, c'est l'apport très supérieur de la percolation du réservoir de GR4J aux pluies efficaces qui explique cette relation

d'ordre. Il apparaît par ailleurs que les différences de pluie efficace entre GR4J et GR4J_sym sont d'autant plus marquées que le paramètre α est grand. Cette observation confirme que ce paramètre supplémentaire rend la production du modèle plus flexible. La possibilité de GR4J_sym d'adapter sa production avec deux paramètres au lieu d'un seul permet de réguler plus fortement la pluie efficace moyenne.

8.2.4.b. Discussion des améliorations constatées sur les bassins « outliers »

Dans certains bassins français où GR4J modélisait des apports d'eau souterraine en complément des précipitations pour boucler le bilan hydrologique, le modèle GR4J_sym accentue ce phénomène (en haut à droite du troisième graphe de la Figure 8.17).
Les bassins concernés où le paramètre α est le plus grand (> 4) sont la Laines à Soulaines-Dhuys (H1333010), la Teyssonne à Changy (K1084010), le Saint-Nicolas à Rougemont-le-Château (U2305210), la Corrèze à Saint-Yrieix-le-Déjalat (P3322510) et la Siagne à Callian (Y5514040). Ils sont généralement caractérisés par une petite taille (< 50 km^2, exceptée la Siagne à Callian : 165 km^2) et une orientation marquée vers une direction cardinale. Mis à part le bassin de la Laines à Soulaines-Dhuys, ces bassins appartiennent aux 25% des bassins les plus pentus de l'échantillon français.
Le gain de performance apporté par GR4J_sym atteint généralement plusieurs centièmes de point de KGE, en particulier à travers une meilleure corrélation entre débits simulés et observés que GR4J. Plusieurs éléments permettent de mettre en doute l'existence d'un lien physique entre le processus ajouté à GR4J, à savoir la modulation de la non-linéarité de la réponse du réservoir de production aux forçages atmosphériques et le gain de performance observé.

Le paramètre α atteint sur ces bassins des valeurs élevées, proches de 8, ce qui a pour effet de diminuer en moyenne les précipitations efficaces. Le modèle compense donc par une augmentation des apports souterrains, prévoyant ainsi qu'une part significative de l'eau écoulée provienne de l'extérieur des frontières topographiques du bassin, entre 30% et 50% selon les bassins (70% pour la Laines à Soulaines-Dhuys). Il est possible que l'aire contributive au débit du bassin dépasse les frontières topographiques de ces bassins versants. Cette hypothèse est créditée par les positions particulières de ces bassins dans le diagramme de Budyko (Figure 8.17).

Le rendement moyen des précipitations particulièrement élevé de ces bassins compte tenu de leurs caractéristiques climatiques nous a conduit à vérifier pour chacun d'entre eux les données de précipitations par la comparaison des cumuls des entrées SAFRAN aux données de deux pluviomètres locaux fournies par Météo-France. Cette analyse, présentée en Annexe D, tend à montrer que les données de précipitations dans ces bassins sont assez fiables. Seules les données de la Siagne à Callian semblent être légèrement douteuses comparées aux mesures du pluviomètre le plus éloigné du centre du bassin. Nous avons considéré que ce constat n'était pas suffisant pour écarter le bassin versant de l'échantillon, d'autant plus que le réseau karstique du bassin constitue une explication physique crédible aux écoulements conséquents à l'exutoire du bassin (Durozoy, 1979). Il convient également de noter que le bassin de la Laines à Soulaines-Dhuys est marqué par la présence de craie dans la région (Lansiart, 1979), substrat poreux particulièrement aquifère qui peuvent influencer les dynamiques saisonnières et interannuelles du débit.

En définitive, les caractéristiques hydrogéologiques de certains des cinq bassins considérés et la bonne qualité apparente des données de précipitations semblent suffisantes pour justifier la position de ces bassins dans le graphe de Budyko. De ce fait, les gains en eaux souterraines simulés par GR4J et GR4J_sym semblent crédibles. Toutefois, l'augmentation de la contribution des échanges souterrains à l'écoulement dans GR4J_sym par rapport à GR4J reste discutable, dans la mesure où elle est accompagnée d'une diminution de la contribution des précipitations alors que ces bassins

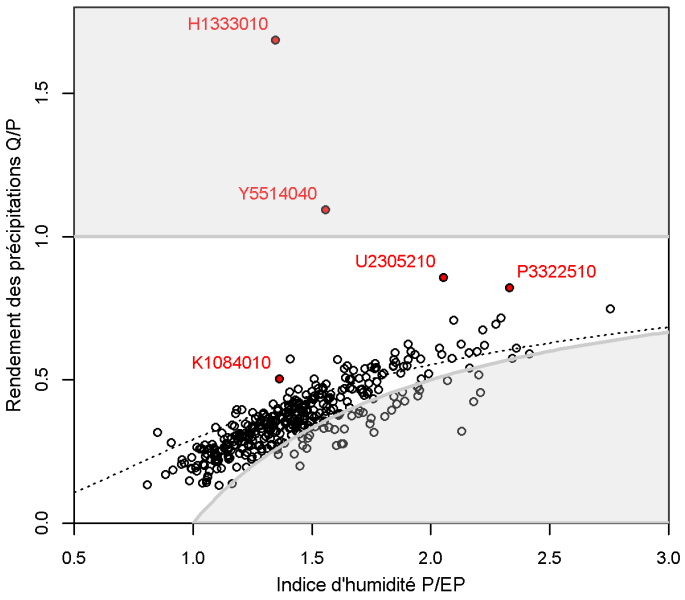


FIGURE 8.17 : Position des bassins de l'échantillon français dans le diagramme de Budyko. Les positions des bassins ont été calculées à partir des données observées sur la période 1960–2018. Les points rouges correspondent aux bassins où GR4J_sym simule des apports d'eau souterrains depuis l'extérieur des bassins plus importants que ceux modélisés par rapport GR4J, et où le paramètre α est supérieur à 4. La ligne pointillée noire correspond au modèle de Turc-Mezentsev. Les zones grisées correspondent aux situations où les précipitations moyennes sont inférieures aux débits observés moyens ($Q/P > 1$), ou aux situations où l'évaporation potentielle est plus inférieure à la différence entre précipitations et débits observés moyens ($E \leq P - Q$).

font déjà partie des bassins de l'échantillon français où les précipitations sont faibles par rapport au débit écoulé. Ces considérations amènent à questionner la signification physique de l'amélioration des performances apportées par GR4J_sym dans ces bassins. Il est possible que pour certains d'entre eux, le modèle apporte de meilleures réponses pour de mauvaises raisons.

8.2.4.c. Compensation des erreurs dans les données d'évaporation potentielle

Les analyses précédentes ont démontré d'une part que le modèle GR4J_sym donnait une plus grande marge de manœuvre à la fonction de production pour moduler son bilan, d'autre part que les bassins qui bénéficiaient le plus de ce degré de liberté supplémentaire étaient les bassins les plus variables et les plus arides. L'étude du bassin de la Canal Creek River a montré que le degré de liberté supplémentaire octroyé à GR4J_sym lui permettait de moduler le rôle de l'évaporation dans le bilan hydrologique, en palliant d'éventuelles erreurs du modèle d'évaporation potentielle. Il convient donc de questionner plus largement le rôle du paramètre α dans la compensation des données, notamment en entrée de modèle, et la responsabilité de ce phénomène dans l'amélioration des performances observée.
Pour tester l'hypothèse selon laquelle les meilleures performances de GR4J_sym sur le sous-échantillon australien sont liées par sa capacité à compenser les erreurs d'évaporation potentielle, nous avons comparé les écarts entre GR4J et GR4J_sym associés à l'évaporation potentielle de Morton. Pour rappel, la formule de Morton est plus adaptée que la formule d'Oudin en Australie, qui tend à sous-estimer la demande évaporatoire, et est associée à de meilleures performances des modèles hydrologiques testés (cf. Chapitre 6). Les écarts de performances entre GR4J et GR4J_sym au regard du KGE sont représentés en Figure 8.18.

Le modèle GR4J_sym forcé par l'évaporation potentielle de Morton apporte un gain de performance par rapport à GR4J sensiblement réduit comparé aux résultats obtenus avec l'évaporation

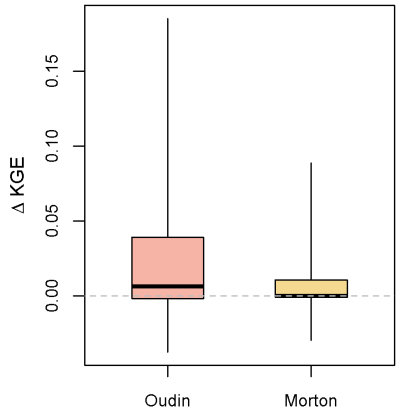


FIGURE 8.18 : Distribution des écarts de KGE obtenus en DSST entre GR4J_sym et GR4J sur le sous-échantillon de bassins australiens, forcés par l'évaporation potentielle d'Oudin ou par l'évaporation potentielle de Morton.

potentielle d'Oudin. En moyenne, les performances de GR4J_sym sont d'ailleurs égales avec les deux modèles d'évaporation (résultat non illustré). Ce constat indique donc que le paramétrage du réservoir de production agit effectivement en compensation des erreurs dans les données d'évaporation potentielle.
D'une certaine façon, les performances de GR4J_sym sont insensibles au modèle d'évaporation choisi, ce qui était déjà presque le cas de GR4J. Les performances très similaires de GR4J_sym avec les deux modèles d'évaporation suggèrent ainsi que le modèle est plus flexible que GR4J, et a priori plus facilement adaptable à des contextes de modélisation différents. Dans la mesure où GR4J_sym repose davantage sur l'évaporation réelle pour fermer son bilan hydrique, plutôt que sur les échanges souterrains, la flexibilité accrue du modèle se traduit à la fois par des meilleures performances et un meilleur réalisme physique que GR4J. Nous suggérons donc que cette flexibilité est ici un avantage.

Les développements présentés jusqu'à ce point montrent que la possibilité laissée à GR4J_sym de moduler la linéarité des équations liées à son réservoir de production est à la fois bénéfique en termes de performance et de réalisme. Toutefois, ils ne permettent pas de comprendre les influences respectives de la modulation de la partition des précipitations nettes et de la partition de l'évaporation nette.
Afin de détailler ces rôles, nous avons proposé dans la suite un modèle asymétrique, conçu de sorte que le paramètre α influence uniquement soit la partition de l'évaporation nette, soit celle des précipitations.

8.3. Évaluation du paramétrage asymétrique du réservoir de production

Dans cette section, nous avons introduit deux versions adaptées de GR4J_sym, nommées GR4J_asymP et GR4J_asymE. Dans GR4J_asymP, α est le paramètre libre de la partition des précipitations nettes dans le réservoir, tandis que le paramètre de l'évaporation nette est fixé à 2 comme dans GR4J. Dans GR4J_asymE, les rôles sont inversés : α est le paramètre libre de la

partition de l'évaporation nette, et le paramètre de la partition des précipitations nettes est fixé à 2. Les résultats montrés dans la suite ont été obtenus en suivant une méthode d'évaluation similaire à celle de GR4J_sym, c'est-à-dire en calage sur toute la période, puis en calage-contrôle avec six expériences de DSST.

8.3.1. Performances de GR4J_asym

8.3.1.a. Comparaison en calage

Comparés à GR4J, les deux modèles asymétriques atteignent de meilleures performances en calage, au regard du critère RKGE racine (Figure 8.19). Les deux modèles améliorent dans l'ensemble légèrement les performances en calage. Dans quelques bassins, GR4J_asymP atteint des performances sensiblement supérieures à celles de GR4J.

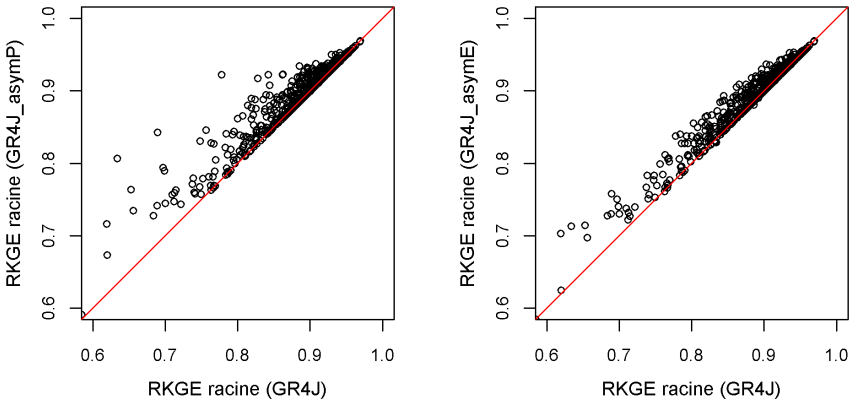


FIGURE 8.19 : Comparaison des valeurs de RKGE racine obtenues en calage sur l'échantillon de bassins avec GR4J (abscisse) et GR4J_asymP ou GR4J_asymE (ordonnées) en laissant le paramètre α libre. Les résultats obtenus sur les sous-échantillons de bassins français et australien ont été mélangés.

La comparaison des performances des modèles asymétriques entre elles confirment la similarité des performances de GR4J_asymP sur GR4J_asymE, excepté sur une poignée de bassins (Figure 8.20). Cette apparente similarité, masque des distributions du paramètre α fortement hétérogènes d'un modèle à l'autre. Dans le modèle GR4J_asymE, le paramètre parcourt amplement la gamme autorisée des valeurs de α : il semble que la borne supérieure fixée à 8 ne soit pas suffisante dans un nombre important de cas.
Ces résultats indiquent vraisemblablement des sensibilités différentes du comportement du modèle au paramétrage des partitions des précipitations et de l'évaporation nettes.

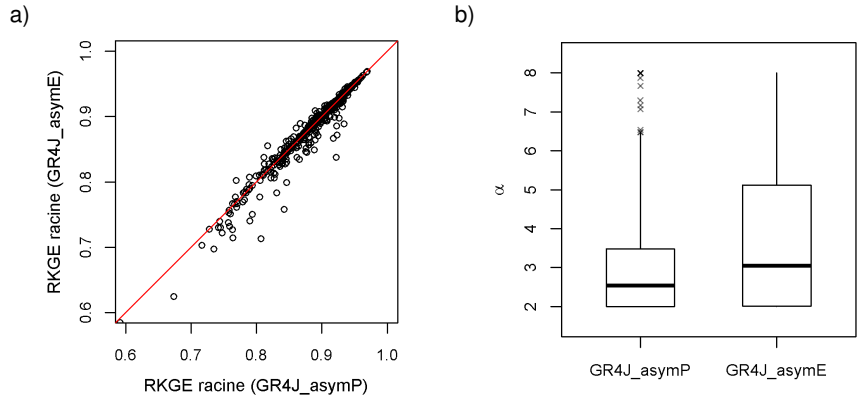


FIGURE 8.20 : a) Comparaison des valeurs de RKGE racine obtenues en calage sur l'échantillon de bassins avec GR4J_asymP (abscisse) et GR4J_asymE (ordonnées) en laissant le paramètre α libre. b) Distributions des valeurs du paramètre α obtenues en calage sur l'échantillon de bassins pour les modèles GR4J_asymP et GR4J_asymE. Les résultats obtenus sur les sous-échantillons de bassins français et australien ont été mélangés.

8.3.1.b. Comparaison en contrôle

En contrôle, les performances en KGE des deux modèles sont généralement équivalentes voire meilleures que celles de GR4J (Figure 8.21). De même qu'en calage, GR4J_asymP et GR4J_asymE ne se démarquent pas clairement l'un de l'autre. La comparaison des valeurs du KGE atteintes par les deux modèles ne montre pas d'avantage clair pour l'un ou l'autre des modèles asymétriques (Figure 8.22). Le paramètre α suit des distributions similaires à celles analysées précédemment en calage sur la période totale.

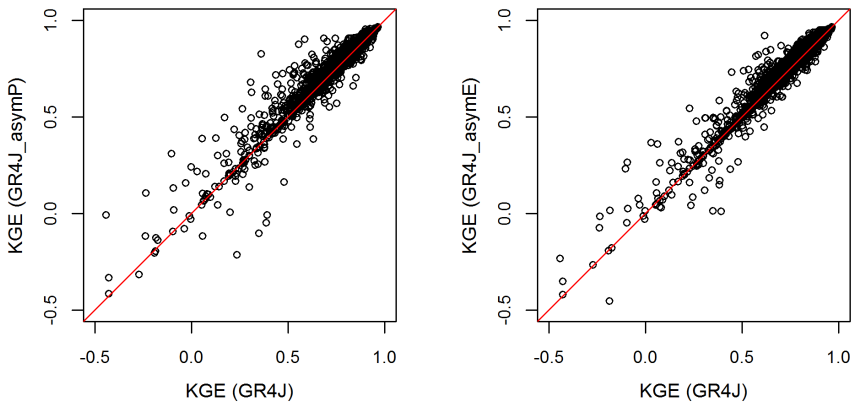


FIGURE 8.21 : Comparaison des valeurs de KGE obtenues en DSST sur l'échantillon de bassins avec GR4J (abscisse) et GR4J_asymP ou GR4J_asymE (ordonnées) en laissant le paramètre α libre (six points par bassin correspondant aux six expériences de DSST). Les résultats obtenus sur les sous-échantillons de bassins français et australien ont été mélangés.

Pour quantifier les différences de sensibilité des modèles à la valeur de α, nous avons évalué dans la suite les performances en contrôle des deux modèles asymétriques non seulement avec un paramètre α libre mais également en fixant le paramètre à des valeurs déterminées. De même que précédemment pour le modèle GR4J_sym, α a donc été soit laissé libre en calage, soit fixé à 2,5 et à 3. La Figure 8.23 présente les performances respectives de ces réalisations au regard des

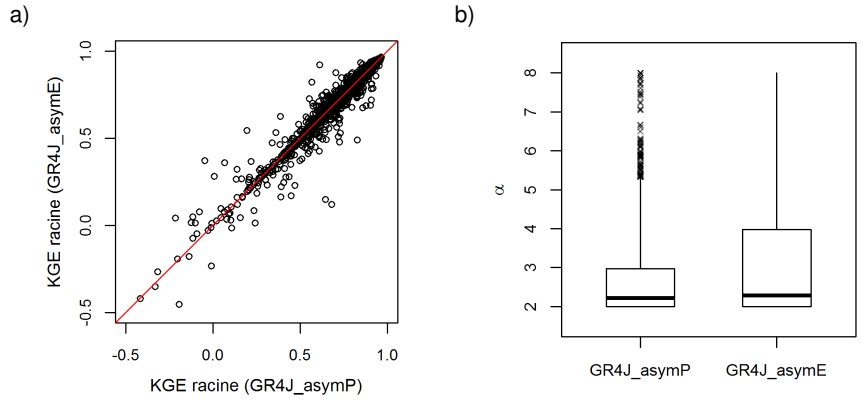


FIGURE 8.22 : a) Comparaison des valeurs de KGE obtenus en DSST sur l'échantillon de bassins avec GR4J_asymP (abscisse) et GR4J_asymE (ordonnées) en laissant le paramètre α libre. b) Distributions des valeurs du paramètre α obtenues en calage sur l'échantillon de bassins pour les modèles GR4J_asymP et GR4J_asymE (six points par bassin correspondant aux six expériences de DSST). Les résultats obtenus sur les sous-échantillons de bassins français et australiens ont été mélangés.

critères de Nash calculés les débits transformés.

De même que pour le modèle symétrique, les meilleures performances des deux modèles asymétriques sont obtenues en contrôle en laissant α libre. Pour les deux modèles asymétriques, ces performances sont généralement meilleures que celles du modèle GR4J original. Il convient de remarquer que les différences entre GR4J et les modèles asymétriques sont très faibles lorsque α est fixé à 2,5 ou 3, contrairement au cas précédent de GR4J_sym, dont les performances moyennes chutaient plus nettement lorsque α était fixé à ces mêmes valeurs. Nous suggérons donc que les effets additionnés d'une augmentation de la non-linéarité des équations de partition des précipitations nettes et de l'évaporation nette dans le modèle GR4J_sym rendent le modèle beaucoup plus sensible à la valeur du paramètre α.
Les performances des modèles testés ont été résumées dans le Tableau 8.2. Dans l'ensemble, le paramétrage de la réponse du réservoir de production améliore très légèrement les performances tant en calage qu'en contrôle.

TABLEAU 8.2 : Tableau récapitulatif des performances moyennes en DSST de GR4J, de GR4J_sym, de GR4J_asymP et de GR4J_asymE. Les performances ont été présentées de sorte que les meilleurs modèles ont des performances les plus proches de zéro.

Modèle	$1-$KGE	$1-$NSE$[\log Q]$	$1-$NSE$[\sqrt{Q}]$	$1-$NSE$[Q]$	PMR$_B$	PMR$_I$	PMR$_H$
GR4J	0,222	0,253	0,194	0,240	0,150	0,103	0,087
GR4J_sym	0,208	0,246	0,180	0,230	0,144	0,097	0,085
GR4J_asymP	0,212	0,248	0,182	0,232	0,146	0,099	0,085
GR4J_asymE	0,212	0,246	0,183	0,230	0,148	0,099	0,085

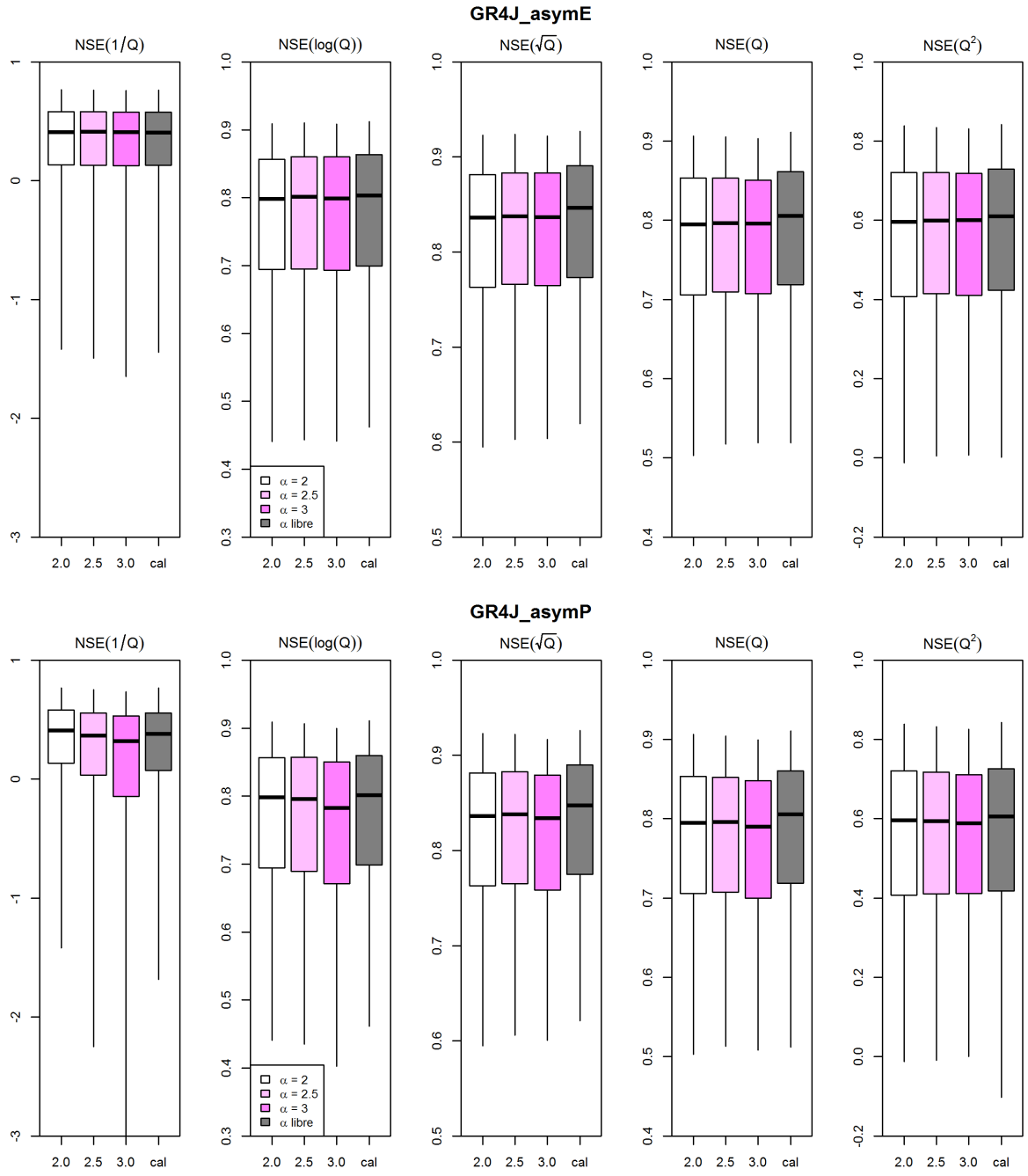


FIGURE 8.23 : Distribution des valeurs de NSE calculé sur les débits transformés par cinq différentes transformations (fonctions inverse, logarithme, racine carrée, identité, et carrée) sur les bassins de l'échantillon obtenues pour le modèle GR4J original ($\alpha = 2$, en blanc), le modèle GR4J_asymP (respectivement GR4J_asymE) avec paramètre α fixé ($\alpha = 2, 5$ ou $\alpha = 3$, en rose) et le modèle GR4J_asymP (respectivement GR4J_asymE) avec paramètre α libre (en gris). Les résultats obtenus sur les sous-échantillons de bassins français et australiens ont été mélangés.

8.3.2. Analyse du comportement du modèle

Les distributions des valeurs de Nash-Sutcliffe permettent donc de confirmer l'hypothèse faite sur les influences inégales des deux paramétrages du réservoir de production. Cependant, si le seul paramétrage libre de la partition des précipitations semble améliorer les performances du modèle à peu près autant que le seul paramétrage libre de la partition de l'évaporation potentielle, les mécanismes mis en œuvre sont différents. De fait, les similarités observées dans les performances sont susceptibles de masquer des différences de comportement. Par exemple, dans le cas où $\alpha = 8$, la quasi-totalité de la pluie nette est captée par le réservoir de production de GR4J_asymP, tandis que l'évaporation s'opère à taux quasi potentiel dans GR4J_asymE. En considérant la taille du réservoir fixée, ces deux comportements participent certes à la diminution de la pluie efficace moyenne, mais de façon différente. Dans le premier cas, une plus grande partie de la pluie nette est simplement captée par le réservoir, dans le second la vidange du réservoir sous l'effet de l'évaporation tend à augmenter sa propension à capter la pluie nette.

8.3.2.a. Comportement des modèles sur le bassin de la Teyssonne à Changy

Pour mieux comprendre les effets des deux paramétrages, nous avons étudié l'état du modèle sur un bassin versant où le calage mène à des valeurs de α proches de 8 dans les deux cas, pour des performances proches. Ainsi, les comportements des deux modèles ont pu être comparés, comme deux chemins menant au même endroit.
Nous avons choisi pour cet exemple le bassin de la Teyssonne à Changy (K1084010), dans le département de la Loire. Ce bassin est l'un de ceux où GR4J_sym obtient les meilleures performances par rapport à GR4J. Il est assez petit (22 km^2), assez bas en altitude et non karstique. L'analyse développée en Section 8.2.4.b a montré qu'il était possible que le degré de liberté supplémentaire permette de compenser des erreurs dans les données d'entrée et le manque de processus de fine échelle, et donc fournisse de bonnes réponses pour de mauvaises raisons. Il constitue cependant un bon exemple pour comprendre les différences entre les deux modèles asymétriques.

Les deux modèles asymétriques GR4J_asymP et GR4J_asymE atteignent respectivement un KGE de 0,82 et 0,83 (0,80 pour GR4J). Les courbes de régime simulées par les deux modèles sont représentées en Figure 8.24.

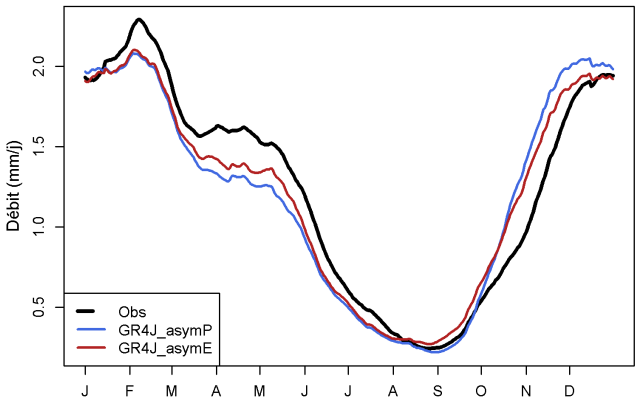


FIGURE 8.24 : Courbes de régime du bassin versant de la Teyssonne à Changy (K1084010) observée (noir) et simulées par les modèles GR4J_asymP (bleu) et GR4J_asymE (rouge).

Alors que leurs performances en termes de KGE sont très proches, quelques aspects des

simulations produites par les deux modèles diffèrent. La récession des écoulements simulée par GR4J_asymP est en effet plus douce que celle de GR4J_asymE, tandis que la reprise des écoulement est plus abrupte. Ces résultats sont cohérents avec les comportements théoriques de la production des deux modèles. Le réservoir de GR4J_asymP est paramétré pour se remplir très vite par captation des pluies nettes. Le réservoir de GR4J_asymE est au contraire paramétré pour se vider très vite par évaporation.
Poir détailler ces comportements, nous avons également comparé les flux des modèles sur le bassin (Figure 8.25).

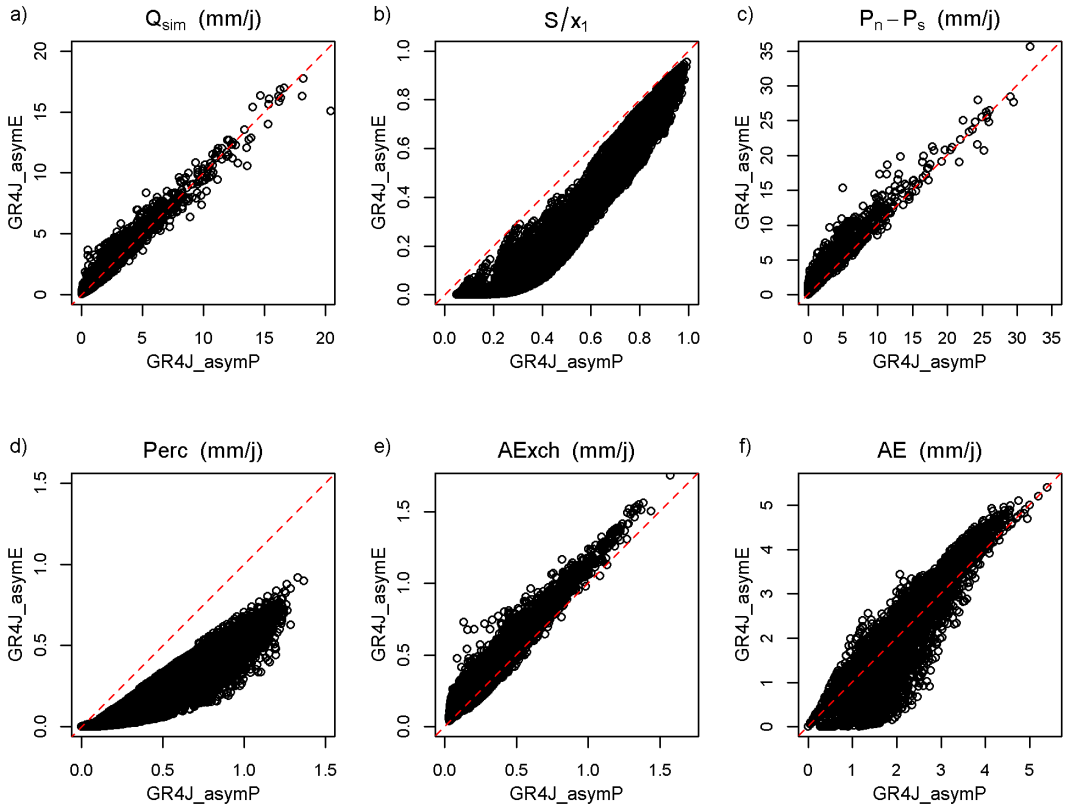


FIGURE 8.25 : Comparaison des flux et états internes de GR4J_asymP (abscisses) et GR4J_asymE (ordonnées) sur le bassin versant de la Teyssonne à Changy (K1084010). Les flux et états internes comparés sont le débit simulé (Q_sim), le taux de remplissage du réservoir de production (S/X_1), les précipitations non captés par le réservoir de production ($P_n - P_s$), la percolation du réservoir de production ($Perc$), les échanges souterrains totaux depuis le routage des modèles (AE_{xch}), et l'évaporation réelle totale (AE).

Le réservoir de production du modèle GR4J_asymP est systématiquement plus rempli dans le modèle (sous-figure b). Le réservoir de GR4J_asymP étant également plus grand (146 mm contre 122 mm), la quantité d'eau stockée par le modèle est donc nettement plus importante que celle stockée par GR4J_asymE. Cette différence de remplissage s'explique par la part plus large des pluies nettes qui est captée par le réservoir de GR4J_asymP à remplissage égal. Une partition égale des précipitations nettes dans les deux modèles suppose donc un taux de remplissage plus élevé du réservoir de production de GR4J_asymP. La part des précipitations nettes captées par le réservoir de GR4J_asymP est néanmoins généralement plus grande (sous-figure c), signe que ce taux de remplissage plus élevé ne suffit pas à compenser les différences de rendement des précipitations des réservoirs de production des deux modèles. Les effets additionnés de la taille et du taux de remplissage plus élevé de GR4J_asymP résultent en une percolation nettement plus importante que celle de GR4J_asymE (sous-figure d). Cette percolation contribue vraisemblable-

ment à compenser la vitesse de remplissage du réservoir de GR4J_asymP.
Les écarts entre la percolation et le volume des précipitations non captées des deux modèles asymétriques illustrent finalement une différence de cheminement de l'eau dans la production : dans GR4J_asymP une grande partie fait un détour par le réservoir de production et est restituée sous forme de percolation (43% de la pluie efficace moyenne, contre 20% dans GR4J_asymE). En moyenne, le volume des précipitations efficaces simulé par GR4J_asymP est légèrement supérieur au volume simulé par GR4J_asymE (0,90 contre 0,88 mm j^{-1}).

Cet écart de précipitations efficaces est compensé par des apports souterrains plus importants pour GR4J_asymE (sous-figure e). Quant à l'évaporation réelle, elle est en moyenne un peu moins élevée dans le cas GR4J_asymP que dans celui de GR4J_asymE, qui évapore à taux presque potentiel sur un spectre très large de remplissage du réservoir de production. Le taux de remplissage du réservoir plus fort de GR4J_asymP compense cependant en partie son plus faible rendement de l'évaporation potentielle. Les évaporations simulées par deux modèles diffèrent en réalité moins par leur volume que par leur variations, le flux simulé par GR4J_asymP étant nettement plus stable que celui de GR4J_asymE.

Finalement, pour comprendre les effets des deux paramétrages sur l'écoulement produit, nous avons calculé la corrélation entre le flux entrant dans le réservoir de routage (Q_9) et le flux sortant du réservoir (Q_R). La corrélation entre ces deux écoulements atteint 0,91 pour GR4J_asymP et 0,85 pour GR4J_asymE. Il convient de remarquer que les tailles du réservoir de routage respectives des deux modèles sont presque égales. Cette différence suggère que les grands pics de pluie efficace simulés par GR4J_asymP coïncident mieux avec les situations de quasi saturation du réservoir de routage, et qu'inversement la pluie efficace est plus faible aux moments où le réservoir de routage est vide. Le réservoir de production du modèle GR4J_asymP assume donc un rôle plus affirmé de filtre passe-bas des précipitations. Nous suggérons que ce rôle est favorisé par la contribution de la percolation aux précipitations efficaces, dont les variations dans le temps sont plus lisses que celles des précipitations.

Il convient de vérifier que les observations relevées dans le cas particulier du bassin de la Teyssonne sont valables en moyenne pour le reste des bassins.

8.3.2.b. Comportement général

La Figure 8.26 présente une comparaison des flux moyens modélisés sur l'ensemble des bassins de l'échantillon par les deux modèles par rapport à ceux de GR4J, de la même façon que la Figure 8.16 présentait les flux de GR4J_sym.
De manière générale, les deux modèles asymétriques se distinguent par rapport à GR4J de la même façon que le modèle symétrique. Ce constat confirme que les effets des paramétrages des partitions des précipitations nettes et de l'évaporation nette s'additionnent dans GR4J_sym. En moyenne, l'évaporation réelle simulée par les deux modèles est supérieure à celle de GR4J, et inversement pour les précipitations efficaces. Les échanges souterrains sont décalés dans le sens d'une diminution des pertes ou de manière équivalente d'une augmentation des apports, selon les bassins considérés.
De manière générale, les écarts avec GR4J sont plus marqués dans le cas de GR4J_asymE. Les précipitations efficaces modélisées par GR4J_asymE sont en moyenne plus faibles que celles de GR4J_asymP, conduisant le modèle à diminuer d'autant les pertes souterraines pour boucler le bilan.

Certains constats formulés dans l'analyse précédente sur le bassin de la Teyssonne sont valables sur le reste de l'échantillon de bassins français et australiens. Le volume moyen de précipitations

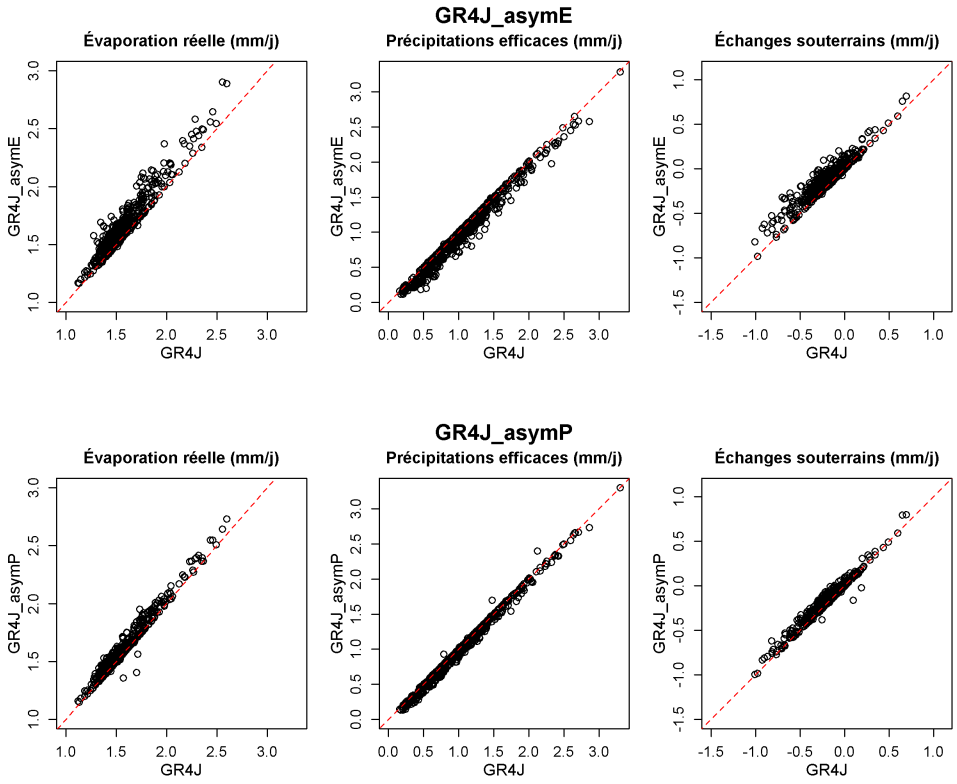


FIGURE 8.26 : Comparaison des flux moyens des modèles GR4J_asymE et GR4J_asymP (ordonnées) par rapport à ceux de GR4J (abscisse) simulés en calage sur les bassins de l'échantillon. Les précipitations efficaces correspondent au flux d'eau P_R en sortie de la fonction de production, alimentant dans les hydrogrammes unitaires. Les échanges souterrains (AE_{xch}) correspondent à la somme des composantes d'échange avec l'extérieur du système du réservoir de routage et du flot direct. L'évaporation réelle (notée AE) est calculée comme la somme des évaporations de l'interception et du réservoir de production.

nettes captées dans le réservoir de production augmente davantage dans le modèle GR4J_asymP (de 7,5% en moyenne) que dans le modèle GR4J_asymE (3,5%). L'influence de la percolation sur les précipitations efficaces est également plus importante dans le cas de GR4J_asymP. Près de 34% de la pluie efficace moyenne simulée par GR4J_asymP est issue de la percolation du réservoir de production, contre 28% pour GR4J_asymE et GR4J. Ce lissage de la pluie efficace a des conséquences similaires sur le routage du modèle : la corrélation moyenne entre les écoulements à l'entrée du réservoir de routage à sa sortie atteint 0,85 pour GR4J_asymP 0,82 pour GR4J_asymE. Les écarts moins marqués entre GR4J_asymP et GR4J_asymE en moyenne sur les bassins de l'échantillon en comparaison du cas de la Teyssonne à Changy s'expliquent par la valeur très élevée du paramètre α dans le cas particulier du bassin versant étudié précédemment, qui accentue les différences de comportement de chaque modèle.

8.3.3. Discussion des résultats

8.3.3.a. Pertinence des choix de modélisation

Dans l'ensemble, les performances des modèles asymétriques à 5 paramètres sont meilleures que celles de GR4J. Bien qu'aucun gain significatif ne soit observé dans la représentation des débits extrêmes (crues et étiages marqués), les débits intermédiaires sont quant à eux mieux simulés.

Aucune différence notable de performance n'est en revanche observable entre les deux modèles asymétriques.

L'analyse détaillée des divergences de comportement de ces versions modifiées de GR4J permet néanmoins de tester certaines hypothèses formulées au moment de la conceptualisation de ces modifications. In fine, elle offre quelques éléments de réflexion pour répondre à la question soulevée par une amélioration moyenne des performances : le modèle est-il meilleur pour les bonnes raisons ? En modélisation, l'analyse interne des états d'un modèle ne permet pas de répondre définitivement à une telle question tant qu'elle ne confronte pas ces états à des observations. En modélisation conceptuelle néanmoins, les composantes du modèle n'ont pas toujours vocation à représenter des éléments ou des processus parfaitement définis et compartimentés. De ce fait, il n'est pas forcément aisé de comparer directement à des observations les états et flux modélisés par ce genre de modèles, bien que cela ait déjà été fait pour certaines variables de GR4J (Le Moine, 2008; Bouaziz *et al.*, 2020). Plutôt que de répondre définitivement à la question des « bonnes » raisons, nous suggérons de répondre plus modestement à celle des intentions : le modèle s'est-il comporté tel que le modélisateur l'escomptait ?

Dans notre cas, le modèle GR4J_asymP a effectivement stocké une part plus importante des précipitations nettes dans le réservoir de production, et le modèle GR4J_asymE a évaporé en moyenne plus d'eau. En revanche, certains processus ont agi en compensation de ces comportements attendus. En premier lieu, la rétroaction négative du remplissage du réservoir de production a contribué à diminuer la part captée des précipitations nettes simulées par GR4J_asymP, ou le taux d'évaporation réelle par rapport à l'évaporation potentielle chez GR4J_asymE. En second lieu, le niveau de remplissage plus élevé du réservoir de production de GR4J_asymP a accru la participation de la percolation à l'écoulement, compensant un peu plus les pertes de précipitations efficaces. Ces rétroactions négatives peuvent brouiller légèrement la perception initiale des aménagements proposés pour le modèle. Néanmoins, elles évitent de perturber complètement le modèle en modifiant radicalement le rôle et la détermination de ses paramètres originaux.

8.3.3.b. Critique du paramétrage des modèles développés

Outre des différences dans leurs flux internes, la comparaison des deux modèles asymétriques a également mis en lumière des sensibilités différentes de leur comportement à la valeur du paramètre α. Le modèle GR4J_asymP offre peut-être un terrain moins favorable à l'équifinalité dans la mesure où des variations restreintes de la valeur du paramètre suffisent en général, tandis que dans le cas GR4J_asymE la limite supérieure (fixée à 8) de sa valeur est atteinte dans un nombre non négligeable de bassins. Ce contraste questionne la pertinence du modèle symétrique étudié au début du chapitre, pour lequel la même valeur de α paramétrise les équations de partition des précipitations et de l'évaporation nettes. Pour prendre ce résultat en considération, il pourrait être envisageable de proposer un modèle fondé sur le même principe avec deux paramètres supplémentaires au lieu d'un seul. Cette approche serait toutefois assez redondante dans la mesure où deux paramètres auraient des effets similaires sur le rendement de la fonction de production du modèle. La redondance des paramètres d'un tel modèle ferait vraisemblablement courir une difficulté accrue pour les algorithmes de calage pour converger vers l'optimum, un risque de sur-apprentissage dans certaines conditions de calage, sans parler de l'allongement du temps nécessaire à l'optimisation des paramètres. Une alternative consisterait éventuellement à lier les paramètres respectifs des deux équations de partition par une relation mathématique, par exemple une fonction affine, pour n'ajouter qu'un seul paramètre libre.

Parmi les trois amendements proposés pour la fonction de production de GR4J, nous suggérons

que le paramétrage asymétrique de la partition des précipitations nettes est le plus pertinent. Les trois paramétrages évalués ayant abouti à des performances similaires, cette suggestion se fonde sur arguments davantage qualitatifs que quantitatifs.
Le paramétrage implémenté dans GR4J_asymP est en effet le plus influent sur le comportement du modèle. Cette influence est matérialisée par une sensibilité accrue des performances du modèle et du rendement moyen des précipitations de la fonction de production à la valeur du paramètre α. Elle se traduit également par des changements notables sur le fonctionnement interne du modèle, en permettant de moduler la contribution relative de la percolation aux précipitations efficaces en sortie de la fonction de production du modèle.
L'interprétation du paramétrage asymétrique est en outre facilitée par rapport à son pendant symétrique. L'écart de la valeur de α à 2 renseigne en effet directement sur les différences entre la non-linéarité du réservoir par rapport aux entrées de précipitations et la non-linéarité du réservoir par rapport aux entrée d'évaporation potentielle.

8.4. Synthèse

Objectifs

L'objectif de ce chapitre était l'évaluation de modifications de la fonction de production de GR4J proposées à l'aune des résultats du Chapitre 7. Par l'introduction d'un cinquième paramètre α, une liberté supplémentaire a été octroyée au modèle pour moduler le rendement des précipitations de la fonction de production, à travers une modulation de la non-linéarité des équations de partition des précipitations et de l'évaporation nettes dans le réservoir de production.
Dans un premier temps, une version « symétrique » du modèle modifié a été testée : le même paramètre α contrôlait la non-linéarité de la réponse du réservoir aux forçages (précipitations et évaporation nettes). Ce modèle a été nommé GR4J_sym. Dans GR4J, ce paramètre est implicitement fixé à 2.
Dans un second temps, la réponse du modèle à l'un des forçages est restée telle que dans le modèle original, tandis que le paramètre α contrôlait la réponse à l'autre forçage. Les deux modèles ont de ce fait été baptisés « asymétriques », respectivement GR4J_asymP et GR4J_asymE, puisque seule la partition des précipitations nettes ou seule la partition de l'évaporation nette était influencée par la valeur de α. Ces deux versions asymétriques du paramétrage ont été évaluées afin de déterminer les contributions respectives du paramétrage des partitions des précipitations et de l'évaporation dans le réservoir de production.

Évaluation du paramétrage symétrique du réservoir de production

L'évaluation du modèle symétrique a montré que le paramétrage du réservoir de production conduisait à une amélioration très légère de la robustesse du modèle. Cette amélioration a notamment été conditionnée par une définition adéquate de l'ensemble des valeurs autorisées par le paramètre α. En effet, les cas où le paramétrage tendait à rendre le comportement du réservoir plus linéaire que celui du modèle original s'est avéré ne pas être judicieux en contrôle. Dans la plupart des bassins de l'échantillon, la valeur du paramètre α fixée dans le modèle GR4J original ($\alpha = 2$) semble donc être une valeur nominale pour le comportement du modèle.
Il est apparu que les bassins où GR4J_sym surclassait GR4J étaient généralement caractérisés par un climat plus sec et un régime hydrologique plus variable. Dans les bassins australiens, le paramétrage a notamment servi à diminuer le rendement moyen du modèle en compensation de données d'évaporation potentielle sous-estimées par la formule d'Oudin. Le paramétrage a en outre

permis d'éviter à la fonction d'échanges souterrains de gérer seule la fermeture du bilan hydrique du modèle, au bénéfice de la robustesse du modèle.

Évaluation du paramétrage asymétrique du réservoir de production

L'analyse des modèles asymétriques a mis en lumière la prépondérance du paramétrage de la partition des précipitations sur le comportement du modèle. Le paramétrage du rendement de l'évaporation potentielle ayant un effet indirect sur le rendement de la fonction de production du modèle, à travers le remplissage du réservoir, ce résultat semble logique.
Bien que les performances atteintes par les deux modèles asymétriques soient presque équivalentes à celles atteintes par le modèle symétrique, les deux modèles modélisent des cheminements de l'eau précipitée différents. La percolation du réservoir de production contribue ainsi nettement plus à la pluie efficace dans le cas de GR4J_asymP. La percolation joue un rôle moins important dans le cas GR4J_asymE, dont le comportement est proche de celui de GR4J. Le fonctionnement du réservoir de production de GR4J_asymP joue de ce fait un rôle accentué de filtre passe-bas des forçages climatiques. Ce comportement ne détériore pas pour autant la qualité de la représentation des crues.
Des trois modèles modèles testés, le modèle GR4_sym a obtenu les meilleures performances, mais l'écart avec les modèles asymétriques reste toutefois très léger.

Perspectives pour le paramétrage de la production de GR4J

Les tentatives de modification de GR4J réalisées dans ce chapitre se révèlent finalement intéressantes : le degré de complexité supplémentaire du modèle est en apparence facile d'interprétation, la robustesse et la transférabilité des paramètres sont légèrement améliorées, et les modifications du comportement du modèle semblent plausibles en première approche. Néanmoins, les changements opérés dans la structure présentent également quelques défauts, liés à la nécessité de calculer les variables par une résolution numérique des équations différentielles du modèle, alors que celles de GR4J étaient intégrables.
Un de ces défauts concerne les temps de calcul, qui sont nettement plus longs (entre 10 et 100 fois plus). Comparé à ceux des modèles dits « à base physique », ce coût reste certes relativement restreint. Toutefois, dans la perspective où les performances du modèle se sont améliorées assez marginalement, il peut être avantageux de chercher à réduire le temps de calcul tout en préservant les gains obtenus.
Un autre défaut concerne la moindre facilité d'interprétation des paramètres du modèle. Bien que la valeur du paramètre supplémentaire α indique clairement la non-linéarité de la réponse du réservoir, en revanche deux tailles de ce réservoir (X_1) ne sont plus comparables pour deux valeurs différentes de α. Dans la mesure où l'objectif de ce chapitre et des suivants est d'améliorer la robustesse de GR4J et de comprendre les raisons des échecs et/ou des réussites, préserver la facilité d'interprétation des quatre paramètres originaux du modèle garantit un certain confort de développement.

En perspective des développements conduits dans ce chapitre, nous suggérons donc une modification alternative de la fonction de production de GR4J, qui pourrait être évaluée dans de futurs travaux. Nous proposons l'implémentation d'un paramétrage inspiré des tentatives de Grigg et Hughes (2018) pour amender GR4J, qui repose sur l'utilisation de deux tailles différenciées du réservoir de production, impliquées respectivement dans la partition des précipitations nettes et dans la partition de l'évaporation nette du réservoir. Cette formulation pourrait permettre de recréer les conditions de modulation de la non-linéarité de la réponse du réservoir introduite dans ce chapitre, tout en respectant l'intégrabilité des équations et la parcimonie du modèle.
Il semble également nécessaire de comparer ces modifications à des ajustements plus triviaux de la fonction de production du modèle, qui ne concernent pas directement le réservoir de production.

Parmi ceux-ci, un paramètre multiplicatif simple de l'évaporation potentielle apparaît comme une solution extrêment simple pour corriger les évaporations potentielles sous-estimées et influencer le rendement du modèle. Un tel paramètrage a déjà été étudié par Le Moine *et al.* (2007) pour remplacer les échanges souterrains dans le modèle et finalement écarté pour des raisons physiques liées à l'hydrogéologie. Il pourrait néanmoins être intéressant d'articuler un facteur correctif de l'évaporation et des échanges souterrains simultanément dans la structure de GR4J. Par ailleurs, l'implémentation d'un réservoir d'interception en lieu et place de la neutralisation a déjà été proposée par Ficchì (2017) pour le modèle équivalent à GR4J au pas de temps horaire, et pourrait également participer à la modulation du rendement moyen des précipitations de la fonction de production.
Enfin, ces essais devraient être jugés à l'aune des modèles GR5J (Le Moine, 2008) et GR6J (Pushpalatha *et al.*, 2011), versions adaptées de GR4J respectivement pour la simulation des variations saisonnières des échanges nappes-rivière et pour la simulation des étiages.

Conclusion générale

Problématique de la thèse

Les modèles sont une représentation imparfaite de la réalité, qui comportent donc des erreurs. Si l'on est capable d'estimer l'ordre de grandeur de ces erreurs sur des périodes historiques, cette estimation devient impossible dans le futur. Or, le risque existe que des déviations de la réalité s'accumulent, rendant les erreurs encore plus fortes. Cet état de fait pose la question de la confiance que l'on peut accorder aux modèles ayant pour but d'estimer des variables futures. A fortiori, ce problème se pose pour les modèles hydrologiques, utilisés pour anticiper les effets du changement climatique sur les cours d'eau terrestres et notamment sur les ressources en eau futures.
Pour que les estimations de ces changements calculées par les modèles hydrologiques soient crédibles, il est nécessaire de s'assurer que ceux-ci répondent à une double exigence. Les modèles hydrologiques doivent en effet non seulement représenter tous les processus ayant une influence notable sur les écoulements, mais également proposer une représentation pertinente de ces processus. La capacité des modèles à simuler le débit des cours d'eau dans des conditions climatiques changeantes, aussi appelée robustesse des modèles hydrologiques, repose sur ces deux conditions. Il est crucial d'être en mesure d'évaluer la robustesse des modèles hydrologiques actuels, afin d'identifier leurs lacunes et de trouver des voies pour leur amélioration.

Afin d'évaluer les modèles hydrologiques dans les conditions défavorables typiquement rencontrées en contexte opérationnel, Klemeš (1986) a proposé dans un article visionnaire une procédure fondée sur la distinction entre période de calage et période de contrôle. Dans cette procédure, les observations de la période de calage sont supposées être utilisées pour déterminer et dimensionner les équations d'un modèle, tandis que les observations de la période de contrôle servent à évaluer les performances du modèle. Lorsque ces périodes sont choisies pour être représentatives de conditions hydro-climatiques aussi différentes que possible, cette procédure est connue sous le nom de *Differential Split-Sample Test* (DSST). Le DSST est aujourd'hui encore la référence pour l'évaluation de la robustesse des modèles hydrologiques, pourvu qu'ils soient suffisamment simples pour permettre la division des chroniques d'observations en plusieurs parties sans empêcher le calage de leurs paramètres. De nombreux modèles hydrologiques ont ainsi pu être confrontés à des conditions de simulation en climat varié et dans des contextes régionaux multiples. Le constat formulé à l'égard de la plupart des modèles ainsi testés est souvent le même : les modèles hydrologiques actuels manquent de robustesse et commettent des erreurs potentiellement préoccupantes concernant l'estimation des volumes moyens écoulés.

Le travail de cette thèse avait pour objectif d'explorer les voies d'amélioration de la robustesse des modèles hydrologiques en conditions climatiques changeantes. Nous avons défini deux axes principaux pour l'orientation de nos recherches. Le premier axe a consisté en l'évaluation des choix de modélisation en dehors de la structure et des équations des modèles. Les choix de modélisation étudiés ont concerné en particulier celui de la fonction objectif optimisée en phase de calage des paramètres des modèles et celui de la formule utilisée pour le calcul des entrées d'évaporation potentielle. Le deuxième axe s'est articulé autour de la critique spécifique de la structure d'un modèle hydrologique. Ces examens ont été accompagnés par des développements sur l'évaluation de la robustesse des modèles hydrologiques, afin de disposer d'outils adéquats pour réaliser les étapes de diagnostic. Tous les travaux ont été réalisés sur un échantillon de 377 bassins versants français et 181 bassins australiens, caractérisé par une grande hétérogénéité topographique, climatique et hydrogéologique.

Synthèse des travaux réalisés

Une nouvelle métrique d'évaluation de la robustesse des modèles hydrologiques

Les premiers travaux ont porté sur une métrique de performance spécifiquement consacrée à l'évaluation de la robustesse des modèles hydrologiques. Cette métrique, nommée *Proxy for Model Robustness* (PMR), s'appuie sur l'évolution du biais des modèles sur des sous-périodes d'une chronique d'observation définies par une fenêtre temporelle glissante. Le PMR permet d'estimer en première approche la capacité d'extrapolation climatique d'un modèle hydrologique sans recourir au DSST. Les valeurs du PMR reflètent en effet assez fidèlement les biais obtenus en appliquant la procédure traditonnelle de calage-contrôle.
Nous avons ensuite appliqué une procédure généralisée de calage-contrôle, appelée *Generalized Split-Sample Test* (GSST), pour diagnostiquer trois modèles hydrologiques conceptuels. Le PMR a été utilisé à cet égard conjointement à d'autres métriques de performance.

Des difficultés pour simuler les variations de la relation précipitations-débit

L'application du GSST à trois modèles hydrologiques a permis une analyse *a posteriori* de l'influence de plusieurs types de changements hydro-climatiques sur les performances des modèles. Parmi les changements considérés, ceux du rendement moyen des précipitations, c'est-à-dire de la part moyenne des précipitations effectivement écoulée à l'exutoire, se sont avérés être les plus défavorables à la qualité des estimations des volumes écoulés. Ce résultat signale une difficulté particulière des modèles évalués à simuler des variations de la réponse des bassins versants aux forçages climatiques, et donc une relative incapacité des modèles à représenter des modifications durables du comportement des bassins versants liées au changement climatique. Le rendement des précipitations n'étant pas une grandeur strictement météorologique, elle ne peut pas être utilisée de manière prédictive pour estimer ou corriger les erreurs des modèles *a posteriori*. Toutefois, l'application de procédures de DSST à partir des variations du rendement des précipitations peut servir à la construction d'un cadre d'évaluation des modèles hydrologiques véritablement exigeant.

Une marge d'amélioration existante mais limitée de la robustesse à travers le choix des techniques de calage

L'influence des conditions de calage des modèles hydrologiques a été examinée à travers la comparaison de quatre fonctions objectif pour le calage dans le cadre du GSST. Ces quatre fonctions objectif, choisies pour leur simplicité, constituent un échantillon assez représentatif à la fois des pratiques habituelles en hydrologie et de propositions récentes trouvées dans la littérature scientifique pour l'amélioration de la robustesse des modèles hydrologiques. Le critère de Kling-Gupta (KGE) donne des résultats très différents selon qu'il est calculé sur les débits non transformés ou sur la racine carrée des débits. Appliqué aux débits non transformés, le KGE est une des fonctions objectif les moins robustes : il est associé aux pertes de performances les plus importantes, notamment pour la simulation des étiages. Au contraire, le KGE appliqué à la racine carrée des débits est à la fois la fonction objectif contribuant aux meilleures performances mais également aux performances les moins sensibles aux conditions de calage. Le KGE appliqué aux débits non transformés étant une des fonctions objectif les plus utilisées dans les études sur la robustesse des modèles hydrologiques, ces conclusions indiquent que les conclusions de ces études sous-estiment vraisemblablement la robustesse réelle des modèles testés.
Par ailleurs, le constat que l'algorithme de calage employé n'était pas optimal pour le calage d'un des modèles hydrologiques nous a conduit à le comparer avec un algorithme alternatif. Alors que l'algorithme s'est révélé plus efficace pour le calage du modèle, aucune différence remarquable n'a

pu être constatée entre les performances en contrôle obtenues avec les deux algorithmes testés. Nos résultats tendent donc à montrer que le choix de la fonction objectif influence la robustesse des modèles hydrologiques. La comparaison des fonctions objectif nous conduit à recommander l'utilisation du KGE calculé sur la racine carrée des débits. Néanmoins, l'amélioration constatée sur la robustesse des modèles reste trop faible pour conclure que le calage des paramètres des modèles est responsable de tous les problèmes de robustesse des modèles hydrologiques.

Quelques recommandations quant au calcul de l'évaporation potentielle

Les voies d'amélioration de la robustesse des modèles hydrologiques ont également été étudiées sous l'angle des données d'évaporation potentielle, utilisées en entrée des modèles hydrologiques et qui déterminent la quantité maximale d'eau perdue quotidiennement par évaporation par un bassin versant. La comparaison de quatre formules d'évaporation potentielle différentes a dans un premier temps été réalisée dans le cadre du GSST. Ces formules, sélectionnées pour leur popularité en modélisation hydrologique, illustrent des conceptualisations variées du phénomène d'évaporation. Au contraire des résultats concernant les fonctions objectif, la comparaison des formules d'évaporation potentielle montre une certaine hétérogénéité selon les régions considérées. Pour les bassins australiens, la formule de Morton est associée à de meilleures performances des modèles hydrologiques en comparaison de la formule d'Oudin. Pour les bassins français, ces deux formules permettent d'obtenir des résultats équivalents et surclassent les performances obtenues avec les formules de Penman et Penman-Monteith.
Une analyse détaillée des bassins où les écarts de performances étaient les plus forts a débouché sur le constat que la détérioration des performances des modèles était simplement liée à une sous-estimation de la demande évaporatoire des formules concernées. Les modèles dont le bilan hydrique n'est modulé que par les pertes par évaporation étaient de ce fait les plus sensibles aux données d'évaporation potentielle sous-estimées. Une analyse des bassins concernés par ces erreurs a permis d'identifier certains facteurs géographiques communs et de proposer une critique de certaines des formules testées. Ces critiques demandent toutefois à être étayées par des études mettant ces formules directement en perspective avec des observations d'évaporation.
En complément de l'analyse des effets de la moyenne de l'évaporation potentielle, nous avons également développé un cadre d'analyse de l'adéquation des variations interannuelles de l'évaporation potentielle et du débit des bassins de l'échantillon sans recourir aux modèles hydrologiques. La méthode, basée sur des régressions linéaires, a mis en exergue des différences légères entre les formules testées. La formule d'Oudin s'est ainsi révélée être la plus à-même d'expliquer les variations interannuelles du débit des bassins français, et la formule de Morton a été identifiée comme la plus pertinente pour les bassins australiens.
Dans l'ensemble, il s'est donc avéré que le choix de la formule d'évaporation potentielle avait une influence équivalente à celui de la fonction objectif de calage : un choix judicieux permet d'améliorer la robustesse des modèles hydrologiques, toutefois cette amélioration est trop légère pour être considérée comme suffisante pour garantir la fiabilité des simulations hydrologiques en contexte de changement climatique. Qui plus est, le choix de la formule d'évaporation potentielle doit être déterminé à partir de considérations sur le contexte local de modélisation.

Une méthodologie d'identification des lacunes structurelles des modèles hydrologiques

Après avoir examiné les rôles respectifs des techniques de calage et du calcul des données d'évaporation potentielle dans les problèmes de robustesse des modèles hydrologiques, nous avons étudié celui de la structure des modèles à travers une méthodologie originale. La méthodologie a été conçue comme une rationalisation des analyses multi-critères, applicable sur de larges échantillons de bassins versants. Divisée en plusieurs étapes, elle vise essentiellement à identifier les composantes des modèles hydrologiques impliquées dans des compétitions entre plusieurs critères de

performance, qui imposent de devoir arbitrer entre eux. L'analyse repose sur la recherche de paramètres optimisant simultanément les performances de sous-ensembles de critères de performance, afin de déterminer des groupes de critères incompatibles entre eux. Les états internes du modèle obtenus avec les paramètres optimisant ces groupes de critères incompatibles sont ensuite comparés entre eux, dans le but de déterminer les composantes se comportant différemment selon les jeux de paramètres choisis.

Des voies d'amélioration du modèle GR4J encore à explorer

La méthodologie précédente a été appliquée à GR4J vis-à-vis du biais et du *Proxy for Model Robustness* sur plusieurs gammes d'écoulement définies ad hoc. L'analyse des résultats a révélé l'existence de compétitions délétères entre les capacités du modèle à simuler les variations interannuelles des débits de plusieurs gammes simultanément. Une première raison à l'origine de cette situation est l'impossibilité pour le modèle de simuler à la fois des dynamiques d'écoulement lentes caractéristiques des basses eaux et des dynamiques rapides caractéristiques des hautes eaux. Une solution proposée pour pallier cette incapacité consiste en l'ajout d'un réservoir de routage supplémentaire au modèle, de comportement différent du réservoir existant. Une deuxième raison est la dépendance très forte du modèle à la fonction d'échanges souterrains pour la gestion du bilan hydrique moyen des bassins. Cette dépendance est avant tout causée par une relative insensibilité du rendement moyen du réservoir de production aux valeurs des paramètres du modèle. Pour assouplir le comportement de la fonction de production du modèle, nous proposons l'ajout d'un paramètrage destiné à moduler ce rendement.
Une tentative de modification du modèle a été mise en œuvre dans le but de paramétrer le rendement du réservoir de production. Un paramètre fixé dans le modèle GR4J original a été laissé libre afin que le comportement du réservoir de production puisse être géré par deux paramètres différents. Ce paramétrage a permis une amélioration légère de la robustesse du modèle. Nos résultats indiquent en particulier un bénéfice pour la simulation du débit des bassins les plus secs et les plus variables de l'échantillon, ainsi que pour les bassins où l'évaporation potentielle était sous-estimée par la formule d'Oudin. Par ailleurs, l'analyse des états internes a clairement montré une atténuation du rôle de la fonction d'échanges souterrains dans la gestion du bilan hydrique du modèle, ce qui était l'objectif recherché.

Perspectives

Les développements menés dans cette thèse ont effleuré de nombreuses disciplines liées à la modélisation hydrologique, sans toutefois approfondir les recherches dans l'une d'entre elles. Cette orientation reflète la grande interconnexion entre ces champs displinaires à travers la structure d'un modèle hydrologique : évaporation des surfaces terrestres et transpiration des plantes, hydrogéologie, optimisation... De ce fait, les perspectives ouvertes sont relativement nombreuses. Toutefois, avant d'évoquer les pistes directement tirées de nos résultats, il est essentiel de remarquer que notre travail n'a pas abordé un sujet pourtant fondamental en hydrologie.
La question de la qualité des données de précipitations et de débits a été mise de côté à dessein pendant la thèse dans la mesure où les développements se sont concentrés sur les champs directement liés à la modélisation. Ces données ne sont pourtant pas parfaites et de nombreuses erreurs peuvent influencer négativement la robustesse des modèles, ne serait-ce qu'en perturbant le calage des paramètres. Améliorer la qualité de ces données devrait donc être un objectif incontournable pour faciliter les progrès futurs en modélisation hydrologique.

À l'aune de la synthèse bibliographique effectuée en début de thèse, nous constatons dans un premier temps le manque de repères communs à l'évaluation des modèles hydrologiques utilisés en projection climatique. Bien qu'ils s'inscrivaient lors de leur développement dans des contextes opérationnels différents, des modèles aussi dissemblables que les simples modèles conceptuels globaux, possédant quelques paramètres, et les modèles de surface continentale, fondés sur la résolution spatiale et temporelle d'équations physiques paramétrées, ont en projection climatique l'objectif commun de simuler des débits à l'exutoire de bassins versants. À cet égard, il nous paraît aujourd'hui important d'établir des comparaisons entre la robustesse de ces modèles pour guider l'estimation des incertitudes et la sélection des modèles. Dans la mesure où il ne requiert pas de procédures de calage-contrôle pour être informatif, le Proxy for Model Robustness développé au cours de la thèse pourrait s'inscrire dans la liste des critères d'évaluation utiles à de telles comparaisons.

D'autre part, les conclusions des travaux sur l'évaporation potentielle pointent les faiblesses de quelques formules dans certains contextes. En ce qui concerne la formule d'Oudin, nos résultats suggèrent une tendance à sous-estimer la demande évaporatoire dans les massifs forestiers. La formule d'Oudin figurant parmi les formules d'évaporation potentielle les plus simples utilisées en modélisation hydrologique, il pourrait être bénéfique d'envisager l'inclusion d'une variable supplémentaire à son calcul en cherchant à améliorer de manière empirique les performances des modèles hydrologiques dans ces bassins. Par ailleurs, les bons résultats de la formule de Morton constatés pour les bassins français invitent à considérer une optimisation de ses paramètres pour favoriser son utilisation en France. La formule de Morton étant construite à partir d'équations physiques, cette optimisation pourrait être réalisée par la confrontation à des mesures d'évaporation, à l'instar de ce qui a préalablement été réalisé avec succès dans le contexte australien. Quels que soient les résultats de ces potentiels travaux, étant donné la difficulté de sélectionner une formule générique adaptée à tous les contextes, nous suggérons que les incertitudes liées au choix de la formule d'évaporation devraient être davantage prises en considération pour les études d'impact du changement climatique.

Enfin, les maux structurels dont souffrent les modèles hydrologiques conceptuels vis-à-vis des simulations en climat changeant commencent à être reconnus et identifiés au sein de la communauté des hydrologues. Les modèles ont une incapacité relative à représenter les dynamiques interannuelles du régime hydrologique liée à un manque de mémoire des conditions climatiques passées. Nous espérons que cette reconnaissance suscitera une dynamique de développement de structures alternatives, spécialement conçues pour proposer une représentation des signaux à la fois basses-fréquences et hautes-fréquences. Pour guider ces efforts, il nous semble important d'appuyer les diagnostics des modèles par l'utilisation de données complémentaires au débit (mesures piézométriques, mesures gravimétriques, mesures de l'indice foliaire, couverture neigeuse, par exemple).

Bibliographie

N ADDOR, HX DO, C ALVAREZ-GARRETON, G COXON, K FOWLER et PA MENDOZA : Large-sample hydrology: recent progress, guidelines for new datasets and grand challenges. *Hydrological Sciences Journal*, 65(5):712–725, 2020.

N ADDOR, O RÖSSLER, N KÖPLIN, M HUSS, R WEINGARTNER et J SEIBERT : Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments. *Water Resources Research*, 50(10):7541–7562, 2014.

AC AITKEN : On least squares and linear combination of observations. *Proceedings of the Royal Society of Edinburgh*, 55:42–48, 1936.

H AJAMI, A SHARMA, LE BAND, JP EVANS, NK TUTEJA, GE AMIRTHANATHAN et MA BARI : On the non-stationarity of hydrological response in anthropogenically unaffected catchments: an Australian perspective. *Hydrology & Earth System Sciences*, 21(1):281–294, 2017.

H AKAIKE : Information theory and an extension of the maximum likelihood principle. *In Second International Symposium on Information Theory*, pages 267–281, 1973.

CD ALLEN, AK MACALADY, H CHENCHOUNI, D BACHELET, N MCDOWELL, M VENNETIER, T KITZBERGER, A RIGLING, DD BRESHEARS, EHT HOGG *et al.* : A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest ecology and management*, 259(4):660–684, 2010.

RG ALLEN, LS PEREIRA, D RAES, M SMITH *et al.* : Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome*, 300(9):D05109, 1998.

B AMBROISE, JL PERRIN et D REUTENAUER : Multicriterion validation of a semidistributed conceptual model of the water cycle in the Fecht Catchment (Vosges Massif, France). *Water Resources Research*, 31(6):1467–1481, 1995.

V ANDRÉASSIAN : *Impact de l'évolution du couvert forestier sur le comportement hydrologique des bassins versants*. Thèse de doctorat, Géosciences et Ressources Naturelles, Université Paris VI, 781 pp., 2002.

V ANDRÉASSIAN, L CORON, J LERAT et N LE MOINE : Climate elasticity of streamflow revisited–an elasticity index based on long-term hydrometeorological records. *Hydrology and Earth System Sciences*, 20(11):4503–4524, 2016.

V ANDRÉASSIAN, A HALL, N CHAHINIAN et J SCHAAKE : Introduction and synthesis: Why should hydrologists work on a large number of basin data sets? *In Large sample basin experiments for hydrological parametrization : results of the models parameter experiment- MOPEX. IAHS Red Books Series n°307*, pages 1–5. AISH, 2006.

V ANDRÉASSIAN, N LE MOINE, C PERRIN, MH RAMOS, L OUDIN, T MATHEVET, J LERAT et L BERTHET : All that glitters is not gold: the case of calibrating hydrological models. *Hydrological Processes*, 26:2206– 2210, 2012.

V ANDRÉASSIAN, C PERRIN, L BERTHET, N LE MOINE, J LERAT, C LOUMAGNE, L OUDIN, T MATHEVET, MH RAMOS et A VALÉRY : Crash tests for a standardized evaluation of hydrological models. *Hydrology and Earth System Sciences*, 10(13):1757–1764, 2009.

R ARSENAULT, F BRISSETTE et JL MARTEL : The hazards of split-sample validation in hydrological model calibration. *Journal of Hydrology*, 566:346–362, 2018.

P BAI, X LIU, K LIANG et C LIU : Comparison of performance of twelve monthly water balance models in different climatic catchments of China. *Journal of Hydrology*, 529:1030–1040, 2015.

P BAI, X LIU et C LIU : Improving hydrological simulations by incorporating GRACE data for model calibration. *Journal of Hydrology*, 557:291–304, 2018.

Y BARD : *Nonlinear parameter estimation*. New York, Academic Press, 341 pp., 1974.

S BERGSTRÖM : *Development and Application of a Conceptual Runoff Model for Scandinavian Catchments*. SMHI, Report No. RHO 7, Norköpping, 1976.

S BERGSTRÖM : *The HBV model - its structure and applications*. SMHI Hydrology, RH No.4, Norköpping, 1995.

L BERTHET, V ANDRÉASSIAN, C PERRIN et C LOUMAGNE : How significant are quadratic criteria? Part 2. On the relative contribution of large flood events to the value of a quadratic criterion. *Hydrological Sciences Journal–Journal des Sciences Hydrologiques*, 55(6):1063–1073, 2010.

F BESSON, P ETCHEVERS et S PRATS : Réponse au vœu 2020 de la Commission Hydrologie-Énergie sur l'ETP, 2020. Météo-France.

K BEVEN : Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in water resources*, 16(1):41–51, 1993.

K BEVEN : A manifesto for the equifinality thesis. *Journal of Hydrology*, 320(1-2):18–36, 2006.

K BEVEN et J FREER : Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 249(1-4):11–29, 2001.

K BEVEN et P YOUNG : A guide to good practice in modeling semantics for authors and referees. *Water Resources Research*, 49(8):5092–5098, 2013.

KJ BEVEN et MJ KIRKBY : A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Journal*, 24(1):43–69, 1979.

KJ BEVEN, PJ SMITH et JE FREER : So just why would a modeller choose to be incoherent? *Journal of Hydrology*, 354(1-4):15–32, 2008.

B BISSELINK, M ZAMBRANO-BIGIARINI, P BUREK et A DE ROO : Assessing the role of uncertain precipitation estimates on the robustness of hydrological model parameters under highly variable climate conditions. *Journal of Hydrology : Regional Studies*, 8:112–129, 2016.

S BLAZKOVA et K BEVEN : A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic. *Water Resources Research*, 45(12), 2009.

G BLÖSCHL, M FP BIERKENS, A CHAMBEL, C CUDENNEC, G DESTOUNI, A FIORI, J W KIRCHNER, J J MCDONNELL, H HG SAVENIJE, M SIVAPALAN *et al.* : Twenty-three unsolved problems in hydrology (UPH)–a community perspective. *Hydrological sciences journal*, 64(10):1141–1158, 2019.

G BLÖSCHL, M SIVAPALAN, H SAVENIJE, T WAGENER et A VIGLIONE : *Runoff prediction in ungauged basins: synthesis across processes, places and scales*. Cambridge University Press, 2013.

T BOSSHARD, M CARAMBIA, K GOERGEN, S KOTLARSKI, P KRAHE, M ZAPPA et C SCHÄR : Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections. *Water Resources Research*, 49(3):1523–1536, 2013.

LJE BOUAZIZ, G THIREL, T de BOER-EUSER, LA MELSEN, J BUITINK, CC BRAUER, J DE NIEL, S MOUSTAKAS, P WILLEMS, B GRELIER *et al.* : Behind the scenes of streamflow model performance. *Hydrology and Earth System Sciences Discussions*, pages 1–38, 2020.

MA BOUCHER, MH RAMOS, Q DUAN, F PAPPENBERGER, J THIELEN, A WOOD, HL CLOKE et JC SCHAAKE : Ensemble streamflow forecasts for hydropower systems. *Handb. Hydrometeorol. Ensemble Forecast*, pages 1–19, 2018.

RJ BOUCHET : Evapotranspiration réelle et potentielle, signification climatique. *IAHS Publ*, 62:134–142, 1963.

GEP BOX et DR COX : An analysis of transformations. *Journal of the Royal Statistical Society : Series B (Methodological)*, 26(2):211–243, 1964.

DP BOYLE, HV GUPTA et S SOROOSHIAN : Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resources Research*, 36 (12):3663–3674, 2000.

P BRIGODE, L OUDIN et C PERRIN : Hydrological model parameter instability: A source of additional uncertainty in estimating the hydrological impacts of climate change? *Journal of Hydrology*, 476:410–425, 2013.

C BRODERICK, T MATTHEWS, RL WILBY, S BASTOLA et C MURPHY : Transferability of hydrological models and ensemble averaging methods between contrasting climatic periods. *Water Resources Research*, 52(10):8343–8373, 2016.

W BRUTSAERT : Catchment-scale evaporation and the atmospheric boundary layer. *Water Resources Research*, 22(9S):39S–45S, 1986.

JP CARBONNEL et P HUBERT : Dictionnaire français d'hydrologie : Stationnaire (stationnarité). https://hydrologie.org/glu/FRDIC/DICSTATI.HTM, october 1995.

FHS CHIEW et TA MCMAHON : The applicability of Morton's and Penman's evapotranspiration estimates in rainfall-runoff modeling. *Journal of the American Water Resources Association*, 27 (4):611–620, 1991.

FHS CHIEW, MC PEEL, AW WESTERN *et al.* : Application and testing of the simple rainfall-runoff model SIMHYD. *Mathematical models of small watershed hydrology and applications*, pages 335–367, 2002a.

FHS CHIEW, NJ POTTER, J VAZE, C PETHERAM, L ZHANG, J TENG et DA POST : Observed hydrologic non-stationarity in far south-eastern Australia: implications for modelling and prediction. *Stochastic Environmental Research and Risk Assessment*, 28(1):3–15, 2014.

FHS CHIEW, MJ STEWARDSON et TA MCMAHON : Comparison of six rainfall-runoff modelling approaches. *Journal of Hydrology*, 147(1-4):1–36, 1993.

FHS CHIEW, J TENG, J VAZE, DA POST, JM PERRAUD, DGC KIRONO et NR VINEY : Estimating climate change impact on runoff across southeast Australia: Method, results, and implications of the modeling method. *Water Resources Research*, 45(10), 2009.

FHS CHIEW, QJ WANG, F MCCONACHY, R JAMES, W WRIGHT, G de HOEDT *et al.* : Evapotranspiration maps for Australia. *In Water Challenge : Balancing the Risks : Hydrology and Water Resources Symposium 2002*, page 167. Institution of Engineers, Australia, 2002b.

MP CLARK et D KAVETSKI : Ancient numerical daemons of conceptual hydrological modeling: 1. Fidelity and efficiency of time stepping schemes. *Water Resources Research*, 46(10), 2010.

MP CLARK, D KAVETSKI et F FENICIA : Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resources Research*, 47(9), 2011.

MP CLARK, B NIJSSEN, JD LUNDQUIST, D KAVETSKI, DE RUPP, RA WOODS, JE FREER, ED GUTMANN, AW WOOD, LD BREKKE *et al.* : A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resources Research*, 51(4):2498–2514, 2015a.

MP CLARK, B NIJSSEN, JD LUNDQUIST, D KAVETSKI, DE RUPP, RA WOODS, JE FREER, ED GUTMANN, AW WOOD, DJ GOCHIS *et al.* : A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. *Water Resources Research*, 51(4):2515–2542, 2015b.

MP CLARK, RL WILBY, ED GUTMANN, JA VANO, S GANGOPADHYAY, AW WOOD, HJ FOWLER, C PRUDHOMME, JR ARNOLD et LD BREKKE : Characterizing uncertainty of the hydrologic impacts of climate change. *Current Climate Change Reports*, 2(2):55–64, 2016.

TA COHN et HF LINS : Nature's style: Naturally trendy. *Geophysical research letters*, 32(23), 2005.

L CORON : *Les modèles hydrologiques conceptuels sont-ils robustes face à un climat en évolution ? Diagnostic sur un échantillon de bassins versants français et australiens*. Thèse de doctorat, Institut des Sciences et Industries du Vivant et de l'Environnement, AgroParisTech, 364 pp., 2013.

L CORON : *airGRplus: Additional functions to the airGR package*, 2020. R package version 0.8.1.2.

L CORON, V ANDRÉASSIAN, C PERRIN, J LERAT, J VAZE, M BOURQUI et F HENDRICKX : Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments. *Water Resources Research*, 48(5), 2012.

L CORON, V ANDRÉASSIAN, C PERRIN, M BOURQUI et F HENDRICKX : On the lack of robustness of hydrologic models regarding water balance simulation: a diagnostic approach applied to three models of increasing complexity on 20 mountainous catchments. *Hydrology and Earth System Sciences*, 2014.

L CORON, O DELAIGUE, G THIREL, C PERRIN, C MICHEL, V ANDRÉASSIAN, F BOURGIN, P BRIGODE, N LE MOINE, T MATHEVET *et al.* : airGR: Suite of GR hydrological models for precipitation-runoff modelling (v. 1.0.12.3.2). 2018. URL `https://webgr.inrae.fr/en/airGR/`.

L CORON, G THIREL, O DELAIGUE, C PERRIN et V ANDRÉASSIAN : The suite of lumped GR hydrological models in an R package. *Environmental modelling & software*, 94:166–171, 2017.

C COSANDEY et M ROBINSON : *Hydrologie continentale*. Armand Colin, 2012.

G COXON, J FREER, R LANE, T DUNNE, WJM KNOBEN, NJK HOWDEN, N QUINN, T WAGENER et R WOODS : DECIPHeR v1: Dynamic fluxEs and ConnectIvity for Predictions of HydRology. *Geoscientific Model Development*, 12(6), 2019.

G COXON, J FREER, T WAGENER, NA ODONI et M CLARK : Diagnostic evaluation of multiple hypotheses of hydrological behaviour in a limits-of-acceptability framework for 24 UK catchments. *Hydrological Processes*, 28(25):6135–6150, 2014.

RE CRISS et WE WINSTON : Do Nash values have value? Discussion and alternate proposals. *Hydrological Processes : An International Journal*, 22(14):2723–2725, 2008.

L CROCHEMORE, C PERRIN, V ANDRÉASSIAN, U EHRET, SP SEIBERT, S GRIMALDI, H GUPTA et JE PATUREL : Comparing expert judgement and numerical criteria for hydrograph evaluation. *Hydrological sciences journal*, 60(3):402–423, 2015.

H DAKHLAOUI, D RUELLAND et Y TRAMBLAY : A bootstrap-based differential split-sample test to assess the transferability of conceptual rainfall-runoff models under past and future climate variability. *Journal of Hydrology*, 575:470–486, 2019.

H DAKHLAOUI, D RUELLAND, Y TRAMBLAY et Z BARGAOUI : Evaluating the robustness of conceptual rainfall-runoff models under climate variability in northern Tunisia. *Journal of Hydrology*, 550:201–217, 2017.

DR DAWDY et RW LICHTY : Methodology of hydrologic model building. *IAHS Publ*, 81:347–355, 1968.

T de BOER-EUSER, L BOUAZIZ, J DE NIEL, C BRAUER, B DEWALS, G DROGUE, F FENICIA, B GRELIER, J NOSSENT, F PEREIRA *et al.* : Looking beyond general metrics for model comparison-lessons from an international model intercomparison study. *Hydrology and Earth System Sciences*, 2017.

NJ DE VOS, THM RIENTJES et HV GUPTA : Diagnostic evaluation of conceptual rainfall–runoff models using temporal clustering. *Hydrological Processes*, 24(20):2840–2850, 2010.

O DELAIGUE, B GÉNOT, L LEBECHEREL, P BRIGODE et PY BOURGIN : Base de données hydroclimatiques observées à l'échelle de la France, 2020. Université Paris-Saclay, INRAE, HYCAR Research Unit, Hydrology group, Antony, `https://webgr.inrae.fr/base-de-donnees/`.

C DESER, A PHILLIPS, V BOURDETTE et H TENG : Uncertainty in climate change projections: the role of internal variability. *Climate dynamics*, 38(3-4):527–546, 2012.

A DEVERS : *Vers une réanalyse hydrométéorologique à l'échelle de la France sur les 150 dernières années par assimilation de données dans des reconstructions ensemblistes*. Thèse de doctorat, Université Grenoble Alpes, 288 pp., 2019.

C DOBLER, S HAGEMANN, RL WILBY et J STÄTTER : Quantifying different sources of uncertainty in hydrological projections in an Alpine watershed. *Hydrology and Earth System Sciences*, 16:4343–4360, 2012.

LM DONNELLY-MAKOWECKI et RD MOORE : Hierarchical testing of three rainfall–runoff models in small forested catchments. *Journal of Hydrology*, 219(3-4):136–152, 1999.

JCI DOOGE : Hydrology in perspective. *Hydrological Sciences Journal*, 33(1):61–85, 1988.

Q DUAN, S SOROOSHIAN et HV GUPTA : Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, 28(4):1015–1031, 1992.

D DUETHMANN, G BLÖSCHL et J PARAJKA : Why does a conceptual hydrological model fail to correctly predict discharge changes in response to climate change? *Hydrology and Earth System Sciences*, 24(7):3493–3511, 2020.

G DUROZOY : *Ressources en eau de l'Ouest du département des Alpes Maritimes*. Doc. BRGM/FRANKARST, 79 SGN 324 PAC, Marseille, 1979.

E DUTRA, G BALSAMO, P VITERBO, PMA MIRANDA, A BELJAARS, C SCHÄR et K ELDER : An improved snow scheme for the ECMWF land surface model: Description and offline validation. *Journal of Hydrometeorology*, 11(4):899–916, 2010.

N EDIJATNO : *Mise au point d'un modèle élémentaire pluie-débit au pas de temps journalier*. Thèse de doctorat, Université Louis Pasteur/ENGEES, Strasbourg, 253 pp., 1991.

N EDIJATNO, NO NASCIMENTO, X YANG, Z MAKHLOUF et C MICHEL : GR3J: a daily watershed model with three free parameters. *Hydrological Sciences Journal*, 44(2):263–277, 1999.

T EUSER, HC WINSEMIUS, M HRACHOWITZ, F FENICIA, S UHLENBROOK et HHG SAVENIJE : A framework to assess the realism of model structures using hydrological signatures. *Hydrology and Earth System Sciences, 17 (5), 2013*, 2013.

G EVIN, D KAVETSKI, M THYER et G KUCZERA : Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration. *Water Resources Research*, 49(7):4518–4524, 2013.

F FENICIA, D KAVETSKI et HHG SAVENIJE : Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development. *Water Resources Research*, 47(11), 2011.

JC FERNANDEZ, L MOUNIER et C PACHON : A model-based approach for robustness testing. *In IFIP International Conference on Testing of Communicating Systems*, pages 333–348. Springer, 2005.

A FICCHÌ : *An adaptive hydrological model for multiple time-steps: Diagnostics and improvements based on fluxes consistency*. Thèse de doctorat, Géosciences, Ressources Naturelles et Environnement, Université Pierre et Marie Curie, 281 pp., 2017.

A FICCHÌ, C PERRIN et V ANDRÉASSIAN : Impact of temporal resolution of inputs on hydrological model performance: An analysis based on 2400 flood events. *Journal of Hydrology*, 538:454–470, 2016.

K FOWLER : *Towards improved rainfall-runoff modelling in changing climatic conditions*. Thèse de doctorat, University of Melbourne, 400 pp., 2017.

K FOWLER, G COXON, J FREER, M PEEL, T WAGENER, A WESTERN, R WOODS et L ZHANG : Simulating runoff under changing climatic conditions: a framework for model improvement. *Water Resources Research*, 54(12):9812–9832, 2018a.

K FOWLER, G COXON, J FREER, M PEEL, T WAGENER, A WESTERN, R WOODS et L ZHANG : Flip the bucket: why models based on deficits, not buckets, might be key to runoff projections in a drying climate. *In Geophysical Research Abstracts*, volume 21, 2019.

K FOWLER, W KNOBEN, M PEEL, T PETERSON, D RYU, M SAFT, KW SEO et A WESTERN : Many commonly used rainfall-runoff models lack long, slow dynamics: implications for runoff projections. *Water Resources Research*, 56(5):e2019WR025286, 2020.

K FOWLER, M PEEL, A WESTERN et L ZHANG : Improved rainfall-runoff calibration for drying climate: Choice of objective function. *Water Resources Research*, 54(5):3392–3408, 2018b.

K FOWLER, MC PEEL, AW WESTERN, L ZHANG et TJ PETERSON : Simulating runoff under changing climatic conditions: Revisiting an apparent deficiency of conceptual rainfall-runoff models. *Water Resources Research*, 52(3):1820–1846, 2016.

F GARAVAGLIA : *Méthode SCHADEX de prédétermination des crues extrêmes : Méthodologie, applications, études de sensibilité*. Thèse de doctorat, Terre Univers Environnement, Université de Grenoble, 417 pp., 2011.

R GARÇON : Prévision opérationnelle des apports de la Durance à Serre-Ponçon à l'aide du modèle MORDOR. Bilan de l'année 1994–1995. *La Houille Blanche*, (5):71–76, 1996.

S GHARARI, M HRACHOWITZ, F FENICIA et HHG SAVENIJE : An approach to identify time consistent model parameters: sub-period calibration. *Hydrology and Earth System Sciences*, 17(1):149, 2013.

I GIUNTOLI, JP VIDAL, C PRUDHOMME et DM HANNAH : Future hydrological extremes: the uncertainty from multiple global climate and global hydrological models. *Earth System Dynamics*, 6 (1):267–285, 2015.

RJ GRANGER : An examination of the concept of potential evaporation. *Journal of Hydrology*, 111 (1-4):9–19, 1989.

AH GRIGG et JD HUGHES : Nonstationarity driven by multidecadal change in catchment groundwater storage: A test of modifications to a common rainfall–runoff model. *Hydrological Processes*, 32 (24):3675–3688, 2018.

M GUIMBERTEAU, D ZHU, F MAIGNAN, Y HUANG, C YUE, S DANTEC-NÉDÉLEC, C OTTLÉ, A JORNET-PUIG, A BASTOS, P LAURENT *et al.* : ORCHIDEE-MICT (v8. 4.1), a land surface model for the high latitudes: model description and validation. *Geoscientific Model Development*, 11(1):121–163, 2018.

D GUO, S WESTRA et HR MAIER : Impact of evapotranspiration process representation on runoff projections from conceptual rainfall-runoff models. *Water Resources Research*, 53(1):435–454, 2017a.

D GUO, S WESTRA et HR MAIER : Sensitivity of potential evapotranspiration to changes in climate variables for different Australian climatic zones. *Hydrology and Earth System Sciences*, 21(4): 2107–2126, 2017b.

HV GUPTA, H KLING, KK YILMAZ et GF MARTINEZ : Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1-2):80–91, 2009.

HV GUPTA, C PERRIN, G BLÖSCHL, A MONTANARI, R KUMAR, M CLARK et V ANDRÉASSIAN : Large-sample hydrology: a need to balance depth with breadth. *Hydrology and Earth System Sciences*, 18(2):463–477, 2014.

HV GUPTA, S SOROOSHIAN et PO YAPO : Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research*, 34(4):751–763, 1998.

HV GUPTA, T WAGENER et Y LIU : Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrological Processes : An International Journal*, 22(18):3802–3813, 2008.

F HABETS : *Modélisation du cycle continental de l'eaua l'échelle régionale : application aux bassins versants de l'Adour et du Rhône*. Thèse de doctorat, Université Paul Sabatier, Toulouse, France, 224 pp., 1998.

F HABETS, A BOONE, JL CHAMPEAUX, P ETCHEVERS, L FRANCHISTEGUY, E LEBLOIS, E LEDOUX, P LE MOIGNE, E MARTIN, S MOREL *et al.* : The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France. *Journal of Geophysical Research : Atmospheres*, 113(D6), 2008.

S HAGEMANN, C CHEN, DB CLARK, S FOLWELL, SN GOSLING, I HADDELAND, N HANASAKI, J HEINKE, F LUDWIG, F VOSS *et al.* : Climate change impact on available water resources obtained using multiple global climate and hydrology models. *Earth System Dynamics*, 4(1):129–144, 2013.

J HAN, Y YANG, ML RODERICK, TR MCVICAR, D YANG, S ZHANG et HE BECK : Assessing the steady-state assumption in water balance calculation across global catchments. *Water Resources Research*, 56(7):e2020WR027392, 2020.

N HANSEN, SD MÜLLER et P KOUMOUTSAKOS : Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evolutionary computation*, 11 (1):1–18, 2003.

J HARLIN : Development of a process oriented calibration scheme for the HBV hydrological model. *Hydrology Research*, 22(1):15–36, 1991.

G HARTMANN et A BÁRDOSSY : Investigation of the transferability of hydrological models and a method to improve model calibration. *Advances in Geosciences*, 5:83–87, 2005.

B HINGRAY et M SAÏD : Partitioning internal variability and model uncertainty components in a multimember multimodel ensemble of climate projections. *Journal of Climate*, 27(17):6779–6798, 2014.

RP HOOPER, A STONE, N CHRISTOPHERSEN, E de GROSBOIS et HM SEIP : Assessing the Birkenes model of stream acidification using a multisignal calibration methodology. *Water Resources Research*, 24(8):1308–1316, 1988.

D HOUNAM et M WERNER : The shuttle radar topography mission (srtm). *In Proceedings of ISPRS workshop-Sensors and Mapping from Space*. Citeseer, 1999.

M HRACHOWITZ et MP CLARK : HESS Opinions: The complementary merits of competing modelling philosophies in hydrology. *Hydrol. Earth Syst. Sci*, 21(8):3953–3973, 2017.

J HUGHES, R SILBERSTEIN, A GRIGG et B MOUNTAIN : Extending rainfall–runoff models for use in environments with long–term catchment storage and forest cover changes. *In MODSIM2013, 20th International Congress on Modelling and Simulation*, pages 2471–2477. Modelling and Simulation Society of Australia and New Zealand December, 2013.

JD HUGHES, KC PETRONE et RP SILBERSTEIN : Drought, groundwater storage and stream flow decline in southwestern Australia. *Geophysical Research Letters*, 39(3), 2012.

L HUGHES : Climate change and Australia: trends, projections and impacts. *Austral Ecology*, 28 (4):423–443, 2003.

C HUTTON, T WAGENER, J FREER, D HAN, C DUFFY et B ARHEIMER : Most computational hydrology is not reproducible, so is it really science? *Water Resources Research*, 52(10):7548–7555, 2016.

DI JAYATHILAKE et T SMITH : Assessing the impact of PET estimation methods on hydrologic model performance. *Hydrology Research*, 2020a.

DI JAYATHILAKE et T SMITH : Understanding the role of hydrologic model structures on evapotranspiration-driven sensitivity. *Hydrological Sciences Journal*, 65(9):1474–1489, 2020b.

LA KAUARK-LEITE et N NASCIMENTO : Développement, utilisation et incertitudes des modèles conceptuels en hydrologie. *In Séminaire Eaux Environnement 3 : modélisation du comportement des polluants dans les hydrosystèmes, Paris, 1993*, pages 191–219, 1993.

D KAVETSKI et F FENICIA : Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights. *Water Resources Research*, 47(11), 2011.

D KAVETSKI, SW FRANKS, G KUCZERA *et al.* : Confronting input uncertainty in environmental modelling. *Calibration of watershed models*, 6:49–68, 2003.

CA KELLEHER et SB SHAW : Is ET often oversimplified in hydrologic models? Using long records to elucidate unaccounted for controls on ET. *Journal of Hydrology*, 557:160–172, 2018.

JW KIRCHNER : Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research*, 42(3), 2006.

V KLEMEŠ : Conceptualization and scale in hydrology. *Journal of Hydrology*, 65(1-3):1–23, 1983.

V KLEMEŠ : Operational testing of hydrological simulation models. *Hydrological Sciences Journal*, 31(1):13–24, 1986.

H KLING, M FUCHS et M PAULIN : Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, 424:264–277, 2012.

R KNUTTI et J SEDLÁČEK : Robustness and uncertainties in the new CMIP5 climate model projections. *Nature Climate Change*, 3(4):369–373, 2013.

D KOUTSOYIANNIS et A MONTANARI : Negligent killing of scientific concepts: the stationarity case. *Hydrological Sciences Journal*, 60(7-8):1174–1183, 2015.

P KRAUSE : *Das hydrologische Modellsystem J2000: Beschreibung und Anwendung in großen Flußgebieten*, volume 29 de *Schriften des Forschungszentrums Jülich. Reihe Umwelt*. Forschungszentrum Jülich, Zentralbibliothek, 2001.

P KRAUSE : Quantifying the impact of land use changes on the water balance of large catchments using the J2000 model. *Physics and Chemistry of the Earth, Parts A/B/C*, 27(9-10):663–673, 2002.

P KRAUSE, DP BOYLE et F BÄSE : Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 5:89–97, 2005.

G KRINNER, N VIOVY, N de NOBLET-DUCOUDRÉ, J OGÉE, J POLCHER, P FRIEDLINGSTEIN, P CIAIS, S SITCH et IC PRENTICE : A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, 19(1), 2005.

G KUCZERA : Prediction of water yield reductions following a bushfire in ash-mixed species eucalypt forest. *Journal of Hydrology*, 94(3-4):215–236, 1987.

G KUCZERA, D KAVETSKI, S FRANKS et M THYER : Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters. *Journal of Hydrology*, 331(1-2):161–177, 2006.

G KUCZERA et M MROCZKOWSKI : Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. *Water Resources Research*, 34(6):1481–1489, 1998.

R KUMAR, L SAMANIEGO et S ATTINGER : Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations. *Water Resources Research*, 49(1):360–379, 2013.

M LAFAYSSE, B HINGRAY, A MEZGHANI, J GAILHARD et L TERRAY : Internal variability and model uncertainty components in future hydrometeorological projections: The Alpine Durance basin. *Water Resources Research*, 50(4):3317–3341, 2014.

GN LANCE et WT WILLIAMS : A general theory of classificatory sorting strategies: 1. Hierarchical systems. *The computer journal*, 9(4):373–380, 1967.

M LANSIART : *Étude hydrogéologique de la région de Soulaines-Dhuys (Aube)*. Doc. BRGM/CEA, 86 SGN 432 STO, 1979.

H LAURI, H DE MOEL, PJ WARD, TA RÄSÄNEN, M KESKINEN et M KUMMU : Future changes in Mekong River hydrology: impact of climate change and reservoir operation on discharge. *Hydrology and Earth System Sciences*, 16(5):4603–4619, 2012.

P LE MOIGNE, F BESSON, E MARTIN, J BOÉ, A BOONE, B DECHARME, P ETCHEVERS, S FAROUX, F HABETS, M LAFAYSSE *et al.* : The latest improvements with SURFEX v8. 0 of the Safran–Isba–Modcou hydrometeorological model for France. *Geoscientific Model Development*, 13(9):3925–3946, 2020.

N LE MOINE : *Le bassin versant de surface vu par le souterrain : une voie d'amélioration des performances et du réalisme des modèles pluie-débit ?* Thèse de doctorat, Géosciences et Ressources Naturelles, Université Pierre et Marie Curie, 324 pp., 2008.

N LE MOINE, V ANDRÉASSIAN et T MATHEVET : Confronting surface-and groundwater balances on the La Rochefoucauld-Touvre karstic system (Charente, France). *Water Resources Research*, 44 (3), 2008.

N LE MOINE, V ANDRÉASSIAN, C PERRIN et C MICHEL : How can rainfall-runoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments. *Water resources research*, 43(6), 2007.

L LEBECHEREL : *Sensibilité des calculs hydrologiques à la densité des réseaux de mesure hydrométrique et pluviométrique*. Thèse de doctorat, Géosciences, Ressources Naturelles et Environnement, Université Pierre et Marie Curie, 280 pp., 2015.

MJ LEBLANC, P TREGONING, G RAMILLIEN, SO TWEED et A FAKES : Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia. *Water Resources Research*, 45(4), 2009.

E LEDOUX, G GIRARD, G DE MARSILY, JP VILLENEUVE et J DESCHENES : Spatially distributed modeling: conceptual approach, coupling surface water and groundwater. *In Unsaturated Flow in Hydrologic Modeling*, pages 435–454. Springer, 1989.

I LELEU, I TONNELIER, R PUECHBERTY, P GOUIN, I VIQUENDI, L COBOS, A FORAY, M BAILLON et PO NDIMA : La refonte du système d'information national pour la gestion et la mise à disposition des données hydrométriques. *La Houille Blanche*, (1):25–32, 2014.

JP LHOMME : Towards a rational definition of potential evaporation. *Hydrology and Earth System Sciences*, 1(2):257–264, 1997.

CZ LI, L ZHANG, H WANG, YQ ZHANG, FL YU, DH YAN et N VERHOEST : The transferability of hydrological models under nonstationary climatic conditions. *Hydrology & Earth System Sciences*, 16(4), 2012.

G LINDSTRÖM, B JOHANSSON, M PERSSON, M GARDELIN et S BERGSTRÖM : Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology*, 201(1-4):272–288, 1997.

Y LIU, J FREER, K BEVEN et P MATGEN : Towards a limits of acceptability approach to the calibration of hydrological models: Extending observation error. *Journal of Hydrology*, 367(1-2):93–103, 2009.

Z LIU : Dynamics of interdecadal climate variability: A historical perspective. *Journal of Climate*, 25 (6):1963–1995, 2012.

F LOBLIGEOIS : *Mieux connaître la distribution spatiale des pluies améliore-t-il la modélisation des crues ? Diagnostic sur 181 bassins versants français*. Thèse de doctorat, Institut des Sciences et Industries du Vivant et de l'Environnement, AgroParisTech, 312 pp., 2014.

JA LUQUE-ESPINAR, M CHICA-OLMO, E PARDO-IGÚZQUIZA et MJ GARCÍA-SOLDADO : Influence of climatological cycles on hydraulic heads across a Spanish aquifer. *Journal of Hydrology*, 354 (1-4):33–52, 2008.

H MADSEN : Automatic calibration of a conceptual rainfall–runoff model using multiple objectives. *Journal of Hydrology*, 235(3-4):276–288, 2000.

H MADSEN : Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives. *Advances in water resources*, 26(2):205–216, 2003.

HB MANN et DR WHITNEY : On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, pages 50–60, 1947.

J MARGAT : Carte hydrogéologique de la France à l'échelle 1/1500 000 – Systèmes aquifères, 1980. Marseille, BRGM.

T MATHEVET : *Quels modèles pluie-débit globaux au pas de temps horaire ? Développements empiriques et comparaison de modèles sur un large échantillon de bassins versants*. Thèse de doctorat, ENGREF Paris, 463 pp., 2005.

T MATHEVET, HV GUPTA, C PERRIN, V ANDRÉASSIAN et N LE MOINE : Assessing the performance and robustness of two conceptual rainfall-runoff models on a worldwide sample of watersheds. *Journal of Hydrology*, 585:124698, 2020.

T MATHEVET, C MICHEL, V ANDRÉASSIAN, C PERRIN *et al.* : A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins. *IAHS Publ*, 307:211, 2006.

TA MCMAHON, MC PEEL, L LOWE, R SRIKANTHAN, TR MCVICAR *et al.* : Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis. *Hydrol. Earth Syst. Sci*, 17(4):1331–1363, 2013.

H MCMILLAN, T KRUEGER et J FREER : Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality. *Hydrological Processes*, 26(26):4078–4111, 2012.

H McMillan, I Westerberg et F Branger : Five guidelines for selecting hydrological signatures. *Hydrological Processes*, 31(26):4757–4761, 2017.

TR McVicar, TG Van Niel, LTao Li, ML Roderick, DP Rayner, L Ricciardulli et RJ Donohue : Wind speed climatology and trends for Australia, 1975–2006: Capturing the stilling phenomenon and comparison with near-surface reanalysis output. *Geophysical Research Letters*, 35(20), 2008.

LA Melsen, N Addor, N Mizukami, AJ Newman, PJJF Torfs, MP Clark, R Uijlenhoet et AJ Teuling : Mapping (dis) agreement in hydrologic projections. *Hydrology and Earth System Sciences*, 22(3):1775–1791, 2018.

R Merz, J Parajka et G Blöschl : Time stability of catchment model parameters: Implications for climate impact analyses. *Water Resources Research*, 47(2), 2011.

C Michel : *Hydrologie appliquée aux petits bassins versants ruraux*. Thèse de doctorat, Cemagref, Antony, 1989.

C Michel, C Perrin et V Andréassian : The exponential store: a correct formulation for rainfall—runoff modelling. *Hydrological Sciences Journal*, 48(1):109–124, 2003.

PCD Milly, J Betancourt, M Falkenmark, RM Hirsch, ZW Kundzewicz, DP Lettenmaier et RJ Stouffer : Stationarity is dead: Whither water management? *Earth*, 4:20, 2008.

PCD Milly et KA Dunne : Potential evapotranspiration and continental drying. *Nature Climate Change*, 6(10):946–949, 2016.

C Milzow, PE Krogh et P Bauer-Gottwein : Combining satellite radar altimetry, SAR surface soil moisture and GRACE total storage changes for hydrological model calibration in a large poorly gauged catchment. *Hydrology & Earth System Sciences*, 15(6), 2011.

A Montanari et E Toth : Calibration of hydrological models in the spectral domain: An opportunity for scarcely gauged basins? *Water Resources Research*, 43(5), 2007.

C Monteil, F Zaoui, N Le Moine et F Hendrickx : Multi-objective calibration by combination of stochastic and gradient-like parameter generation rules–the caRamel algorithm. *Hydrology and Earth System Sciences*, 24(6):3189–3209, 2020.

JL Monteith : Evaporation and environment. *In Symposia of the society for experimental biology*, volume 19, pages 205–234. Cambridge University Press (CUP) Cambridge, 1965.

FI Morton : Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. *Journal of Hydrology*, 66(1-4):1–76, 1983.

FI Morton : Evaporation research—a critical review and its lessons for the environmental sciences. *Critical reviews in environmental science and technology*, 24(3):237–280, 1994.

A Mostafaie, Ehsan Forootan, A Safari et M Schumacher : Comparing multi-objective optimization techniques to calibrate a conceptual hydrological model using in situ runoff and daily GRACE data. *Computational Geosciences*, 22(3):789–814, 2018.

DF Motavita, R Chow, A Guthke et W Nowak : The comprehensive differential split-sample test: A stress-test for hydrological model robustness under climate variability. *Journal of Hydrology*, 573:501–515, 2019.

R Moussa, N Chahinian et C Bocquillon : Distributed hydrological modelling of a Mediterranean mountainous catchment–Model construction and multi-site validation. *Journal of Hydrology*, 337 (1-2):35–51, 2007.

MR NAJAFI, H MORADKHANI et IW JUNG : Assessing the uncertainties of hydrologic model selection in climate change impact studies. *Hydrological processes*, 25(18):2814–2826, 2011.

NO NASCIMENTO : *Appréciation à l'aide d'un modèle empirique des effets d'action anthropique sur la relation pluie-débit à l'échelle du bassin versant*. Thèse de doctorat, CERGRENE/ENPC, Paris, 550 pp., 1991.

JE NASH et JV SUTCLIFFE : River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3):282–290, 1970.

P NICOLLE, R PUSHPALATHA, C PERRIN, D FRANÇOIS, D THIÉRY, T MATHEVET, M LE LAY, F BESSON, JM SOUBEYROUX, C VIEL *et al.* : Benchmarking hydrological models for low-flow simulation and forecasting on French catchments. *Hydrology and Earth System Sciences*, 18:2829–2857, 2014.

J NOILHAN et J-F MAHFOUF : The ISBA land surface parameterisation scheme. *Global and planetary Change*, 13(1-4):145–159, 1996.

J NOILHAN et S PLANTON : A simple parameterization of land surface processes for meteorological models. *Monthly weather review*, 117(3):536–549, 1989.

T OKI et S KANAE : Global hydrological cycles and world water resources. *Science*, 313(5790): 1068–1072, 2006.

M OSUCH, RJ ROMANOWICZ et MJ BOOIJ : The influence of parametric uncertainty on the relationships between HBV model parameters and climatic characteristics. *Hydrological Sciences Journal*, 60(7-8):1299–1316, 2015.

L OUDIN : *Recherche d'un modèle d'évapotranspiration potentielle pertinent comme entrée d'un modèle pluie-débit global*. Thèse de doctorat, ENGREF Paris, 495 pp., 2004.

L OUDIN, V ANDRÉASSIAN, T MATHEVET, C PERRIN et C MICHEL : Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations. *Water Resources Research*, 42(7), 2006a.

L OUDIN, F HERVIEU, C MICHEL, C PERRIN, V ANDRÉASSIAN, F ANCTIL et C LOUMAGNE : Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. *Journal of Hydrology*, 303(1-4):290–306, 2005a.

L OUDIN, C MICHEL et F ANCTIL : Which potential evapotranspiration input for a lumped rainfall-runoff model?: Part 1—can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs? *Journal of Hydrology*, 303(1-4):275–289, 2005b.

L OUDIN, C PERRIN, T MATHEVET, V ANDRÉASSIAN et C MICHEL : Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models. *Journal of Hydrology*, 320(1-2):62–83, 2006b.

RK PACHAURI, MR ALLEN, VR BARROS, J BROOME, W CRAMER, R CHRIST, JA CHURCH, L CLARKE, Q DAHE, P DASGUPTA *et al.* : *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. IPCC, 2014.

TC PAGANO, A W WOOD, MH RAMOS, HL CLOKE, F PAPPENBERGER, MP CLARK, M CRANSTON, D KAVETSKI, T MATHEVET, S SOROOSHIAN *et al.* : Challenges of operational river forecasting. *Journal of Hydrometeorology*, 15(4):1692–1707, 2014.

V PARETO : *Manuel d'économie politique*. Giard, París, 703 pp., 1909.

S PATHIRAJA, D ANGHILERI, P BURLANDO, A SHARMA, L MARSHALL et H MORADKHANI : Time-varying parameter models for catchments with land use change: the importance of model structure. *Hydrology and Earth System Sciences*, 22(5):2903–2919, 2018.

MC PEEL : Hydrology: catchment vegetation and runoff. *Progress in Physical Geography*, 33 (6):837–844, 2009.

A PELLETIER et V ANDRÉASSIAN : Hydrograph separation: an impartial parametrisation for an imperfect method. *Hydrology and Earth System Sciences*, 24(3):1171–1187, 2020.

A PELLETIER, V ANDRÉASSIAN et O DELAIGUE : *baseflow : Computes Hydrograph Separation*, 2020. URL `https://cran.r-project.org/package=baseflow`. R package version 0.12.3.

HL PENMAN : Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 193(1032):120–145, 1948.

HL PENMAN : Weather, plant and soil factors in hydrology. *Weather*, 16(7):207–219, 1961.

C PERRIN : *Vers une amélioration d'un modèle global pluie-débit*. Thèse de doctorat, Institut National Polytechnique de Grenoble-INPG, 530 pp., 2000.

C PERRIN, V ANDRÉASSIAN, C ROJAS SERNA, T MATHEVET et N LE MOINE : Discrete parameterization of hydrological models: Evaluating the use of parameter sets libraries over 900 catchments. *Water Resources Research*, 44(8), 2008.

C PERRIN, C MICHEL et V ANDRÉASSIAN : Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, 279(1-4):275–289, 2003.

C PERRIN, L OUDIN, V ANDRÉASSIAN, C ROJAS-SERNA, C MICHEL et T MATHEVET : Impact of limited streamflow data on the efficiency and the parameters of rainfall-runoff models. *Hydrological Sciences Journal*, 52(1):131–151, 2007.

HO PÖRTNER, DC ROBERTS, V MASSON-DELMOTTE, P ZHAI, E POLOCZANSKA, K MINTENBECK, M TIGNOR, A ALEGRÍA, M NICOLAI, A OKEM, J PETZOLD, B RAMA et NM WEYER : IPCC, 2019: Technical Summary. *In* HO PÖRTNER, DC ROBERTS, V MASSON-DELMOTTE, P ZHAI, E POLOCZANSKA, K MINTENBECK, M TIGNOR, A ALEGRÍA, M NICOLAI, A OKEM, J PETZOLD, B RAMA et NM WEYER, éditeurs : *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*. In press, 2019.

WH PRESS, SA TEUKOLSKY, BP FLANNERY et WT VETTERLING : *Numerical recipes in Fortran 77: volume 1, volume 1 of Fortran numerical recipes: the art of scientific computing*. Cambridge university press, 1992.

CHB PRIESTLEY et RJ TAYLOR : On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, 100(2):81–92, 1972.

R PUSHPALATHA, C PERRIN, N LE MOINE et V ANDRÉASSIAN : A review of efficiency criteria suitable for evaluating low-flow simulations. *Journal of Hydrology*, 420:171–182, 2012.

R PUSHPALATHA, C PERRIN, N LE MOINE, T MATHEVET et V ANDRÉASSIAN : A downward structural sensitivity analysis of hydrological models to improve low-flow simulation. *Journal of Hydrology*, 411(1-2):66–76, 2011.

P REED et V DEVIREDDY : Groundwater Monitoring Design: A case study combining epsilon dominance archiving and automatic parameterization for the NSGA-II. *In Applications of multi-objective evolutionary algorithms*, pages 79–100. World Scientific, 2004.

JC REFSGAARD et J KNUDSEN : Operational validation and intercomparison of different types of hydrological models. *Water Resources Research*, 32(7):2189–2202, 1996.

JC REFSGAARD, H MADSEN, V ANDRÉASSIAN, K ARNBJERG-NIELSEN, TA DAVIDSON, M DREWS, DP HAMILTON, E JEPPESEN, E KJELLSTRÖM, JE OLESEN *et al.* : A framework for testing the ability of models to project climate change and its impacts. *Climatic change*, 122(1-2):271–282, 2014.

P REICHERT et M OMLIN : On the usefulness of overparameterized ecological models. *Ecological Modelling*, 95(2-3):289–299, 1997.

P RIBOUST, G THIREL, N LE MOINE et P RIBSTEIN : Revisiting a simple degree-day model for integrating satellite data: implementation of SWE-SCA hystereses. *Journal of Hydrology and Hydromechanics*, 67(1):70–81, 2019.

M ROCHE : *Hydrologie de surface*. Gauthier-Villars, Paris, ORSTOM, 430 pp., 1963.

P ROYER-GASPARD, V ANDRÉASSIAN et G THIREL : Technical note: PMR – a proxy metric to assess hydrological model robustness in a changing climate. *Hydrology and Earth System Sciences Discussions*, 2021:1–19, 2021.

M SAFT, MC PEEL, AW WESTERN, JM PERRAUD et Lu ZHANG : Bias in streamflow projections due to climate-induced shifts in catchment response. *Geophysical Research Letters*, 43(4):1574–1581, 2016a.

M SAFT, MC PEEL, AW WESTERN et L ZHANG : Predicting shifts in rainfall-runoff partitioning during multiyear drought: Roles of dry period and catchment characteristics. *Water Resources Research*, 52(12):9290–9305, 2016b.

L SAMANIEGO, R KUMAR et S ATTINGER : Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research*, 46(5), 2010.

A SANKARASUBRAMANIAN, Richard M VOGEL et James F LIMBRUNNER : Climate elasticity of streamflow in the United States. *Water Resources Research*, 37(6):1771–1781, 2001.

L SANTOS, G THIREL et C PERRIN : Continuous state-space representation of a bucket-type rainfall-runoff model: a case study with the GR4 model using state-space GR4 (version 1.0). 11(4):1591–1605, 2018a.

L SANTOS, G THIREL et C PERRIN : Pitfalls in using log-transformed flows within the KGE criterion. 22(8):4583–4591, 2018b.

B SCHAEFLI et HV GUPTA : Do Nash values have value? *Hydrological Processes*, 21(15):2075–2080, 2007.

B SCHAEFLI et E ZEHE : Hydrological model performance and parameter estimation in the wavelet-domain. *Hydrology and Earth System Sciences*, 13(10):1921–1936, 2009.

J SCHEWE, J HEINKE, D GERTEN, I HADDELAND, NW ARNELL, DB CLARK, R DANKERS, S EISNER, BM FEKETE, FJ COLÓN-GONZÁLEZ *et al.* : Multimodel assessment of water scarcity under climate change. *Proceedings of the National Academy of Sciences*, 111(9):3245–3250, 2014.

J SEIBERT : Reliability of model predictions outside calibration conditions: Paper presented at the Nordic Hydrological Conference (Røros, Norway 4-7 August 2002). *Hydrology Research*, 34 (5):477–492, 2003.

J SEIBERT et KJ BEVEN : Gauging the ungauged basin: how many discharge measurements are needed? *Hydrology and Earth System Sciences*, 13(6):883–892, 2009.

J SEIBERT, KH BISHOP et L NYBERG : A test of TOPMODEL's ability to predict spatially distributed groundwater levels. *Hydrological processes*, 11(9):1131–1144, 1997.

G SEILLER, F ANCTIL et C PERRIN : Multimodel evaluation of twenty lumped hydrological models under contrasted climate conditions. 16(4):1171–1189, 2012.

G SEILLER, F ANCTIL et R ROY : Design and experimentation of an empirical multistructure framework for accurate, sharp and reliable hydrological ensembles. *Journal of Hydrology*, 552:313–340, 2017.

M SHAFII et BA TOLSON : Optimizing hydrological consistency by incorporating hydrological signatures into model calibration objectives. *Water Resources Research*, 51(5):3796–3814, 2015.

E SHAMIR, B IMAM, HV GUPTA et S SOROOSHIAN : Application of temporal streamflow descriptors in hydrologic model parameter estimation. *Water Resources Research*, 41(6), 2005a.

Eylon SHAMIR, Bisher IMAM, Efrat MORIN, Hoshin V GUPTA et Soroosh SOROOSHIAN : The role of hydrograph indices in parameter estimation of rainfall–runoff models. *Hydrological Processes : An International Journal*, 19(11):2187–2207, 2005b.

WJ SHUTTLEWORTH : Evaporation. *In* DR MAIDMENT, éditeur : *Handbook of hydrology*. McGraw-Hill New York, 1993.

SK SINGH et A BÁRDOSSY : Calibration of hydrological models on hydrologically unusual events. *Advances in Water Resources*, 38:81–91, 2012.

S SOROOSHIAN et JA DRACUP : Stochastic parameter estimation procedures for hydrologie rainfall-runoff models: Correlated and heteroscedastic error cases. *Water Resources Research*, 16(2): 430–442, 1980.

CM STEPHENS, LA MARSHALL, FM JOHNSON, L LIN, LE BAND et H AJAMI : Is past variability a suitable proxy for future change? A virtual catchment experiment. *Water Resources Research*, 56(2):e2019WR026275, 2020.

S M STERLING, A DUCHARNE et J POLCHER : The impact of global land-cover change on the terrestrial water cycle. *Nature Climate Change*, 3(4):385–390, 2013.

KS TAN, FHS CHIEW, RB GRAYSON, PJ SCANLON et L SIRIWARDENA : Calibration of a daily rainfall-runoff model to estimate high daily flows. *In MODSIM 2005 International Congress on Modelling and Simulation, Melbourne*, 2005.

C TEUTSCHBEIN et J SEIBERT : Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of hydrology*, 456:12–29, 2012.

G THIREL, V ANDRÉASSIAN et C PERRIN : On the need to test hydrological models under changing conditions. *Hydrological Sciences Journal*, 60(7-8):1165–1173, 2015a.

G THIREL, V ANDRÉASSIAN, C PERRIN, J-N AUDOUY, L BERTHET, Pamela EDWARDS, N FOLTON, C FURUSHO, Adèle KUENTZ, J LERAT *et al.* : Hydrology under change: an evaluation protocol to investigate how hydrological models deal with changing catchments. *Hydrological Sciences Journal*, 60(7-8):1184–1199, 2015b.

CW THORNTHWAITE : An approach toward a rational classification of climate. *Geographical review*, 38(1):55–94, 1948.

M THYER, B RENARD, D KAVETSKI, G KUCZERA, SW FRANKS et S SRIKANTHAN : Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. *Water Resources Research*, 45(12), 2009.

KE TRENBERTH : The definition of El-Niño. *Bulletin of the American Meteorological Society*, 78 (12):2771–2778, 1997.

KE TRENBERTH : Conceptual framework for changes of extremes of the hydrological cycle with climate change. *In Weather and Climate Extremes*, pages 327–339. Springer, 1999.

PA TROCH, T LAHMERS, A MEIRA, R MUKHERJEE, JW PEDERSEN, T ROY et R VALDÉS-PINEDA : Catchment coevolution: A useful framework for improving predictions of hydrological change? *Water Resources Research*, 51(7):4903–4922, 2015.

PA TROCH, GF MARTINEZ, VRN PAUWELS, M DURCIK, M SIVAPALAN, C HARMAN, PD BROOKS, H GUPTA et T HUXMAN : Climate and vegetation water use efficiency at catchment scales. *Hydrological Processes*, 23(16):2409–2414, 2009.

AM UKKOLA, IC PRENTICE, TF KEENAN, AIJM VAN DIJK, NR VINEY, RB MYNENI et J BI : Reduced streamflow in water-stressed climates consistent with CO_2 effects on vegetation. *Nature Climate Change*, 6(1):75–78, 2016.

UNFCCC : Aggregate effect of the intended nationally determined contributions: an update. *In Synthesis Report by the Secretariat FCCC/CP/2016/2*, Marrakech, Maroc, 75 pp., 11 2016.

A VALÉRY : *Modélisation précipitations débit sous influence nivale : Élaboration d'un module neige et évaluation sur 380 bassins versants*. Thèse de doctorat, Institut des Sciences et Industries du Vivant et de l'Environnement, AgroParisTech, 417 pp., 2010.

A VALÉRY, V ANDRÉASSIAN et C PERRIN : 'As simple as possible but not simpler': What is useful in a temperature-based snow-accounting routine? Part 2–Sensitivity analysis of the Cemaneige snow accounting routine on 380 catchments. *Journal of hydrology*, 517:1176–1187, 2014.

CHM VAN BAVEL : Potential evaporation: the combination concept and its experimental verification. *Water Resources Research*, 2(3):455–467, 1966.

W VAN ESSE, C PERRIN, M BOOIJ, D AUGUSTIJN, F FENICIA, D KAVETSKI et F LOBLIGEOIS : The influence of conceptual model structure on model performance: a comparative study for 237 French catchments. *Hydrology and Earth System Sciences*, 17(10):4227–4239, 2013.

J VAZE, FHS CHIEW, JM PERRAUD, N VINEY, D POST, J TENG, B WANG, J LERAT et M GOSWAMI : Rainfall-runoff modelling across southeast Australia: datasets, models and results. *Australasian Journal of Water Resources*, 14(2):101–116, 2011.

J VAZE, DA POST, FHS CHIEW, JM PERRAUD, NR VINEY et J TENG : Climate non-stationarity–validity of calibrated rainfall–runoff models for use in climate change studies. *Journal of Hydrology*, 394(3-4):447–457, 2010.

JP VIDAL, B HINGRAY, C MAGAND, E SAUQUET et A DUCHARNE : Hierarchy of climate and hydrological uncertainties in transient low-flow projections. *Hydrology and Earth System Sciences*, 20 (9):3651–3672, 2016.

JP VIDAL, E MARTIN, L FRANCHISTÉGUY, M BAILLON et JM SOUBEYROUX : A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *International Journal of Climatology*, 30(11):1627–1644, 2010.

JA VRUGT, HV GUPTA, SC DEKKER, S SOROOSHIAN, T WAGENER et W BOUTEN : Application of stochastic parameter optimization to the Sacramento Soil Moisture Accounting model. *Journal of Hydrology*, 325(1-4):288–307, 2006.

JA VRUGT et BA ROBINSON : Improved evolutionary optimization from genetically adaptive multimethod search. *Proceedings of the National Academy of Sciences*, 104(3):708–711, 2007.

JA VRUGT et M SADEGH : Toward diagnostic model calibration and evaluation: Approximate Bayesian computation. *Water Resources Research*, 49(7):4335–4345, 2013.

Frank W : Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1(6):80–83, 1945. ISSN 00994987.

T WAGENER, N MCINTYRE, MJ LEES, HS WHEATER et HV GUPTA : Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis. *Hydrological processes*, 17 (2):455–476, 2003.

T WAGENER, M SIVAPALAN, PA TROCH, BL MCGLYNN, CJ HARMAN, HV GUPTA, P KUMAR, PSC RAO, NB BASU et JS WILSON : The future of hydrology: An evolving science for a changing world. *Water Resources Research*, 46(5), 2010.

K WANG et RE DICKINSON : A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability. *Reviews of Geophysics*, 50(2), 2012.

QJ WANG : The genetic algorithm and its application to calibrating conceptual rainfall-runoff models. *Water Resources Research*, 27(9):2467–2471, 1991.

S WESTRA, M THYER, M LEONARD, D KAVETSKI et M LAMBERT : A strategy for diagnosing and interpreting hydrological model nonstationarity. *Water Resources Research*, 50(6):5090–5113, 2014.

HS WHEATER, AJ JAKEMAN et KJ BEVEN : Progress and directions in rainfall-runoff modelling – Chapter 5. *In* A J JAKEMAN, M B BECK et M J MCALEER, éditeurs : *Modelling Change in Environmental Systems*, pages 101–132. John Wiley and Sons Ltd, 1993.

RL WILBY : Uncertainty in water resource model parameters used for climate change impact assessment. *Hydrological Processes : An International Journal*, 19(16):3201–3219, 2005.

RL WILBY et I HARRIS : A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK. *Water resources research*, 42(2), 2006.

AP WILLIAMS, CD ALLEN, AK MACALADY, D GRIFFIN, CA WOODHOUSE, DM MEKO, TW SWETNAM, SA RAUSCHER, R SEAGER, HD GRISSINO-MAYER *et al.* : Temperature as a potent driver of regional forest drought stress and tree mortality. *Nature climate change*, 3(3):292–297, 2013.

CJ WILLMOTT, SM ROBESON et K MATSUURA : A refined index of model performance. *International Journal of Climatology*, 32(13):2088–2094, 2012.

HC WINSEMIUS, B SCHAEFLI, A MONTANARI et HHG SAVENIJE : On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information. *Water Resources Research*, 45(12), 2009.

WMO : *Guide to Hydrological Practices*, volume 1. World Meteorological Organization, 296 pp., Genève, Suisse, 1996.

T WÖHLING, L SAMANIEGO et R KUMAR : Evaluating multiple performance criteria to calibrate the distributed hydrological model of the upper Neckar catchment. *Environmental Earth Sciences*, 69 (2):453–468, 2013.

KK YILMAZ, HV GUPTA et T WAGENER : A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. *Water Resources Research*, 44(9), 2008.

L ZHANG, N POTTER, K HICKEL, Y ZHANG et Q SHAO : Water balance modeling over variable time scales based on the Budyko framework–Model development and testing. *Journal of Hydrology*, 360(1-4):117–131, 2008.

Annexes

Annexe A.
Suppléments à la note technique

A.1. Characterisation of model bias across DSST setups

Model biases in the DSST have been calculated in an absolute way in the Results section so that they could be compared with PMR values. This resulted in a loss of information about the sign of model errors. In this appendix, it is shown how the sign of these errors relates to the different DSST experiments. The biases obtained for GR4J for each of the six types of DSST setup are plotted in Figure A.1 without taking their absolute values.
Model bias follows different patterns depending on the climatic variable used to define the calibration and evaluation periods of the DSST. While the distribution of model errors seems comparatively even for periods characterized by different average precipitation, transferring model parameters between periods with different runoff ratios clearly triggers opposite model bias, whether the transfer is performed in one way or in another. For most catchments, GR4J indeed underestimates streamflow volumes when runoff ratio increases and, conversely, overestimates streamflow volumes when runoff ratio decreases. DSSTs based on temperature yield situations in between, since median model bias is slightly negative (respectively positive) when calibrated on warmer (respectively colder) periods. When calculated in absolute terms, model bias was larger in DSSTs based on runoff ratio than for experiments based on temperature and precipitation (Figure A.1). Therefore, robustness issues for the model appear to be caused less by changes in climatic changes than by modification of the catchment response to precipitation. This result is in line with the conclusion of Saft *et al.* (2016a), who tested a number of hydrological models in southeastern Australia during prolonged droughts. The authors observed that many of these models would produce biased simulations of streamflow during the drought if, and only if, the catchments had experienced shifts in the rainfall-runoff relationship from pre-drought to drought conditions. Our results extend this statement for GR4J to situations where runoff ratio increases and shows opposite model biases depending on the sign of the change.

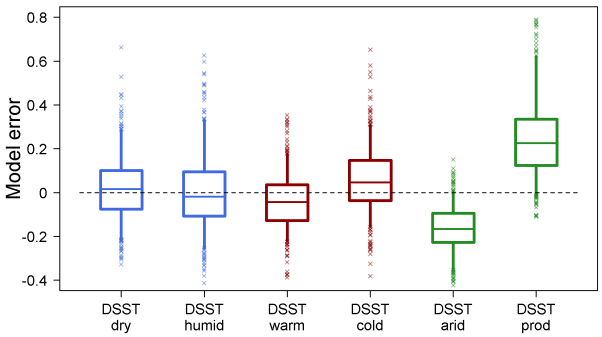


FIGURE A.1 : Distribution of model biases in DSST for each type of setup. The boxplots represent the 5, 25, 50, 75 and 95 quantiles and the little crosses denote the outliers. Blue, red and green boxplots are respectively associated to DSST setups based on precipitation, temperature and runoff ratio.

A.2. The choice of an adequate mathematical expression

The mathematical expression of the PMR also results from a choice that needs to be discussed. For example, Coron *et al.* (2014) proposed computing the flatness of the moving bias curve as the

standard deviation of model bias on the sub-periods.

$$Alt\ PMR = 2 \times \frac{1}{N} \times \sum_{i=1}^{N} \left(\frac{\bar{Q}_{s,i}}{\bar{Q}_{o,i}} - \frac{\bar{Q}_s}{\bar{Q}_o} \right)^2 \tag{A.1}$$

We discussed the mathematical form chosen for the PMR by comparing the metrics defined in Equation (4.1) and Equation (A.1). Figure B.1 shows the differences between the metrics in Pearson's correlation with model biases obtained in the DSST performed on 5-year periods. The length of the sub-periods used in the PMR varies from 1 to 10 years.

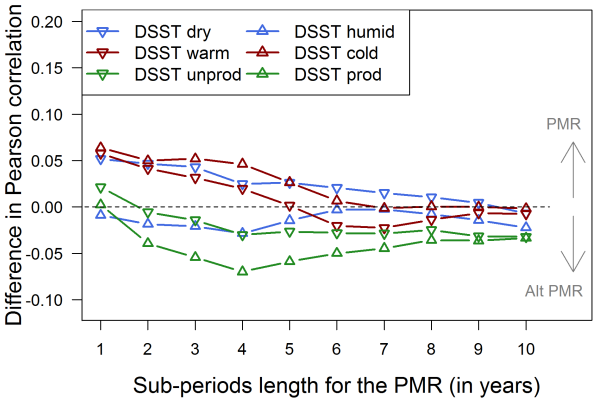


FIGURE A.2 : Differences between the correlations of PMR and alternative PMR with DSST biases (performed on 6-year calibration and evaluation periods), depending on the DSST experiment and on the length of the sub-periods used for the computation of the PMR. Values above zero indicate better agreement between PMR and DSST results compared to alternative PMR, and vice versa.

It appears that short sub-periods confer greater benefit to the reliability of the PMR (Equation (4.1)), whereas longer sub-periods benefit the alternative PMR (Equation (A.1)). Choosing a 5-year sub-period for the computation of the PMR does not, on average, favor either the one or the other formulation of the metric. As mentioned previously, we sought to formulate the PMR so that it maximizes correlation with DSST biases while enhancing the sensitivity of the metric. For this reason, better agreement for the PMR as formulated in Equation (4.1) with DSST biases obtained for shorter sub-periods, where the PMR is most sensitive to model robustness, makes it more suitable. Therefore, the PMR computed as the sum of absolute average model error on 5-year sub-periods is best suited to evaluating model robustness.

In addition, we note that the alternative PMR corresponds better overall to DSST experiments based on runoff ratio, which, we remind the reader, yielded the larger model biases. The fact that model biases are squared in the computation of the alternative PMR enhances the weight of sub-periods where the model simulations are the worst, and thus potentially where the observed runoff ratio differs the most from the average. It is possible that calculating the error differences in absolute terms rather than squared makes the metric less dependent on years when the model is drastically worse or on years with large measurement errors.

Furthermore, the PMR as defined in Equation (4.1) may have other advantages compared to Equation (A.1). The fact that model error on each sub-period is normalized by the average observed streamflow during the total period instead of the average observed streamflow during the sub-period may put less emphasis on very dry years when observed streamflow is close to zero. Model bias on such dry years can be undesirably large, thus the PMR as defined in Equation (4.1) could be a better option for arid catchments. This also makes it possible to compute the PMR in catchments where rivers might eventually cease to flow for long periods of time without any further adjustments to the data. In addition, the interpretation of the PMR is perceived as more straightforward in Equation (4.1) than in Equation (A.1), as model error is simply compared to the observed streamflow averaged on

the whole time series rather than to a quantity that varies across the sub-periods.

A.3. Reliability of the metric for different DSST sub-period lengths

Our results on the sensitivity of the PMR to the length of the sub-periods on which it is computed suggest that reducing the length of the sub-periods involved in the computation of the PMR might slightly reduce loss of information about model robustness, and thus that selecting 1-year sub-periods is the best option. Computed this way, PMR would represent average annual model bias. Since in DSST experiments model bias is usually computed on the whole evaluation period, i.e., on periods that may vary in length from 1 to many years, it is unclear whether the metric would be representative of model biases as computed on periods longer than 1 year.
To evaluate the representativeness of PMR in such conditions, we computed the correlation between PMR values for different sub-period lengths of the moving bias curve, as in Figure 4.6, and also for different sub-period lengths for the DSST experiments. Note that, as previously, PMR was computed on the whole time series after calibrating the model on the whole time series. The heat map of correlations between PMR values and average DSST biases for varying sub-period lengths is displayed on Figure C.1.

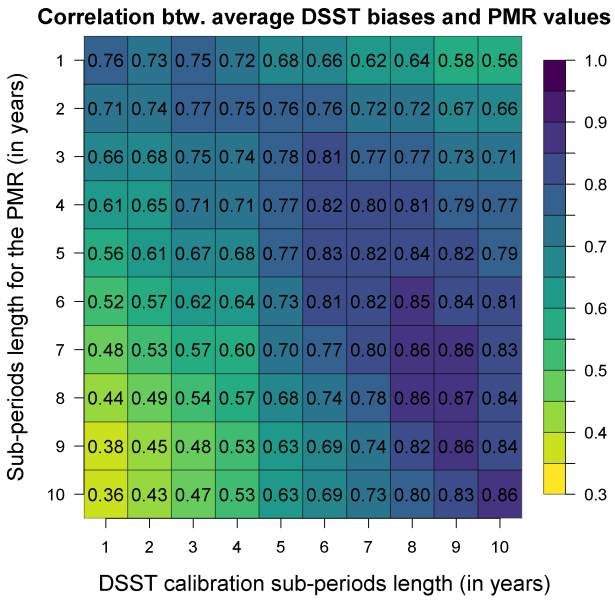


FIGURE A.3 : Pearson's correlations between the PMR computed on varying sub-period lengths (vertical axis) and the average DSST biases obtained on varying sub-period lengths for calibration and evaluation (horizontal axis).

The heat map clearly shows that generally shorter sub-periods for the computation of the PMR relate better to shorter calibration periods in the DSST experiments. Conversely, longer sub-periods for the computation of the PMR relate better to longer calibration periods in the DSST experiments. This result is not surprising given that PMR computed on n-year sub-periods represents the average model bias as computed on n years, and therefore should show similar patterns to model biases computed in DSST experiments involving n-year periods.
However, some sub-period lengths for the PMR computation exhibit a high correlation with a wider

range of DSST setups. By computing row-wise averages in the matrix, we observed that PMR computations based on 3- to 5-year sub-periods reach an average correlation of 0.74 with DSST biases across the range of sub-period lengths. In comparison, the correlation coefficient of PMR values computed on 1-year sub-periods is on average 0.67. Therefore, defining sub-periods with lengths between 3 and 5 years may be the most suitable choice to ensure PMR representativeness across a wide spectrum of possible DSST experiments.

Annexe B.

Discussions sur la définition des gammes de débit

B.1. Introduction

Pour mieux identifier les lacunes structurelles des modèles hydrologiques avec la méthode présentée au Chapitre 7, nous avons cherché à définir des gammes de débit robustes et représentatives autant que possible des divers régimes hydrologiques des bassins de l'échantillon. Évidemment, ces régimes ne sont en réalité pas divisés en trois catégories de débits bien définies, et de ce fait aucun cloisonnement précis ne saurait être pleinement satisfaisant dans l'absolu. C'est pourquoi après avoir décrit ce que représentent différentes définitions de gammes de débit du point de vue des hydrogrammes, nous nous sommes attachés à les évaluer dans la perspective de mieux diagnostiquer les modèles hydrologiques.

Pour comparer entre elles les différentes gammes de débit, nous sommes partis d'un constat simple : si les critères introduits dans le Chapitre 7 doivent permettre d'évaluer la robustesse et la polyvalence d'un jeu de paramètres donné, on peut également retourner le problème en évaluant la robustesse du jeu de paramètres obtenu en optimisant ces critères. La pertinence du jeu de paramètres identifié comme le plus polyvalent au sens des biais et des PMR sur ces gammes est de fait questionnable. Si un tel jeu de paramètres permet de maximiser des scores de performance du modèle jugés comme importants, il n'est pas certain que cet optimum soit également satisfaisant par rapport à des critères d'évaluation plus classiques. Il n'est pas certain non plus qu'il soit mieux transférable que des jeux de paramètres identifiés par d'autres méthodes. À ce titre, il pourrait ainsi être intéressant de comparer l'optimum ainsi obtenu avec celui obtenu avec le KGE racine, identifié comme la fonction objectif de référence dans le Chapitre 5.
Puisqu'il n'est pas n'est pas possible de diviser objectivement un hydrogramme en trois gammes de débits, nous avons pris le parti de juger de la qualité de la définition de ces gammes à l'aune des résultats obtenus en optimisant les critères de biais et de PMR qui leur sont associés.

B.2. Problématique : définition d'une gamme de débit

B.2.1. Gammes définies à partir de parts fixes des pas de temps

Une manière simple et intuitive pour diviser un hydrogramme en trois parties consiste à raisonner sur la distribution des débits observés, à partir de quantiles des débits choisis à l'avance. C'est la méthode proposée par Yilmaz *et al.* (2008) pour un bassin versant en Oklahoma, aux États-Unis. En procédant ainsi, une gamme de débit est construite sur une base fixe en nombre de pas de temps : par exemple les 20% des débits les plus faibles peuvent constituer la gamme des « basses eaux », les 20% des débits les plus forts celle des « hautes eaux », et les 60% restants appartiennent à la gamme des eaux intermédiaires. Considérons que les seuils de 20% pour la définition des deux gammes extrêmes soient ceux retenus pour définir trois gammes de débits pour tous les bassins versants de notre échantillon. Ce seuil permet d'utiliser suffisamment de données dans les gammes pour intégrer de nombreux événements différents et est *a priori* suffisamment bas pour accentuer les différences entre les gammes.

Afin de se représenter ce à quoi correspondent des gammes débit ainsi définies, nous avons étudié les cas des bassins versants de l'Adour à l'Aire-sur-Adour (Q1100010), dans les Pyrénées françaises, et de la Castle Creek River à Arcadia (AU405246), en Australie. Les parties de la courbe de régime correspondants aux gammes des basses eaux, des eaux intermédiaires, et des hautes

eaux sont représentées en Figure B.1.

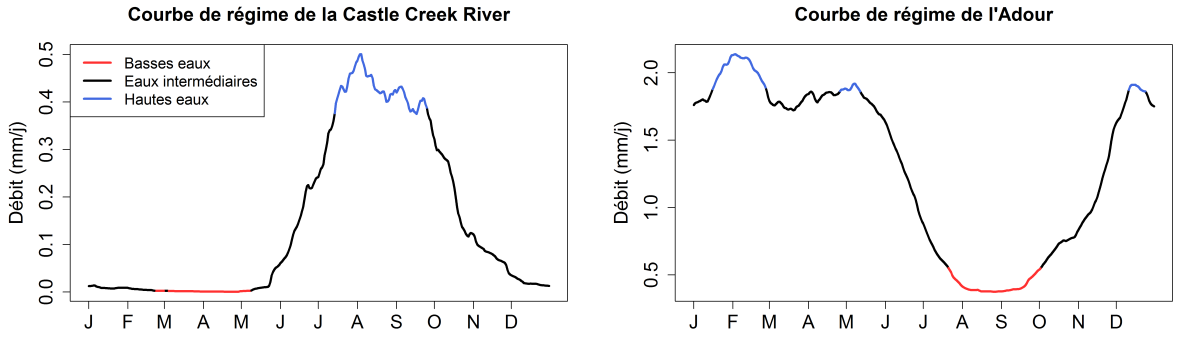


FIGURE B.1 : Courbes de régime des bassins de la Castle Creek River at Arcadia (à gauche) et de l'Adour à l'Aire-sur-l'Adour (à droite). Les parties rouges et bleues des courbes illustrent les gammes de débit, respectivement des basses eaux et des hautes eaux, qu'on obtiendrait si on les définissait à partir de la courbe de régime du bassin.

La définition des gammes de débit considérée semble bien convenir au régime hydrologique du bassin de l'Adour. Pour le bassin de la Castle Creek River en revanche, une grande partie de l'hydrogramme est constituée de débits d'étiage. La gamme des basses eaux choisie représentant ni plus ni moins que 20% des pas de temps où le débit est le plus bas, elle semble définie sur une portion trop faible des écoulements. De ce point de vue, la gamme des eaux intermédiaires définie entre les hautes eaux et les basses eaux contient des écoulements qui pourraient être considérés comme appartenant plutôt à la gamme des basses eaux.

B.2.2. Interprétation des problèmes soulevés par les parts fixes des pas de temps

Pour éviter un tel écueil, il semble bien plus avantageux de considérer une part des pas de temps définissant chaque gamme adaptée à chacun des bassins versants de l'échantillon. La difficulté d'une telle approche réside dans la manière dont cette part est choisie. Si des analyses visuelles peuvent être une bonne solution pour des échantillons de quelques bassins, l'application de ce procédé à de grands échantillons prendrait beaucoup de temps et serait fortement subjective. Une autre approche consiste à distinguer les gammes de débits à partir de leur contribution respective au volume écoulé total.
Cette approche peut être expliquée par une analogie à la richesse entre les membres d'une société. Il est fréquent de lire dans les médias que quelques pourcents des plus riches possèdent près de la totalité des richesses de la planète. Cette observation peut être visualiée au moyen de la courbe de Gini. Cette courbe permet de mesurer graphiquement la richesse totale cumulée par X% de la population, du plus pauvre au plus riche (en partant de la gauche sur l'axe des abscisses) ou du plus riche au plus pauvre (en partant de la droite sur l'axe des abscisses). Inversement, elle permet de mesurer la part de la population détenant X% de la richesse totale cumulée en retrouvant les abscisses à partir des ordonnées. Transposée dans les conditions de notre étude, la « courbe de Gini hydrologique » montre le volume total cumulé de l'écoulement en fonction de la part des pas de temps prise en compte, du pas de temps où le débit est le plus faible (le plus « pauvre » en eau) jusqu'au pas de temps où le débit est le plus fort (le plus « riche »). À partir de cette courbe, il est donc possible d'étudier l'inégalité de répartition du volume écoulé entre les gammes de débit. Selon les définitions précédentes, la gamme des basses eaux agrège les 20% des pas de temps

les plus pauvres, la gamme des hautes eaux agrège les 20% des pas de temps les plus riches, et la gamme intermédiaire agrège le reste.
L'exemple de la Castle Creek River illustre un cas de répartition très inégale du volume écoulé dans le temps : il faudrait retenir environ 80% des pas de temps pour définir une gamme des basses eaux contribuant autant aux débits cumulés totaux que la gamme des basses eaux de l'Adour (Figure B.2). A l'opposé de la courbe, une très grande partie du volume cumulé de la Castle Creek River s'écoule sur une fraction très faible des pas de temps. En définissant la gamme des hautes eaux comme la gamme agrégeant 20% des pas de temps les plus riches en eau, celle-ci amalgamerait la quasi-totalité du volume total écoulé dans la gamme des hautes eaux.

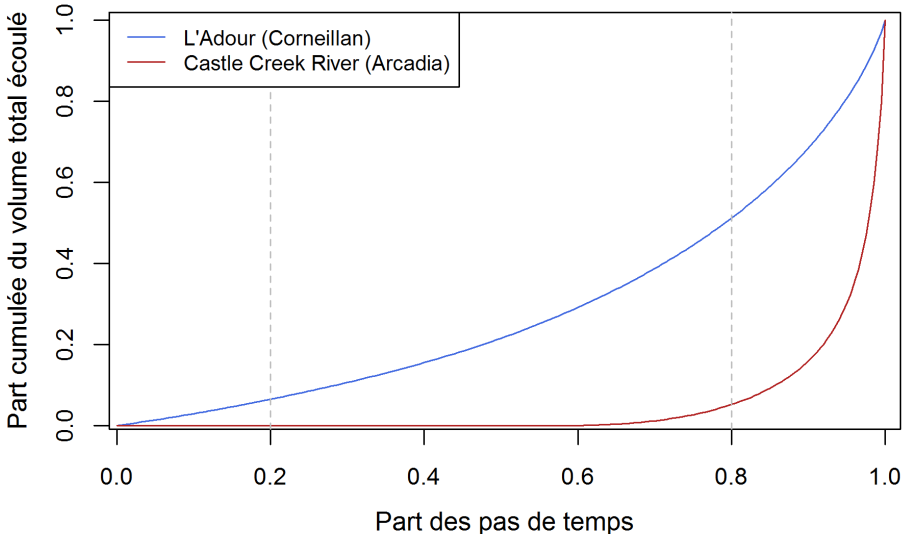


FIGURE B.2 : Cumul des débits en part du volume total en fonction du pourcentage des pas de temps cumulés pour les bassins de l'Adour à Corneillan (bleu) et de la Castle Creek River à Arcadia (rouge).

Il convient de noter qu'une répartition parfaitement équitable correspond à une courbe de Gini en forme de droite. En hydrologie, cette situation est purement théorique. L'Adour est par exemple un des bassins les plus « égalitaires » de l'échantillon, et pourtant il faut comptabiliser près de 70% des pas de temps où le débit est le plus faible pour atteindre 33% du volume total écoulé.
Ainsi, définies sur une part fixe des pas de temps, les gammes de débits peuvent contenir des parts du volume cumulé total très variables d'un bassin à l'autre (Figure B.3). Le flou de cette définition des gammes de débit amène parfois la gamme des hautes eaux à contenir la quasi-totalité du volume écoulé, ne laissant presque rien pour les deux gammes inférieures. Dans de plus rares cas, la gamme intermédiaire peut cumuler une plus grande part du volume écoulé que la gamme des hautes eaux. Enfin, dans un peu plus d'un quart des bassins versants de l'échantillon, la gamme des basses eaux représente moins de 1% du volume cumulé total.

Au regard des travaux du Chapitre 7, ces variations dans les poids respectifs des gammes de débits peuvent poser un problème pour définir des mesures de l'erreur des modèles hydrologiques. En effet, comparer les volumes moyens observés et simulés d'une même gamme de débit n'a pas la même signification sur les performances du modèle selon que cette gamme concentre la quasi-totalité ou au contraire une part très faible du volume écoulé cumulé. Calculer le biais du modèle sur la gamme des hautes eaux de la Castle Creek River revient par exemple presque à calculer simplement le biais du modèle, puisque les erreurs sur les deux gammes de débits inférieures doivent être colossales pour compter dans le calcul du biais du modèle sur l'ensemble des débits. Calculer le biais sur une gamme qui concentre une part très faible du volume total écoulé revient au contraire à donner une importance considérable à des erreurs du modèle ne pesant que très faiblement sur

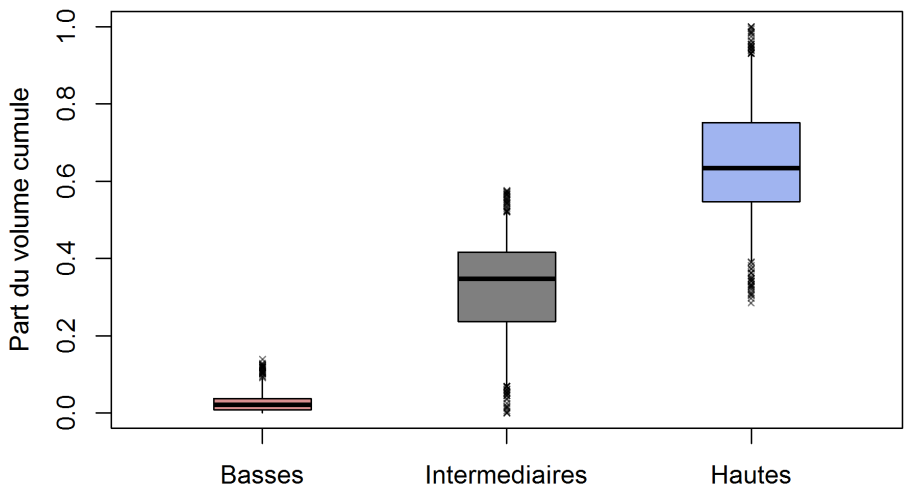


FIGURE B.3 : Distribution des parts du volume écoulé total représentées par les différentes gammes de débit dans l'échantillon d'étude, si ces gammes sont définies avec une part fixe des pas de temps (20% des débits les plus faibles pour la gamme des basses eaux, 20% des débits les plus forts pour la gamme des hautes eaux).

son bilan. Ainsi, les significations et comportements très différents de l'erreur du modèle sur une même gamme de débit d'un bassin versant à l'autre diminuent la lisibilité des résultats.

B.2.3. Gammes définies à partir de part fixes du débit cumulé

L'analyse précédente sur les poids respectifs des gammes de débits en matière de volume cumulé suggère de repenser la définition de ces gammes à partir de ce critère. Ainsi redéfinies, les gammes ne s'étendraient plus sur une part fixe des pas de temps mais sur une part fixe du volume cumulé total. Cette approche présente l'avantage d'affiner la perception des capacités du modèle à simuler le bilan en eau d'un bassin dans différentes conditions d'écoulement. Ce choix permet aussi de mener cette analyse quel que soit le régime du bassin versant modélisé. En revanche, cela peut conduire à définir une seule gamme de débits sur la quasi-totalité des débits si leur distribution est très asymétrique, ou au contraire sur quelques pas de temps seulement, ce qui rend le calcul des erreurs du modèle davantage dépendant des erreurs de mesures.
Pour définir ces gammes de débit de façon à respecter à peu près la typologie des bassins de l'échantillon d'étude et à diminuer l'effet des contreparties précédemment évoquées, nous avons étudié la relation entre débits cumulés et pas de temps concernés. La Figure B.4 montre les « courbes de Gini hydrologiques » de 11 bassins qui correspondent aux quantiles 0, 10, 20,... et 100 de l'inégalité de répartition des volumes écoulés.

La diversité des distributions témoigne de la difficulté de définir des gammes de débit universelles. Fixer la part des pas de temps revient à laisser exister une grande diversité de répartition des volumes entre gammes (courbes verticales pointillées). Inversement, fixer la part du volume cumulé revient à laisser exister une grande diversité de répartition des pas de temps inclus dans chaque gamme (courbes horizontales pointillées). Pour proposer des définitions adéquates pour une majorité de bassins versants de l'échantillon, il peut être avantageux de raisonner avec les courbes les plus centrales, qui représentent la médiane des comportements des bassins versants. La courbe de répartition médiane et ses voisines croisent l'abscisse 50% à peu près à l'ordonnée 10%, et l'abscisse 90% à l'ordonnée 50%.

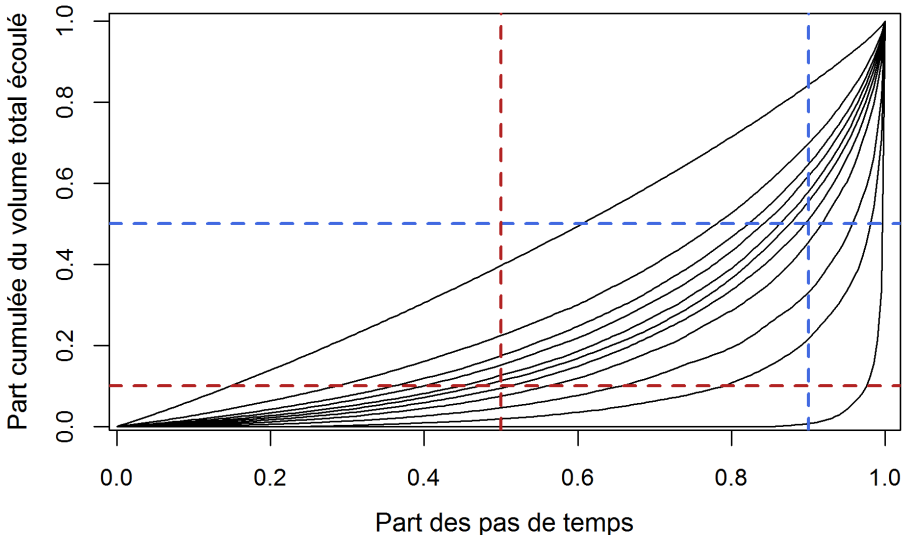


FIGURE B.4 : Cumul des débits en part du volume total en fonction du pourcentage des pas de temps cumulés, pour 11 bassins représentatifs des différences d'inégalités de répartition à travers l'échantillon d'étude. Le trait pointillé bleu vertical (respectivement horizontal) indique la part cumulé du volume total (resp. des pas de temps) obtenu pour les différents bassins en définissant la gamme des hautes eaux comme 10% des débits les plus élevés (resp. comme les débits les plus élevés cumulant 50% du volume total écoulé). Les traits pointillés rouges représentent la même chose pour deux définitions de la gamme des basses eaux (50% des plus bas débits ou débits les plus bas dont le cumul représente 10% de l'écoulement total).

En considérant que les basses eaux puissent être définies comme l'ensemble des débits les plus faibles contribuant à 10% de l'écoulement cumulé total, la gamme ainsi définie pourrait représenter à la fois les faibles écoulements des bassins arides, distribués sur une grande part des pas de temps, et les bassins où les débits sont distribués de façon plus symétrique. Le choix des 10% du volume cumulé permet également de considérer une gamme dont la contribution au volume total est suffisamment faible pour caractériser des débits d'étiage, mais suffisamment haute pour que les erreurs du modèle sur cette gamme s'en ressentent sur le volume total simulé et contribuent au biais du modèle.
De même que les basses eaux, les débits de crue correspondent à des parts de l'écoulement très différentes d'un bassin à l'autre. Dans certains cas, ils peuvent représenter une part très importante du débit cumulé en peu de pas de temps. De ce fait, une gamme des hautes eaux définie à partir d'une part fixe élevée du débit cumulé total peut correspondre à n'inclure qu'une part très faible des pas de temps. Cette situation n'est pas idéale dans le mesure où les erreurs de mesure auraient alors un poids non négligeable dans la caractérisation de la gamme. À l'inverse, augmenter la part fixe du débit cumulé dans la définition de la gamme des hautes eaux afin d'augmenter la part des pas de temps inclus peut conduire à diminuer fortement la part du débit cumulé restante pour définir les gammes inférieures. De ce fait, définir les hautes eaux comme celles concentrant la moitié de l'écoulement cumulé semble être un arbitrage équilibré entre d'une part, la part des pas de temps concernés, et d'autre part la part du débit cumulé dans le bilan en eau.

Pour illustrer cet arbitrage, nous proposons d'adopter une définition de la gamme des hautes eaux comme les débits les plus forts contribuant à 50% du débit cumulé total, une définition des basses eaux comme les débits les plus faibles contribuant à 10% de ce total, et de définir la gamme des eaux intermédiaires sur les débits restants, qui contribuent à 40% du débit cumulé total. La Figure B.5 montre le pourcentage de pas de temps contenant dans les trois gammes de débits ainsi définies. Comme attendu, il arrive que les gammes soient représentées de façon très déséquilibrée dans le temps, les hautes eaux sur une part très faible des pas de temps (0.5% au minimum) et les basses eaux sur la quasi-totalité (97.5% au maximum).

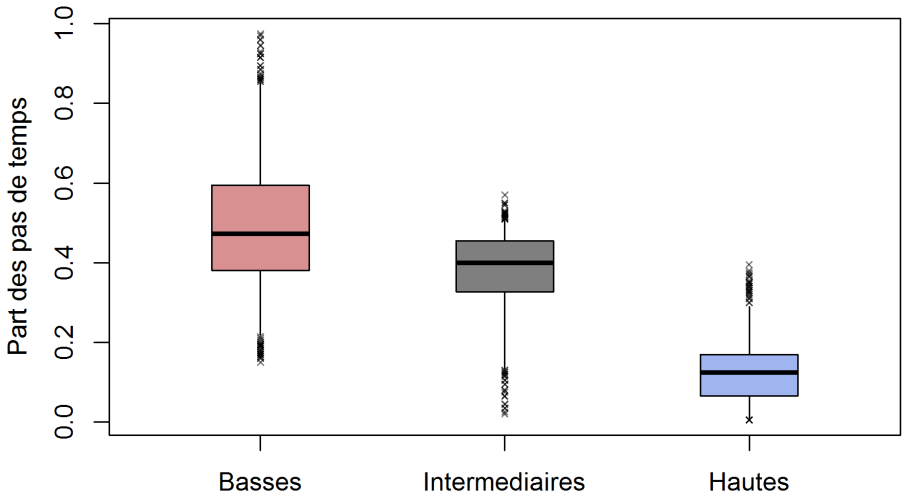


FIGURE B.5 : Distribution des parts des pas de temps inclus dans chaque gamme de débit dans l'échantillon d'étude, si ces gammes sont définies avec une part fixe des contributions au volume écoulé total (débits les plus faibles accumulant 10% du volume total pour la gamme des basses eaux, débits les plus forts accumulant 50% du volume total pour la gamme des hautes eaux).

B.2.4. Objectifs

Les définitions proposées des gammes de débit, qu'elles s'appuient sur une part fixe des pas de temps considérés pour chaque gamme ou sur une part fixe de la contribution au volume écoulé total, souffrent donc toutes de défauts entre lesquels il n'est pas possible d'arbitrer définitivement. Puisqu'il n'est pas possible de juger absolument de la pertinence des multiples définitions des gammes de débit évoquées précédemment à partir des hydrogrammes, nous proposons de le faire grâce à une approche originale fondée sur les performances des modèles hydrologiques. Dans la suite du chapitre, nous avons évalué plusieurs fonctions objectif construites à partir des critères du Chapitre 7 relatifs aux gammes de débit. Pour cela, nous avons supposé que la robustesse du modèle GR4J calé avec les différentes fonctions objectif permettrait de distinguer les définitions pertinentes des gammes de débit des moins pertinentes, au regard de la modélisation hydrologique.

Les objectifs poursuivis dans cette annexe sont donc les suivants :

- Comparer différentes définitions des gammes de débits à partir de l'évaluation de fonctions objectif intégrant des critères relatifs à ces gammes ;
- Comparer les performances du modèle calé avec ces fonctions objectif avec celles obtenues avec le KGE racine.

B.3. Méthodologie

B.3.1. Définitions des gammes de débit et des fonctions objectif associées

Pour une définition générique des gammes de débit, nous introduisons une fonction objectif CGD, pour *Critère sur les Gammes de Débit*, définie de la manière suivante :

$$CGD = 1 - \sqrt{\begin{array}{l}(r\,[Q]-1)^2 + \left(\frac{\mu_{B,s}\,[Q]}{\mu_{B,o}\,[Q]}-1\right)^2 + \left(\frac{\mu_{I,s}\,[Q]}{\mu_{I,o}\,[Q]}-1\right)^2 + \left(\frac{\mu_{H,s}\,[Q]}{\mu_{H,o}\,[Q]}-1\right)^2 \\ \\ + (PMR_B\,[Q])^2 + (PMR_I\,[Q])^2 + (PMR_H\,[Q])^2\end{array}}$$

Les termes $\mu_{B,o}$, $\mu_{I,o}$ et $\mu_{H,o}$ correspondent au débit observé moyen sur les gammes des basses eaux, des eaux intermédiaires et des hautes eaux. Les termes $\mu_{B,s}$, $\mu_{I,s}$ et $\mu_{H,s}$ correspondent au débit simulé moyen sur les gammes des basses eaux, des eaux intermédiaires et des hautes eaux. Enfin les termes PMR_B, PMR_I et PMR_H représentent les critères PMR calculés sur chacune des gammes de débit. La construction des CGD s'appuie sur l'agrégation de plusieurs sous-critères par un sous forme de distance euclidienne, à l'instar du KGE, dont seul le terme de corrélation r a été repris afin de contraindre les dynamiques infra-annuelles simulées par le modèle.

À l'aune des discussions de la section précédente, plusieurs définitions des gammes de débits ont été proposées pour être évaluées à travers les CGD, et répertoriées dans le Tableau B.1.

TABLEAU B.1 : Récapitulatif des fonctions objectif testés et des définitions des gammes associées. Le terme « pdt » correspond à « pas de temps ».

Nom du critère	Basses eaux	Eaux intermédiaires	Hautes eaux
CGD1	20% des pdt où le débit est le plus faible	60% des pdt restants	20% des pdt où le débit est le plus fort
CGD2	50% des pdt où le débit est le plus faible	40% des pdt restants	10% des pdt où le débit est le plus fort
CGD3	débits les plus faibles dont le cumul vaut 10% du volume total	débits restants dont le cumul vaut 40% du volume total	débits les plus forts dont le cumul vaut 50% du volume total

B.3.2. Méthode d'évaluation des fonctions objectif fondées sur les différentes définitions des gammes de débit

La comparaison des fonctions objectif CGD s'appuie sur des expériences de calage-contrôle du modèle GR4J similaires à celles conduites au Chapitre 8.
Une première étape a consisté en une évaluation succincte des performances de GR4J en calage

sur la période totale disponible de chaque bassin. Dans un second temps, ses performances ont été évaluées en contrôle. Pour chaque bassin, six périodes de calage ont été constituées à partir des six années les plus pluvieuses, des six années les moins pluvieuses, des six années les plus chaudes, des six années les plus froides, des six années les plus productives (au sens du rendement moyen des précipitations), et enfin des six années les moins productives. Pour chacune des expériences réalisées, les six années sélectionnées ne sont pas nécessairement consécutives. Les performances des modèles ont été systématiquement évaluées sur l'ensemble de la chronique disponible. Il convient de noter que les périodes de calage et la période de contrôle ne sont pas strictement indépendantes, les premières étant inclues dans la deuxième.

Le choix des précipitations et de la température pour la définition des périodes de calage permet d'inscrire ce travail dans la continuité des études de robustesse relevées dans la littérature scientifique. Celui du rendement moyen des précipitations répond à l'exigence de tester les modèles dans des conditions très défavorables, tel que cela a été démontré aux Chapitres 4 et 5.
Le choix de constituer des périodes de calage de six ans a été déterminé d'une part afin de limiter le chevauchement des périodes de calage et de la période de contrôle, et d'autre part pour éviter que certains paramètres ne soient sous-contraints par une quantité insuffisante de données en calage.

Afin de comparer les performances du modèle calé avec les fonctions objectif CGD à une référence, un calage des paramètres de GR4J a été réalisé avec le KGE racine, dont nous avons démontré la pertinence au Chapitre 5 pour maximiser la robustesse des modèles hydrologiques. L'évaluation des performances a d'abord été réalisée en calage sur la période totale par rapport au KGE racine. L'évaluation en contrôle est basée sur la comparaison des valeurs de KGE d'une part, et de NSE calculés sur les débits transformés avec plusieurs transformations d'autre part. Les transformations utilisées pour les calculs de NSE sont les fonctions inverse, logarithme, racine carrée, identité (pas de transformation) et carrée. Ces transformations permettent d'évaluer le modèle sur des parties variées de l'hydrogramme.

B.4. Comparaison des fonctions objectif fondées sur les gammes de débit

B.4.1. Performances en calage

Dans la perspective de comparaison des performances obtenues avec les différentes fonctions objectif, la Figure B.6 montre les valeurs de KGE racine obtenues en calage avec chacun des trois critères, comparées à celles obtenues en calage avec le KGE racine.
Les performances générales obtenues en calage avec les trois fonctions objectif sont logiquement moins bonnes que celles obtenues avec le KGE racine. Parmi les fonctions objectif testées, les moins bons résultats sont donnés par la fonction objectif CGD1 (KGE racine moyen de 0,869). Les deux autres montrent des performances assez similaires (respectivement 0,891 et 0,893 en moyenne pour les critères CGD2 et CGD3). Cette plus grande proximité avec les valeurs obtenues en calant directement avec le KGE racine est susceptible de signaler une meilleure robustesse des paramètres que ceux obtenus en calant avec le critère CGD1, puisque les paramètres issus du calage du KGE racine sont souvent robustes. Toutefois, l'existence probable de jeux de paramètres équifinaux rend cette hypothèse incertaine.

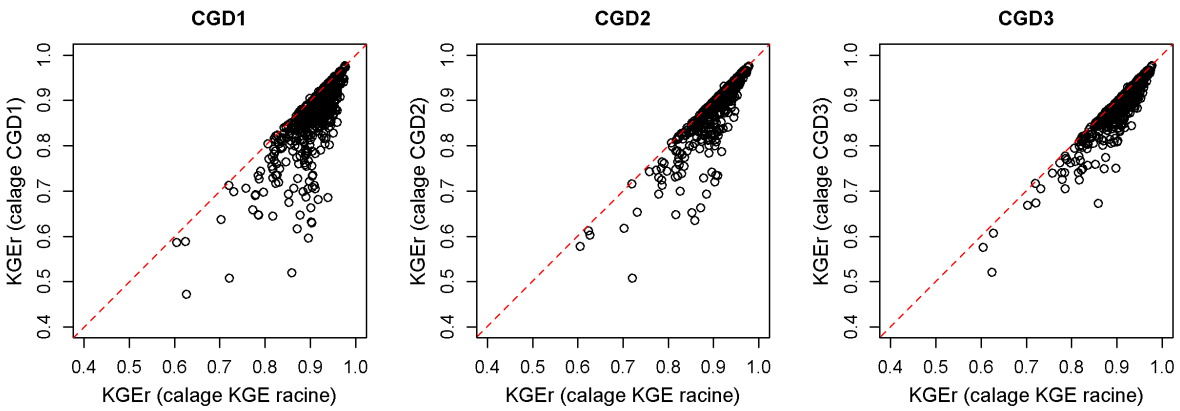


FIGURE B.6 : Comparaison des KGE racine obtenus pour le modèle GR4J en calage avec le KGE racine (abscisses) et avec les trois critères CGD1, CGD2 puis CGD3 (ordonnées). Chaque figure mélange les valeurs obtenues pour les bassins français et australiens.

B.4.2. Performances en contrôle

B.4.3. Comparaison des valeurs de KGE en contrôle

Les valeurs en KGE atteintes en contrôle par GR4J calé avec les fonctions objectif CGD sont comparées à celles obtenues avec le KGE racine en Figure B.7, en distinguant les sous-échantillons de bassins français et australien.

Dans l'ensemble, les fonctions objectif ne se distinguent pas nettement entre elles. Les performances les plus dégradées par rapport à celles obtenues avec le KGE racine sont plutôt obtenues avec le CGD1 (KGE moyens de respectivement 0,782 et 0,578 en France et en Australie). Les écarts entre les performances obtenues avec les deux autres critères sont, comme précédemment en calage, plus difficiles à observer. En France, le CGD2 obtient les valeurs les plus hautes en moyenne (0,819, contre 0,806 avec le CGD3). En Australie, les rôles sont néanmoins inversés (0,618 pour le CGD2, contre 0,653 pour le CGD3).
Interpréter les différences entre CGD2 et CGD3 n'est pas vraiment possible au vu de ces résultats. Il est intéressant de constater en revanche que le CGD2 fait en moyenne toujours mieux que le CGD1. Ceci tend à indiquer la moindre pertinence de choisir des parts égales des pas de temps pour définir les gammes des basses eaux et des hautes eaux (i.e. respectivement 20% des débits les plus faibles et 20% des débits les plus forts). Un découpage asymétrique comme celui porté par le CGD2, prend vraisemblablement mieux en considération la répartition temporelle des débits. Les étiages étant généralement des événements aux dynamiques davantage mensuelles que journalières, au contraire des crues, définir une gamme des basses eaux sur une part plus large de l'hydrogramme semble mieux adapté à la diversité de notre échantillon.

B.4.4. Comparaison des valeurs de NSE calculé sur les débits transformés en contrôle

Les distributions des NSE, calculés sur les débits transformés, atteints en contrôle par GR4J calé avec les fonctions objectif CGD et le KGE racine sont présentés en Figure B.8, en distinguant les

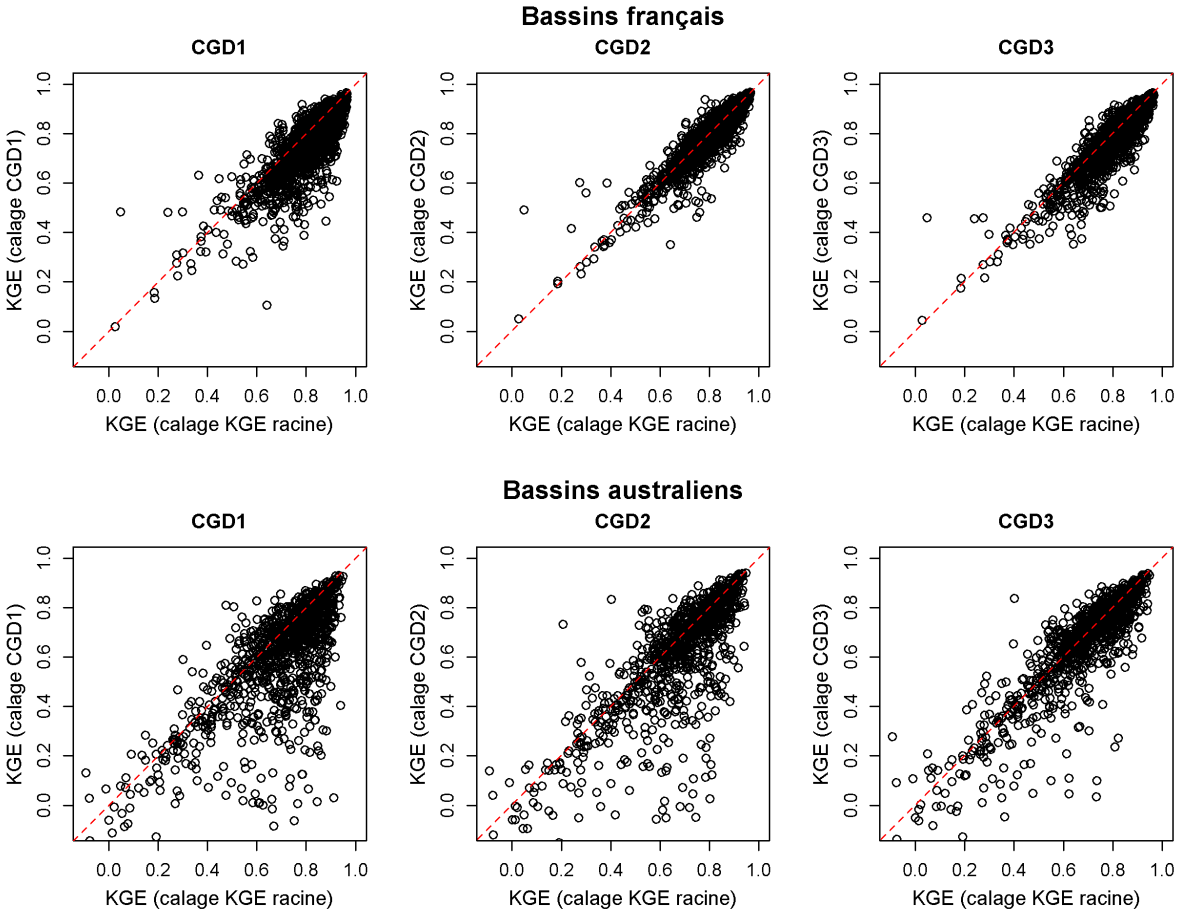


FIGURE B.7 : Comparaison des KGE obtenus pour le modèle GR4J en contrôle avec le KGE racine (abscisses) et avec les trois critères CGD1, CGD2 puis CGD3 (ordonnées), sur les sous-échantillons de bassins français et australiens.

sous-échantillons de bassins français et australien.

Il convient de noter en premier lieu que le KGE racine s'impose nettement comme une fonction objectif plus adéquate que les autres dans l'identification de paramètres robustes, malgré des résultats en étiage moins bons que ceux obtenus avec la fonction CGD1. En ce qui concerne la comparaison des trois fonctions objectif CGD entre elles, les performances en contrôle s'accordent assez bien avec les constats formulés précédemment. Sur les bassins australiens, le critère CGD3 produit une meilleure représentation des débits intermédiaires et de crue, très légèrement aux dépens de celle des étiages. Au contraire, le critère CGD1 est moins adapté sur l'ensemble des gammes de débit excepté les forts étiages (NSE calculé sur l'inverse des débits) quel que soit l'échantillon de bassins considéré. Le critère CGD2 est moins bon que le critère CGD3 sur les bassins australiens, et démontre des performances comparables sur les bassins français, légèrement meilleures au regard des NSE calculés sur les débits non transformés et sur les carrés des débits.

Le critère CGD3 utilisé en calage de ces modèles produit des simulations assez équilibrées entre représentation des basses eaux, des eaux intermédiaires et des hautes eaux, tant pour les débits des bassins australiens que français. Le choix de délimiter la gamme des basses eaux en garantissant qu'elle représente une part faible mais néanmoins non négligeable (10%) du volume total écoulé, ainsi que la limitation de la gamme des hautes eaux à la moitié de ce volume écoulé,

est peut-être une des raisons de cet équilibre. L'erreur calculée sur les basses eaux représente en effet une part des éventuels problèmes de bilan du modèle, la réduire contribue donc à l'ajustement de ce bilan, et l'optimisation de l'erreur du modèle calculée sur les hautes eaux ne se fait pas trop aux dépens de celle des autres gammes de débit. Ces considérations nous amènent donc à privilégier la définition du critère CGD3, fondé sur les volumes écoulés cumulés, pour distinguer les différentes gammes de débit est la plus appropriée pour évaluer la robustesse de GR4J.

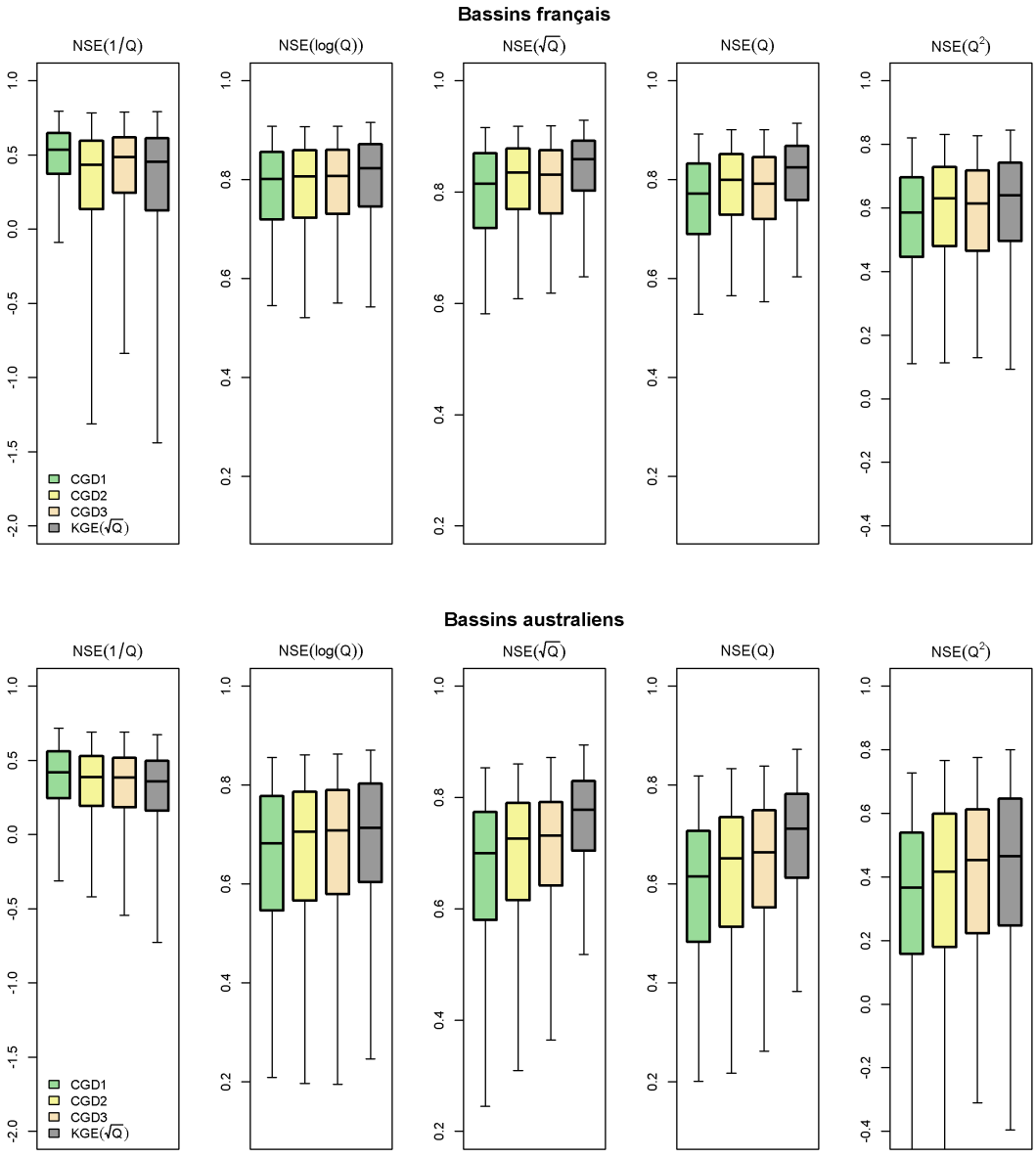


FIGURE B.8 : Distribution des valeurs de NSE calculés sur les débits transformés pour le modèle GR4J en contrôle sur les bassins français et australiens avec les trois fonctions objectif fondées sur les différentes définitions des gammes de débit (CGD1 en vert, CGD2 en jaune, CGD3 en rose pastel), et avec le KGE racine (gris).

Les qualités évoquées du critère CGD3 ne suffisent toutefois pas à en faire une fonction objectif aussi satisfaisante que le KGE racine. Cela est peut-être dû au fait d'avoir retiré le terme de variabilité du kge dans le calcul des critères CGD. Par ailleurs, la prise en considération de sept sous-critères revient à calculer une distance à l'optimum dans un espace à sept dimensions. Ré-

duire à un scalaire les performances du modèle dans un espace si grand résulte vraisemblablement à une perte trop importante d'information pour un calage robuste des paramètres d'un modèle. Le KGE racine reste donc le critère de référence pour l'optimisation des paramètres de GR4J.

B.5. Synthèse

Dans cette annexe, nous avons proposé plusieurs alternatives à la définition de gammes de débit pour la conception de critères de biais et de robustesse utilisés dans le Chapitre 7.
À partir de discussions sur une définition des gammes de débit fondée sur une part fixe des pas de temps (par exemple, un gamme des basses eaux définie sur les 20% des débits les plus faibles), nous avons montré que ce choix pouvait ne pas être pertinent sur tous les bassins de l'échantillon à la fois. En reconnaissant qu'une part variable des pas de temps puisse être nécessaire, nous avons discuté une alternative consistant à définir les gammes de débits à partir de contributions fixes au volume écoulé total. En procédant ainsi, la part des pas de temps inclus dans chaque gamme varie d'un bassin à l'autre, mais conduit de ce fait à définir certaines gammes sur une des parts très faibles des débits pour certains bassins. Aucune définition ne semble donc satisfaisante d'un point de vue théorique.
Pour arbitrer entre ces définitions, nous avons choisi de comparer la robustesse des jeux de paramètres de GR4J obtenus en optimisant des fonctions objectif construites à partir de plusieurs définitions alternatives. L'objectif était de sélectionner la plus adaptée à l'évaluation des modèles hydrologiques, et éventuellement de comparer ces fonctions objectif au KGE racine, fonction objectif établie comme celle garantissant la meilleure transférabilité des paramètres des modèles conceptuels (Chapitre 5).

Trois définitions des gammes de débit ont été proposées : deux définitions fondées sur des parts fixes des pas de temps (définitions 1 et 2), et une définition fondée sur des parts fixes des contributions au volume écoulé total (définition 3). La base commune des fonctions objectif testées pour chacune de ces trois définitions des gammes de débit a consisté en une adaptation du KGE, dont les sous-critères étaient les biais et les PMR du modèle sur les trois gammes de débit. Les trois fonctions objectif (CGD1, CGD2 et CGD3) ainsi créées ont été comparées dans le cadre d'une expérience de calage-contrôle du modèle GR4J similaire à celle implémentée dans le Chapitre 8.

Nos résultats montrent que, mis à part en ce qui concerne les débits très faibles, le critère CGD3 utilisé comme fonction objectif permet d'obtenir les paramètres les plus robustes. Il semble donc que définir les gammes de débit sur la base de part fixe du volume total au lieu de part fixe des pas de temps soit une meilleure approche pour le calage de GR4J. L'adaptabilité de cette définition à des hydrogrammes aux caractéristiques statistiques variées explique peut-être les meilleures performances obtenues avec le critère CGD3, notamment parce qu'elle évite de calculer des erreurs relatives à de très faibles débits, dont la contribution au volume écoulé total est infime, comme cela peut-être le cas si les basses eaux sont définies sur une part fixe des débits les plus faibles. Au contraire, l'optimisation du biais et de la robustesse sur une gamme participant à 10% des écoulements cumulés évite cet écueil. C'est donc la définition des gammes de débit qui a été retenue pour les concevoir les critères d'évaluation du Chapitre 7.
Néanmoins, aucune des fonctions objectif testées n'est aussi pertinente que le KGE racine pour le calage de GR4J, ces fonctions objectif n'ont donc pas été réutilisées à d'autres fins que la comparaison des différentes définitions des gammes de débit réalisée dans cette annexe.

Annexe C.

Implémentation d'une fonction objectif incluant explicitement la robustesse

C.1. Introduction

La comparaison de plusieurs définitions des gammes de débit réalisée en Annexe B s'est appuyée sur la comparaison de fonctions objectifs pour le calage de GR4J. Celles-ci ont été construites *ad hoc* à partir des critères utilisés au Chapitre 7. Les résultats obtenus, s'ils ont permis de distinguer une définition des gammes de débit plus robuste que les autres, ont montré que les fonctions objectif construites n'étaient pas du tout aussi fiables que le KGE racine pour le calage de GR4J.
Dans cette annexe, nous proposons d'étudier une implémentation simple du PMR dans le KGE racine, sans distinction des gammes de débit. Le but de cette étude est de vérifier les possibilités de contraindre explicitement les paramètres des modèles hydrologiques vers plus de robustesse, celle-ci étant matérialisée par le PMR, pour l'aplatissement de la courbe des biais glissants. Cette vérification est nécessaire dans la mesure où il est possible que l'intégration du PMR, qui est censée améliorer la robustesse en « interpolation », conduise au contraire à détériorer les performances des modèles hydrologiques telles que calculées par d'autres critères. Pour évaluer correctement l'intégration du PMR dans une fonction objectif adaptée du KGE racine, nous avons donc envisagé plusieurs façons de calculer le PMR et comparé plusieurs fonctions objectif candidates pour caler le modèle GR4J

C.2. Méthodologie

C.2.1. Description des fonctions objectif évaluées

Nous proposons d'intégrer simplement le PMR comme une quatrième composante au KGE racine. La fonction objectif obtenue est appelée RKGE, pour *Robust KGE*. Le calcul du RKGE sur la racine carrée des débits s'écrit donc ainsi :

$$RKGE\left[\sqrt{Q}\right] = 1 - \sqrt{\begin{aligned}&\left(r\left[\sqrt{Q}\right]-1\right)^2 + \left(\frac{\sigma_s\left[\sqrt{Q}\right]}{\sigma_o\left[\sqrt{Q}\right]}-1\right)^2 + \left(\frac{\mu_s\left[\sqrt{Q}\right]}{\mu_o\left[\sqrt{Q}\right]}-1\right)^2 \\ &+ \left(PMR\left[\sqrt{Q}\right]\right)^2\end{aligned}} \tag{C.1}$$

$$\text{où } PMR\left[\sqrt{Q}\right] = \frac{1}{N}\frac{1}{\mu_o\left[\sqrt{Q}\right]}\sum_{i=1}^{N}\left|\left(\mu_{s,i}\left[\sqrt{Q}\right]-\mu_{o,i}\left[\sqrt{Q}\right]\right)-\left(\mu_s\left[\sqrt{Q}\right]-\mu_o\left[\sqrt{Q}\right]\right)\right| \tag{C.2}$$

Pour rappel, $r\left[\sqrt{Q}\right]$ correspond à la corrélation entre les racines carrées des débits simulés et observés, $\sigma_s\left[\sqrt{Q}\right]$ et $\sigma_o\left[\sqrt{Q}\right]$ sont les écarts-type des racines carrées des débits simulés et observés, et $\mu_s\left[\sqrt{Q}\right]$ et $\mu_o\left[\sqrt{Q}\right]$ sont les moyennes des racines carrées des débits simulés et observés. Les termes $\mu_{s,i}\left[\sqrt{Q}\right]$ et $\mu_{o,i}\left[\sqrt{Q}\right]$ désignent les moyennes des racines carrées des débits simulés et observés sur la sous-période i. Le PMR$\left[\sqrt{Q}\right]$ correspond à l'aire sous la courbe des biais glissants de la racine carrée des débits.

Plusieurs manières de calculer le PMR ont été envisagées, en parallèle de la forme utilisée dans le reste de la thèse. Dans le Chapitre 4, nous avons par exemple montré que le choix de longueur des périodes de calage avait une influence sur les valeurs de PMR. Il était notamment apparu que le calcul du PMR sur des sous-périodes de d'un an augmentait la sensibilité du critère aux erreurs

des modèles hydrologiques. Par ailleurs, dans l'Annexe A, nous avons analysé les différences entre le calcul du PMR tel que présenté au Chapitre 4 et le calcul de l'écart-type des biais glissants tel qu'utilisé par Coron (2013). Une des différences entre ces critères réside notamment dans le calcul de la moyenne des écarts des biais sur les sous-périodes par rapport à la moyenne des biais, calculé comme la moyenne des écarts absolus dans le cas du PMR, et calculé comme la moyenne quadratique des écarts dans le cas des travaux de Coron (2013). L'utilisation de la moyenne quadratique conduit notamment à donner plus de poids à certaines années où les écarts sont les plus grands, c'est-à-dire où les erreurs du modèle sont les plus différentes de ses erreurs « typiques ».

À partir de ces considérations, nous avons proposé quatre façons de calculer le PMR. Nous avons d'abord considéré le calcul du PMR sur des sous-périodes de 5 ans (comme dans toute la thèse) ou sur des sous-périodes d'un an. Nous avons ensuite envisagé de calculer la moyenne des écarts absolus ou la moyenne quadratique des écarts. Les quatre fonctions objectif RKGE racine définies par ces quatre façons de calculer le PMR sont répertoriées dans le Tableau C.1.

TABLEAU C.1 : Définitions des quatre fonctions objectif de type RKGE comparées.

Nom	Caractéristiques du calcul du PMR		
	Longueur des sous-périodes	Moyenne des écarts	Transformation des débits
RKGE5.1$[\sqrt{Q}]$	5 ans	écarts absolus	racine carrée
RKGE5.2$[\sqrt{Q}]$	5 ans	quadratique	racine carrée
RKGE1.1$[\sqrt{Q}]$	1 an	écarts absolus	racine carrée
RKGE1.2$[\sqrt{Q}]$	1 an	quadratique	racine carrée

C.2.2. Méthode d'évaluation des fonctions objectif

La comparaison des fonctions objectif n'a été réalisée qu'en calage. Les modèles ont été calés sur la totalité des chroniques disponibles des bassins de l'échantillon et leurs performances évaluées sur cette même période. Les résultats obtenus sont donc représentatifs des performances du modèle en simulation sur une longue période susceptible de contenir des conditions climatiques variables. Le KGE racine est utilisé comme fonction objectif de référence, que les fonctions objectif RKGE sont censées dépasser.

Le PMR a été utilisé dans la suite comme critère d'évaluation des simulations, notamment pour différencier les performances sur les gammes des basses eaux, des eaux intermédiaires et des hautes eaux (voir Chapitre 7). Il convient de noter que le PMR a été calculé dans ce cadre comme dans tout le reste de la thèse, c'est-à-dire sur des sous-périodes de 5 ans, en considérant la moyenne des écarts absolus.

C.3. Résultats

Afin d'avoir un aperçu des comportements de chacune des fonctions objectif testées, nous avons d'abord comparé les valeurs de KGE racine atteintes par GR4J en calage avec les fonctions objectif RKGE racine par rapport aux valeurs optimales atteintes en calant avec le KGE racine (Figure C.1). Les performances obtenus pour les sous-échantillons de bassins versants français et australien ont été mélangés pour plus de concision.

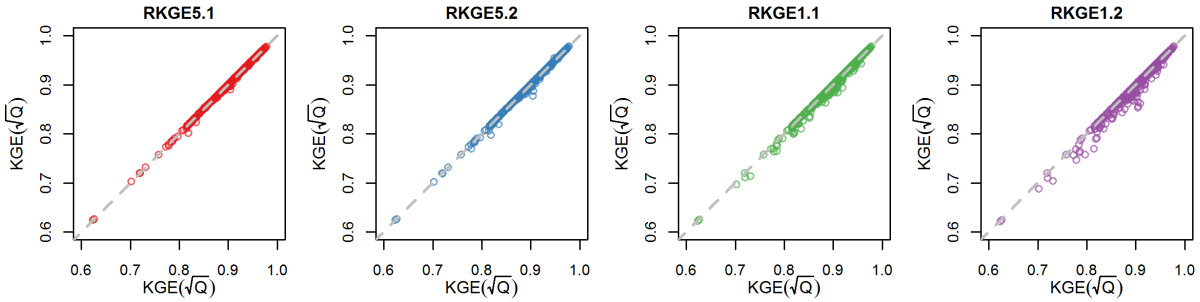


FIGURE C.1 : Comparaison des valeurs de KGE racine de GR4J en calage obtenues en calant avec les fonctions objectif RKGE5.1 (rouge), RKGE5.2 (bleu), RKGE1.1 (vert) et RKGE1.2 (violet) calculées sur la racine carrée des débits par rapport aux valeurs obtenues en calant avec le KGE racine. Les sous-échantillons de bassins français et australien ont été mélangés.

Tandis que les valeurs de KGE racine atteintes en calage avec le RKGE5.1 et RKGE5.2 racine sont sensiblement les mêmes que celles atteintes en calage avec le KGE racine, des différences sont davantage visibles avec les RKGE1.1 et RKGE1.2 racine. Pour quelques dizaines bassins, les valeurs de KGE racine sont en effet nettement inférieures avec ces fonctions objectif. Ceci s'explique par la plus grande sensibilité du PMR calculé sur des sous-périodes d'un an aux erreurs du modèle, et aux valeurs supérieures que le critère atteint dans ce cas-là en comparaison du PMR calculé sur des sous-périodes de 5 ans. Cette sensibilité accrue est susceptible de donner un rôle plus important à l'optimisation du PMR parmi les composantes du RKGE racine et donc d'éloigner davantage le comportement des RKGE1.1 et RKGE1.2 racine de celui du KGE racine.

Les biais de GR4J en calage sont montrés en Figure C.2. De manière générale, le modèle est moins biaisé dans la gamme des basses eaux lorsque ses paramètres sont calés avec les RKGE racine, quels qu'ils soient, plutôt qu'avec le KGE racine, qui est associé à des simulations légèrement sous-estimée des débits de la gamme. À l'opposé, les simulations de la gamme des hautes eaux semblent moins sous-estimées en calage avec le KGE racine qu'avec les RKGE racine. Ce constat est d'autant plus vrai pour les RKGE1.1 et RKGE1.2 racine. Pour la gamme des eaux intermédiaires, les meilleurs résultats sont obtenus avec le RKGE5.1 racine.

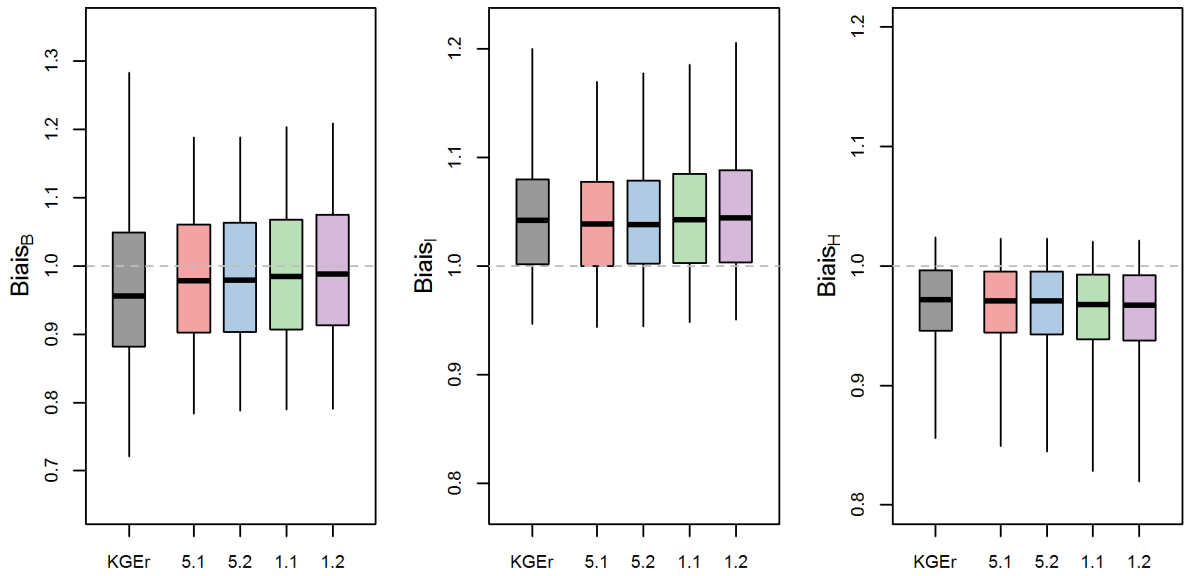


FIGURE C.2 : Distributions des biais de GR4J calculés sur les gammes de débit en calage en calant les fonctions objectif RKGE5.1 (rouge), RKGE5.2 (bleu), RKGE1.1 (vert) et RKGE1.2 (violet) calculées sur la racine carrée des débits et le KGE racine (gris).

Les constats précédents sont assez similaires pour les valeurs de PMR de chaque gamme de débit (Figure C.3). Il semble que conformément aux résultats obtenus en Figure C.1, les différences entre les performances associées au KGE racine et au RKGE racine sont d'autant plus grandes que le PMR est calculé dans le RKGE de sorte à le rendre plus sensible : les différences sont plus marquées entre le KGE racine et le RKGE1.2 racine qu'entre le KGE racine et le RKGE5.1 racine. Enfin, il convient de noter que les différences dans les distributions se situent en particulier pour les valeurs extrêmes, c'est-à-dire pour les bassins où le modèle est le plus biaisé (ou le moins robuste au sens du PMR).

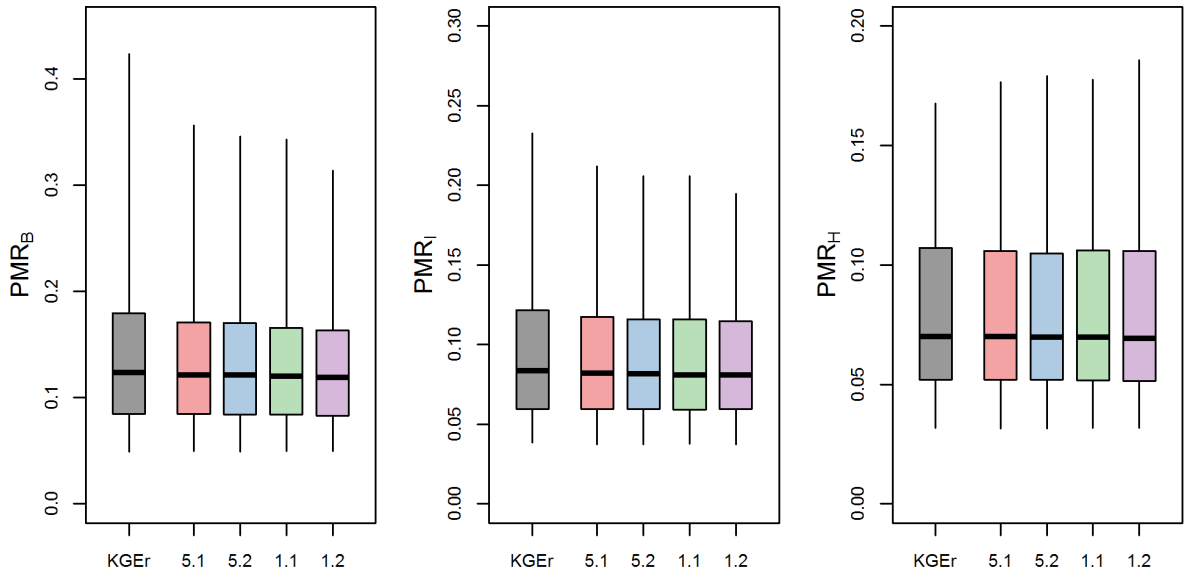


FIGURE C.3 : Distributions des valeurs de PMR de GR4J calculés sur les gammes de débit en calage en calant les fonctions objectif RKGE5.1 (rouge), RKGE5.2 (bleu), RKGE1.1 (vert) et RKGE1.2 (violet) calculées sur la racine carrée des débits et le KGE racine (gris).

Afin d'envisager les simulations des différentes gammes de débit sous un autre angle, nous avons également représenté les distributions des valeurs du critère de Nash-Sutcliffe calculé sur les débits transformés par cinq transformations différentes (fonctions inverse, logarithme, racine carrée, identité, carrée). Les résultats obtenus sur les sous-échantillons de bassin français et australien ont cette fois été séparés, afin de déterminer leurs contributions respectives aux écarts constatés précédemment (Figure C.4).

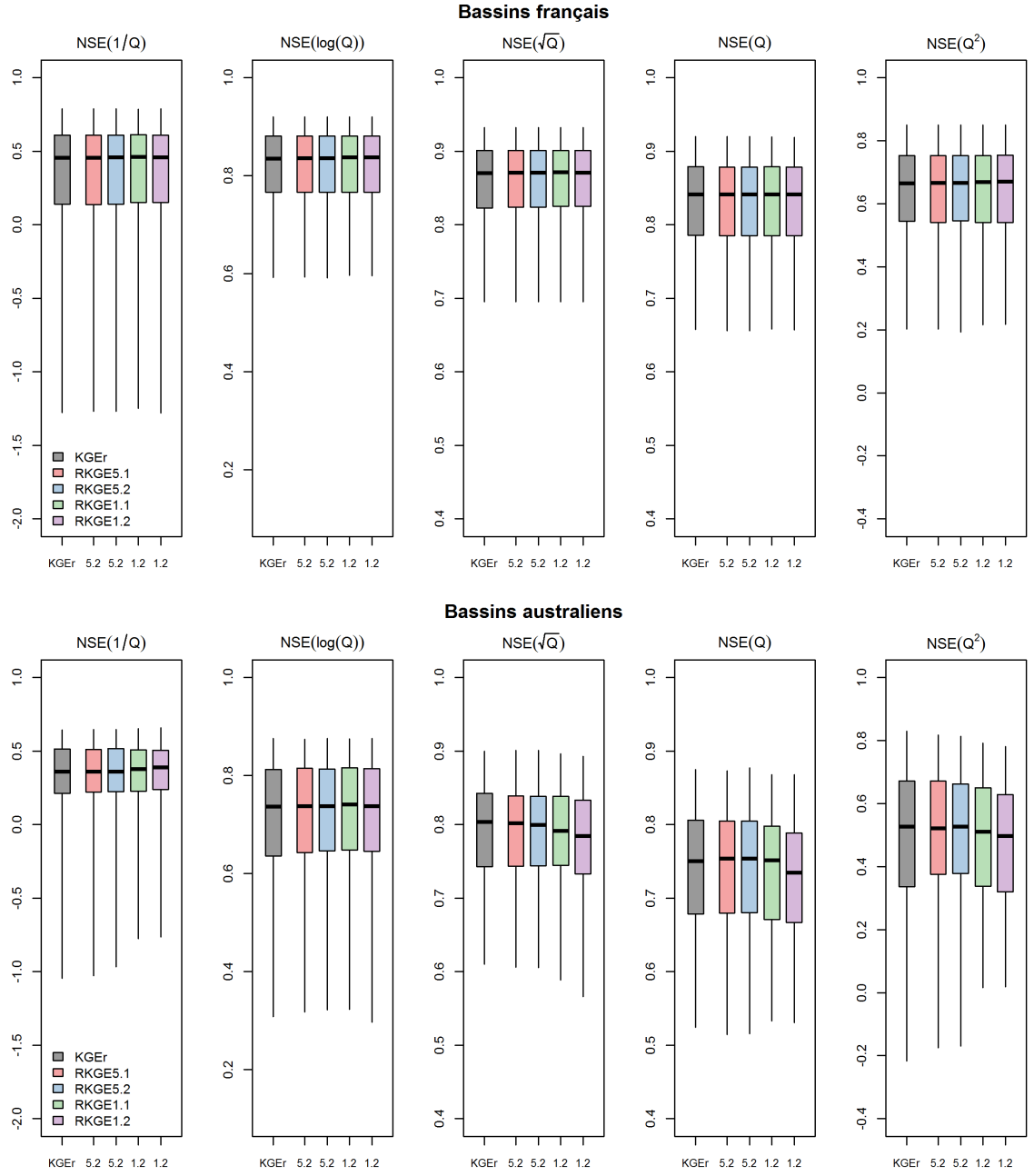


FIGURE C.4 : Distribution des valeurs de NSE calculé sur les débits transformés par cinq différentes transformations (fonctions inverse, logarithme, racine carrée, identité, et carrée) sur les bassins du sous-échantillon français et du sous-échantillon australien obtenues en calage pour le modèle GR4J, calé avec le KGE racine (gris), ou le RKGE5.1 (rouge), RKGE5.2 (bleu), RKGE1.1 (vert) et le RKGE1.2 (violet) calculées sur la racine carrée des débits.

En France, aucune différence n'est visible entre les distributions des performances obtenues en

calage avec les fonctions objectif testées. En Australie, les performances sur les moyennes et hautes eaux (transformations racine carré carrée, identité et carrée) montrent une détérioration des performances médianes lorsque GR4J est calé avec les fonctions RKGE1.1 et RKGE1.2 racine par rapport au KGE racine, mais montrent en parallèle une amélioration pour les bassins les moins bien modélisés par GR4J. Ce résultat est cohérent avec les précédents vis-à-vis des distributions de biais et de PMR, où les simulations de GR4J calé avec le KGE racine étaient plus robustes en hautes eaux. Au regard des basses eaux (transformation inverse), les fonctions objectif RKGE1.1 et RKGE1.2 racine montrent des performances légèrement améliorées.

Dans l'ensemble, les différences de performance en calage associées au différentes fonctions objectif sont assez faibles. Malgré cela, ces résultats montrent que l'intégration du PMR comme quatrième composante au KGE racine tend à favoriser la simulation des débits en basses eaux, plutôt au détriment des simulations en hautes eaux. Ce constat est particulièrement remarquable lorsque le PMR est calculé sur des sous-périodes d'un an.

C.4. Discussion

D'après la similarité des performances en calage en termes de KGE racine (Figure C.1), il semble vraisemblable de supposer que les faibles différences de comportement entre le KGE racine et les RKGE racine testées sont imputables à de faibles contraintes imposée par le PMR sur la sélection des paramètres de GR4J. De ce fait, le choix d'employer le KGE racine ou l'un des RKGE racine ne peut se faire qu'en délibérant sur des détails propres à chacune des fonctions objectif.

Malgré la grande cohérence des résultats entre les différentes implémentations du RKGE racine, il semble que le RKGE1.2 se détache légèrement du groupe par la favorisation plus marquée de la simulation des basses eaux, accompagnée par une détérioration de celle des hautes eaux. En Australie, l'amélioration des performances pour les bassins les moins bien modélisés par GR4J est égale à celle observée avec le RKGE1.1, mais s'accompagne par une détérioration plus marquée des performances médianes. Sur la base de ces constats, qui, rappelons-le, sont des remarques de détail, nous proposons d'éviter d'utiliser l'utilisation du RKGE1.2 racine.
La sélection d'un RKGE racine parmi les trois options restantes ne peut pas se faire à partir de considérations sur les performances, qui sont toutes assez similaires. En revanche, en remarquant que la définition de sous-périodes de 5 ans exige de disposer de chroniques suffisamment longues pour disposer d'un nombre raisonnable de sous-périodes pour le calcul du PMR, il apparaît que le calcul du RKGE1.1 est les plus facile à mettre en œuvre dans un contexte générique. Calculer le PMR sur des sous-périodes de un an laisse en effet imaginer que le RKGE puisse être utilisé même dans des situations où seulement quelques années d'observation sont à disposition pour le calage de modèles hydrologiques. Nous considérons par ailleurs que la baisse légère des performances médianes en hautes eaux est compensée par l'amélioration légère des performances en basses eaux et par la limitation des erreurs dans les bassins « outliers ».
De ce fait, le RKGE1.1 semble être la fonction objectif le plus pratique à utiliser tout en offrant des performances comparables à celles obtenues avec le KGE racine.

Enfin, il convient de remarquer que la simplicité du modèle GR4J est peut-être une des causes des mécanismes de compensations entre la hausse des performances en basses eaux et baisse des performances en hautes eaux. Pour des modèles plus complexes, il est possible que l'intégration du PMR comme composante supplémentaire au KGE racine puisse bénéficier plus nettement à leur robustesse. En effet, le PMR pourrait contribuer à la contrainte des paramètres de modules spéci-

fiquement consacrés aux dynamiques interannuelles, dont GR4J semble manquer (voir Chapitre 7 et travaux de Fowler *et al.* 2020).

C.5. Synthèse

L'objectif de cette analyse était l'évaluation des possibilités d'intégrer le PMR comme une quatrième composante au KGE racine, afin d'ajouter une contrainte supplémentaire au calage des paramètres de modèles hydrologiques visant explicitement à améliorer leur robustesse. Pour ce faire, plusieurs manières de calculer le PMR ont été envisagées à partir de considérations sur la longueur des sous-périodes de calcul du PMR et sur le poids donné aux années portant les erreurs les plus remarquables du modèle. Au total, quatre fonctions objectif RKGE, pour *Robust KGE*, ont été implémentées et testées en les appliquant à la racine carrée des débits. L'évaluation a été réalisée à travers la comparaison des performances du modèle GR4J en calage sur toutes les données disponibles des bassins de l'échantillon pour chacune des fonctions objectif RKGE racine, avec comme référence les performances obtenues en calage avec le KGE racine.

Nos résultats montrent que les fonctions objectif RKGE racine sont associées à une hausse légère des performances du modèle pour la simulation des eaux basses et intermédiaires, accompagnée par une baisse légère des performances en hautes eaux. Les pertes les plus remarquables sont obtenues pour le RKGE1.2 racine, construit pour donner davantage de poids aux années où le modèle est le plus biaisé.
Des considérations sur la facilité d'utilisation du RKGE nous ont conduit à privilégier le RKGE1.1 racine, pour lequel le PMR est calculé comme la moyenne des écarts aboslus des biais du modèle sur des sous-périodes d'un an. Ce critère a été baptisé RKGE racine et a été utilisé pour caler GR4J et ses variantes dans le Chapitre 8.

En perspective de cette analyse, il semble qu'une étude approfondie du RKGE racine soit intéressante à réaliser en contrôle.

Annexe D.

Vérification des données de précipitations de cinq bassins versants français

D.1. Méthodologie

Cette annexe expose l'analyse de la qualité des données de précipitations de cinq bassins versants identifiés au Chapitre 8 comme potentiellement concernés par des erreurs de ce type (Figure D.1). Le contrôle des données de précipitations a été réalisé par la comparaison des cumuls des précipitations issues de SAFRAN aux cumuls des précipitations mesurées par des pluviomètres voisins des bassins versants étudiés.
Les mesures pluviométriques ont été récupérées dans la Publithèque de Météo-France[1]. Chaque pluviomètre a été choisi à partir d'un compromis entre proximité géographique avec le bassin versant étudié, longueur des chroniques disponibles, et nombre de déplacements du pluviomètre recensés pendant la période de mesure.

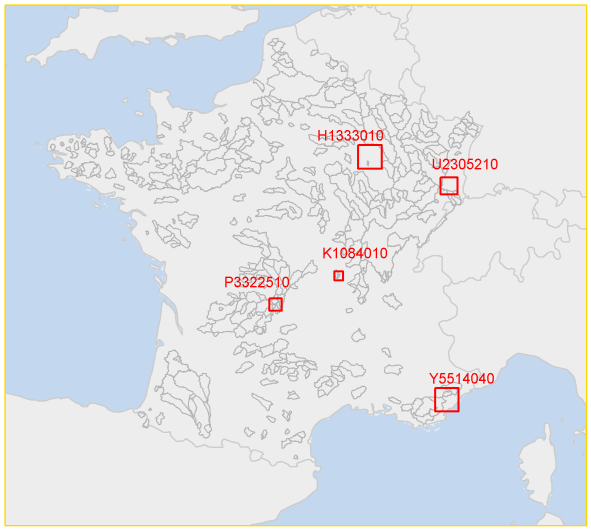


FIGURE D.1 : Carte des bassins où la qualité des données précipitations a été vérifiée.

Pour chaque bassin, une carte indique la position des pluviomètres par rapport aux frontières topographiques du bassin, ainsi que la topographie de la région (MNT à 200 m de résolution Hounam et Werner, 1999) afin d'estimer la cohérence entre les précipitations mesurées par les pluviomètres et les précipitations de bassin versant. L'analyse est ensuite fondée sur l'interprétation des courbes de doubles-cumuls des précipitations SAFRAN et des précipitations des pluviomètres. Enfin, l'évolution du ratio entre les précipitations SAFRAN et les précipitations mesurées par les pluviomètres est confrontée à l'évolution du biais du modèle (courbe des biais glissants).

D.2. Bassin de la Laines à Soulaines-Dhuys (H1333010)

Le bassin de la Laines à Soulaines-Dhuys est situé dans une région relativement plate (Figure D.2). Aucun des pluviomètres utilisés pour la comparaison n'est situé à l'intérieur des frontières géographiques du bassin versant, mais le pluviomètre 1 est localisé assez proche de l'exutoire. Le faible relief de la région permet de supposer que les précipitations sont relativement homogènes dans la région, à l'échelle journalière. Aucun des deux pluviomètres n'a été déplacé pendant la période de mesure.

[1] `publitheque.meteo.fr`

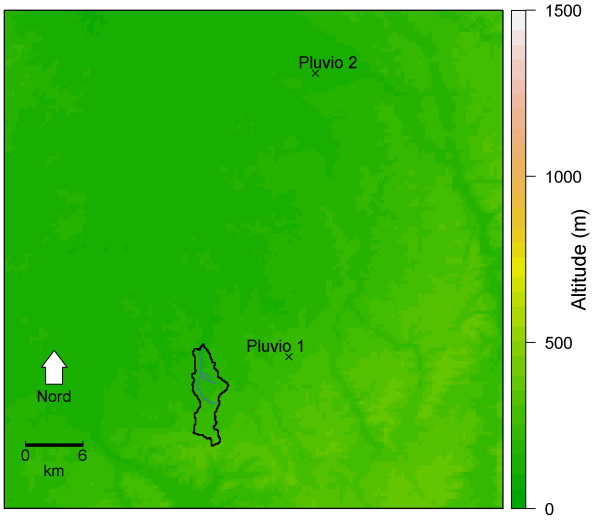


FIGURE D.2 : Carte topographique du bassin de la Laines à Soulaines-Dhuys.

La Figure D.3 montre que les précipitations cumulées mesurées par les deux pluviomètres sont très cohérentes dans le temps. De même, aucune rupture de pente marquée n'est observable.
Enfin, l'évolution des biais du modèle ne semble pas corrélée à celles des rapports entre les précipitations observées dans les différents jeux de données (Figure D.4).

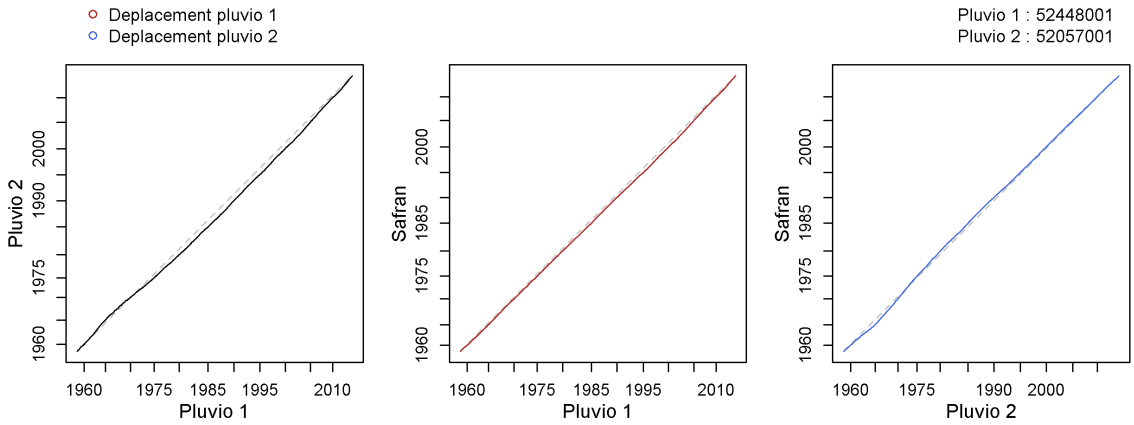


FIGURE D.3 : Courbes de doubles-cumuls des précipitations issues des pluviomètres et de SAFRAN sur le bassin de la Laines à Soulaines-Dhuys.

Aucun indice ne suggère donc que les précipitations SAFRAN sont de mauvaises qualités sur le bassin de la Laines à Soulaines-Dhuys, ni que les éventuelles erreurs n'aient une influence sur les simulations de débit de GR4J.

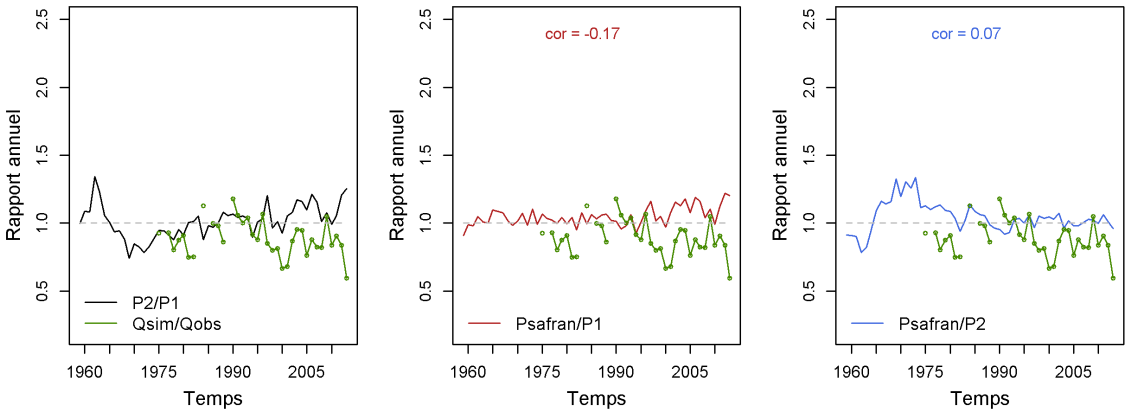


FIGURE D.4 : Comparaison des erreurs sur les volumes annuels du modèle GR4J sur le bassin de la Laines à Soulaines-Dhuys (en vert) aux rapports entre les précipitations annuelles entre les pluviomètres (en noir) ou entre les pluviomètres et les précipitations SAFRAN (en rouge ou en bleu).

D.3. Bassin de la Teyssonne à Changy (K1084010)

Le bassin de la Teyssonne à Changy est situé dans une région présentant un relief prononcé (Figure D.5). Aucun des pluviomètres utilisés pour la comparaison n'est situé à l'intérieur des frontières géographiques du bassin versant. Les deux pluviomètres sont dans des positions géographiques opposées par rapport au bassin. Le pluviomètre 2 est situé dans une vallée au sud du bassin de la Teyssonne, où le relief peut avoir une influence sur le régime des précipitations. Le pluviomètre 1 est situé dans la plaine à l'aval du bassin de la Teyssonne, et capte donc des précipitations peut-être plus cohérentes avec celles du bassin étudié. Aucun des deux pluviomètres n'a été déplacé pendant la période de mesure.

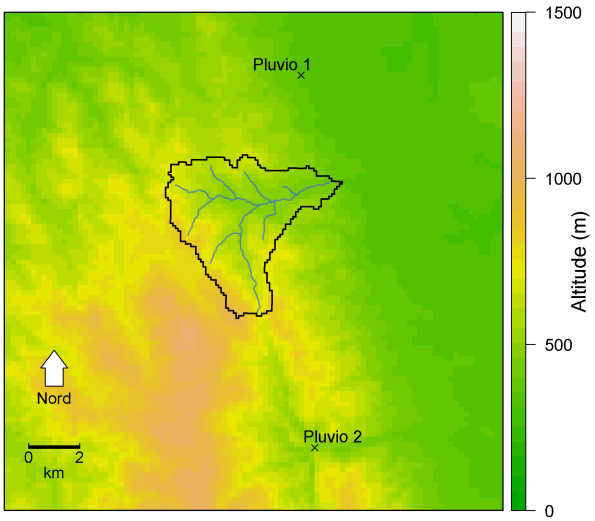


FIGURE D.5 : Carte topographique du bassin de la Teyssonne à Changy.

La Figure D.6 montre que les précipitations cumulées mesurées par les deux pluviomètres sont très cohérentes dans le temps. De même, aucune rupture de pente marquée n'est observable.
Les biais de GR4J sont remarquablement corrélés aux rapports entre les précipitations SAFRAN et les précipitations mesurées par les deux pluviomètres. De ce fait, il est possible qu'une partie des erreurs du modèle proviennent de précipitations mal estimées certaines années. Toutefois, les rapports entre précipitations de bassin et précipitations de pluviomètres restent assez stables dans

le temps, en centrés sur 1, ce qui suggère que les erreurs de mesure sont vraisemblablement marginales.

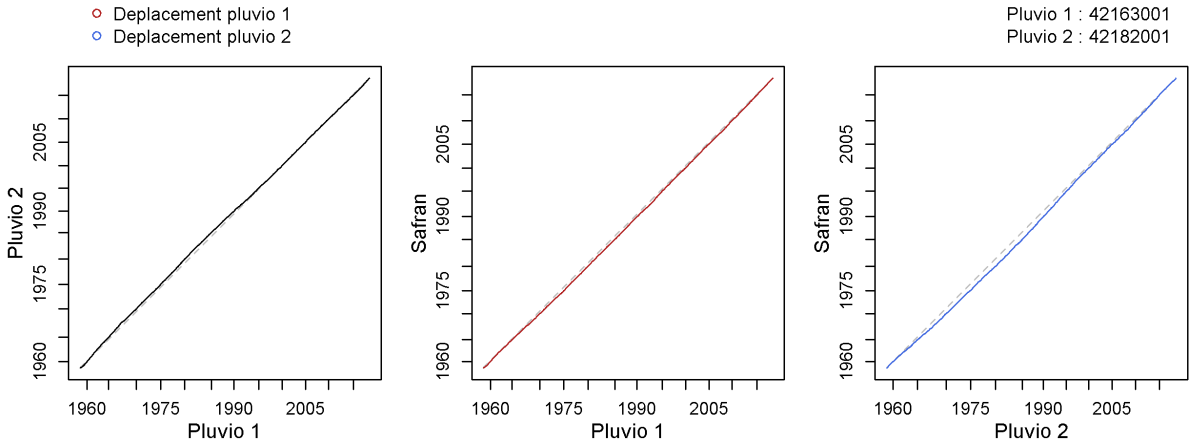


FIGURE D.6 : Courbes de doubles-cumuls des précipitations issues des pluviomètres et de SAFRAN sur le bassin de la Teyssonne à Changy.

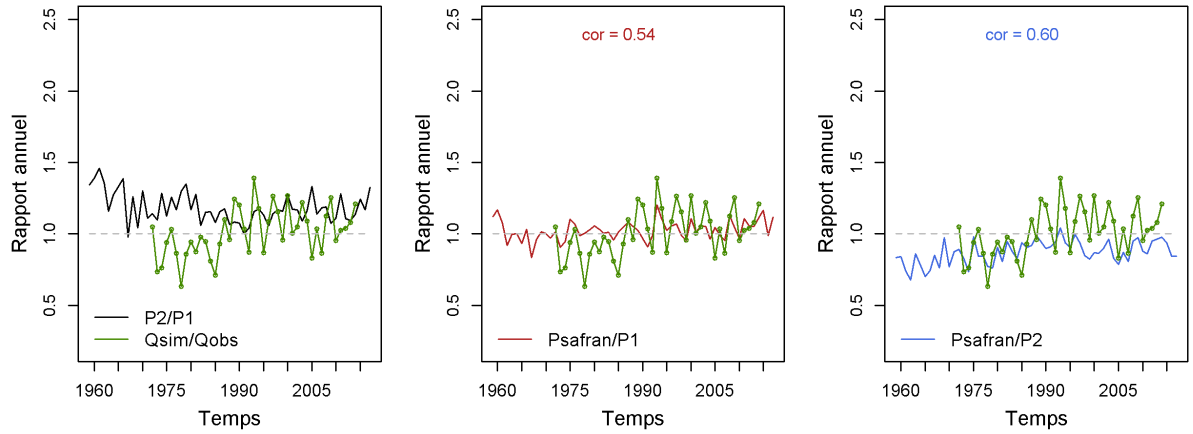


FIGURE D.7 : Comparaison des erreurs sur les volumes annuels du modèle GR4J sur le bassin de la Teyssonne à Changy (en vert) aux rapports entre les précipitations annuelles entre les pluviomètres (en noir) ou entre les pluviomètres et les précipitations SAFRAN (en rouge ou en bleu).

Finalement, il est possible que les précipitations SAFRAN soient erronées sur le bassin de la Teyssonne à Changy et que cela ait une influence sur les biais de GR4J en simulation. Néanmoins, les courbes de doubles-cumuls suggèrent que ces erreurs sont faibles.

D.4. Bassin de la Corrèze à Saint-Yrieix-le-Déjalat (P3322510)

Le bassin de la Corrèze à Saint-Yrieix-le-Déjalat est au bord d'un plateau (Figure D.8). Aucun des pluviomètres utilisés pour la comparaison n'est situé à l'intérieur des frontières géographiques du bassin versant. Le pluviomètre 1 est à l'ouest du bassin versant, en bas du plateau, tandis que le pluviomètre 2 est situé au nord, sur le plateau.
Aucune rupture marquée n'est visible sur les courbes de doubles-cumuls (Figure D.9).

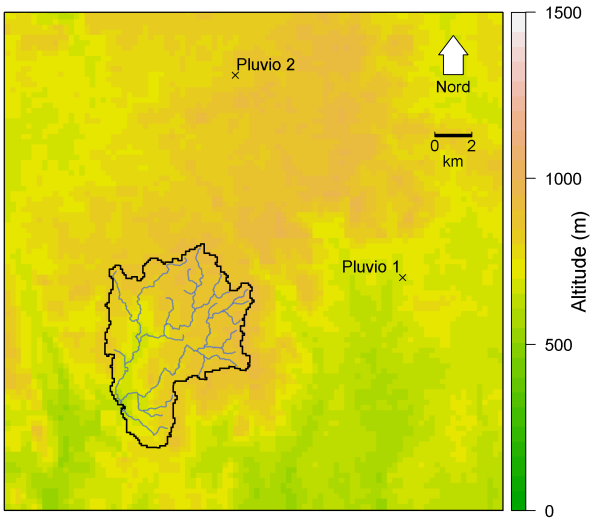


FIGURE D.8 : Carte topographique du bassin de la Corrèze à Saint-Yrieix-le-Déjalat.

Les biais de GR4J sont remarquables dans la mesure où ils sont généralement négatifs avant 1990, et deviennent plutôt positifs après cette date (Figure D.10). Il semble que les précipitations SAFRAN augmentent très légèrement inférieures par rapport aux précipitations mesurées par le pluviomètre 2 après 1990. Les biais du modèle sont d'ailleurs assez corrélés avec le ratio des deux produits de précipitations. Le phénomène est toutefois trop peu marqué pour pouvoir conclure de manière, d'autant plus qu'il est invisible sur la courbe de doubles-cumuls. Le pluviomètre 2 est d'ailleurs le plus éloigné du bassin versant.

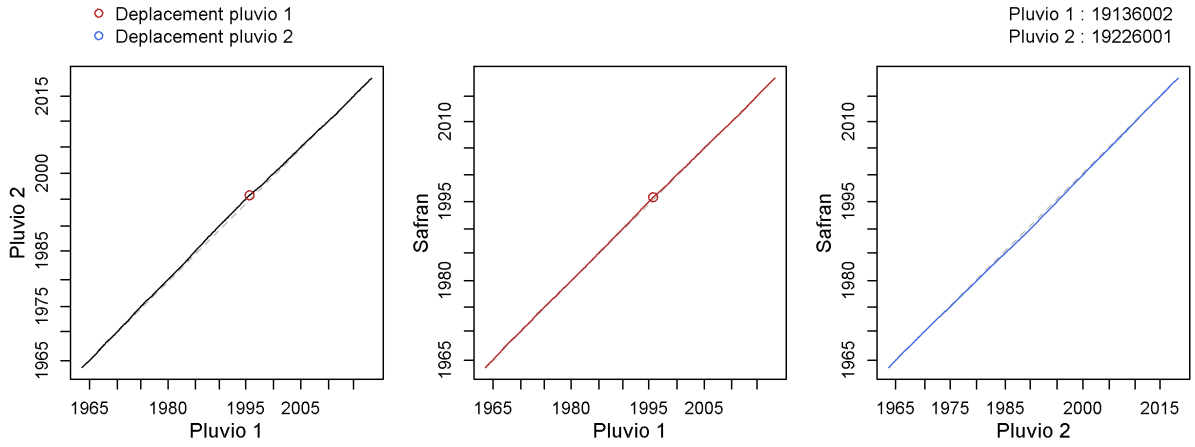


FIGURE D.9 : Courbes de doubles-cumuls des précipitations issues des pluviomètres et de SAFRAN sur le bassin de la Corrèze à Saint-Yrieix-le-Déjalat.

De même que pour le bassin la Teyssonne à Changy, il n'est pas exclu que des erreurs dans les précipitations existent et influencent les biais dans les débits modélisés par GR4J pour le bassin de la Corrèze à Saint-Yrieix-le-Déjalat.

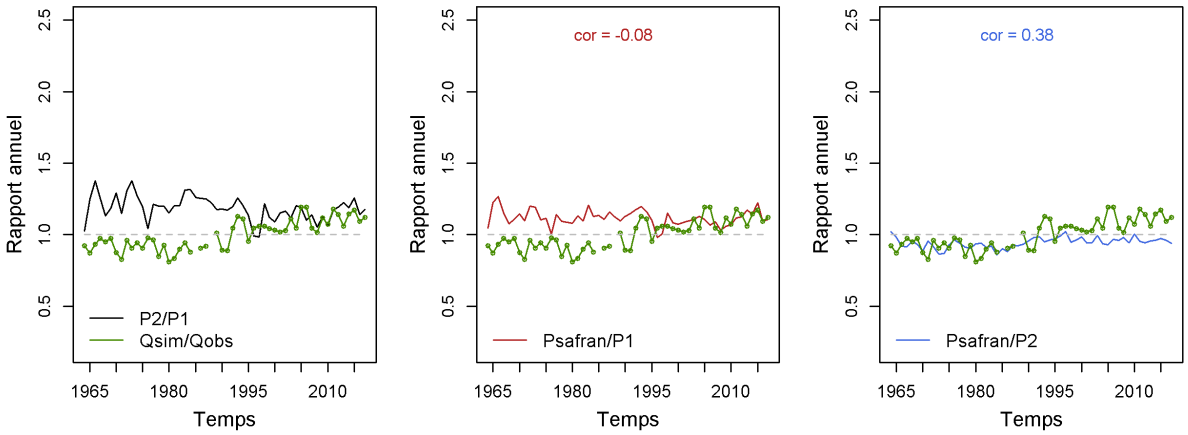


FIGURE D.10 : Comparaison des erreurs sur les volumes annuels du modèle GR4J sur le bassin de la Corrèze à Saint-Yrieix-le-Déjalat (en vert) aux rapports entre les précipitations annuelles entre les pluviomètres (en noir) ou entre les pluviomètres et les précipitations SAFRAN (en rouge ou en bleu).

D.5. Bassin du Saint-Nicolas à Rougemont-le-Château (U2305210)

Les pluviomètres utilisés pour vérifier les précipitations SAFRAN sur le bassin du Saint-Nicolas sont assez éloignés du bassin, en comparaison de sa taille (Figure D.11). Le bassin est dans une zone de relief marqué, tout comme le pluviomètre 2, situé dans un bassin versant au nord du bassin étudié. Le pluviomètre 1 est à l'Ouest du bassin, en plaine. Ces pluviomètres ont subi plusieurs repositionnements entre 1965 et 2005.

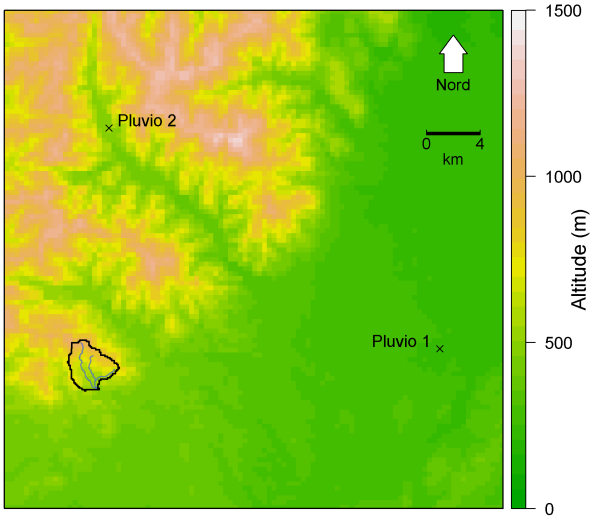


FIGURE D.11 : Carte topographique du bassin du Saint-Nicolas à Rougemont-le-Château.

Aucune rupture marquée n'est visible sur les courbes de doubles-cumuls (Figure D.12).
Il est intéressants de noter que les précipitations moyennes mesurées par les deux pluviomètres diffèrent d'un facteur, et que la moyenne des précipitations SAFRAN est entre ces deux extrêmes (Figure D.13). Il semble que l'hétérogénéité spatiale des précipitations de la région soit donc assez forte. Néanmoins, le bassin étant de très petite taille, il paraît raisonnable de faire l'hypothèse selon laquelle les précipitations au sein du bassin versant sont homogènes en première approche. La corrélation des biais du modèle avec le ratio des précipitations SAFRAN par rapport aux précipita-

tions mesurées par le pluviomètre 2 est assez forte, mais est difficilement explicable par l'analyse visuelle de l'évolution des deux signaux. De plus, l'hétérogénéité spatiale des précipitations de la zone amène à questionner la pertinence de l'usage de ces pluviomètres pour la critique des données SAFRAN.

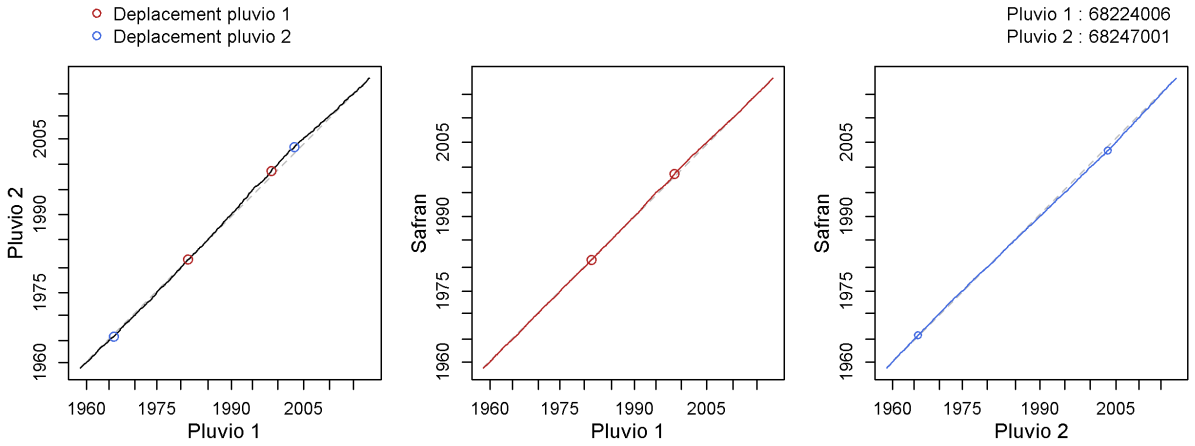


FIGURE D.12 : Courbes de doubles-cumuls des précipitations issues des pluviomètres et de SAFRAN sur le bassin du Saint-Nicolas à Rougemont-le-Château.

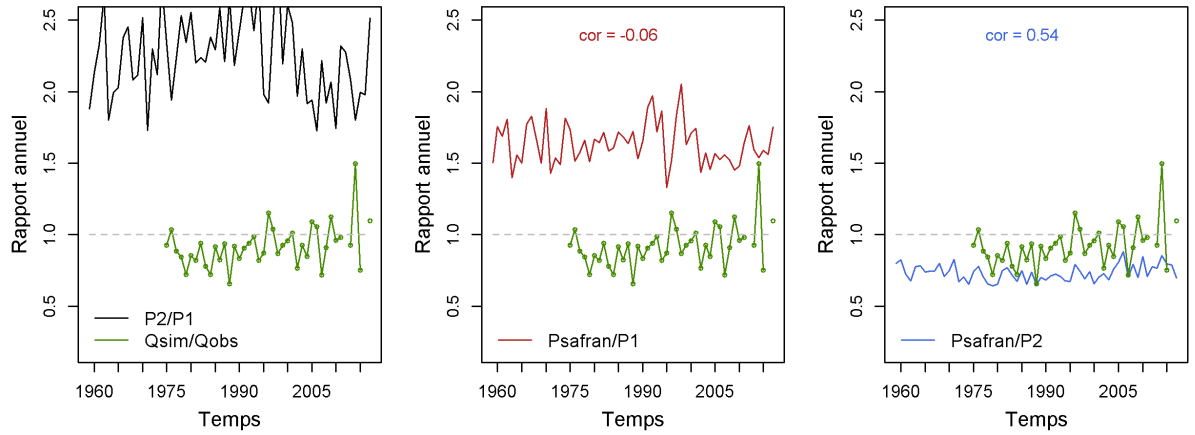


FIGURE D.13 : Comparaison des erreurs sur les volumes annuels du modèle GR4J sur le bassin du Saint-Nicolas à Rougemont-le-Château (en vert) aux rapports entre les précipitations annuelles entre les pluviomètres (en noir) ou entre les pluviomètres et les précipitations SAFRAN (en rouge ou en bleu).

Aucun indice ne suggère donc que les précipitations SAFRAN sont de mauvaise qualité sur le bassin du Saint Nicolas à Rougemont-le-Château. Il est toutefois possible que les éventuelles erreurs de mesures aient une influence sur les simulations de débit de GR4J.

D.6. Bassin de la Siagne à Callian (Y5514040)

Le bassin de la Siagne, situé près de la Mer Méditerranée, est marqué par un très fort relief (Figure D.14). Un des pluviomètres (pluviomètre 2) utilisés est à l'intérieur des frontières topographiques du bassin versant. Ce pluviomètre n'a jamais été déplacé pendant la période étudiée. Le pluviomètre 1 est en revanche assez loin du bassin versant, proche de la côte, et a été repositionné une fois.

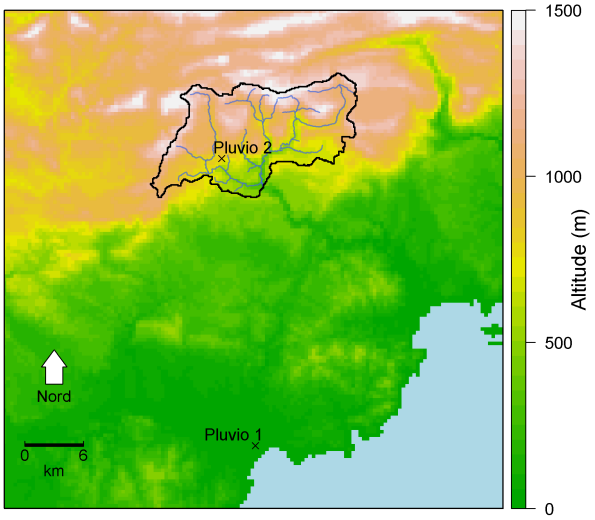


FIGURE D.14 : Carte topographique du bassin de la Siagne à Callian.

Aucune rupture marquée n'est visible sur les courbes de doubles-cumuls (Figure D.15).

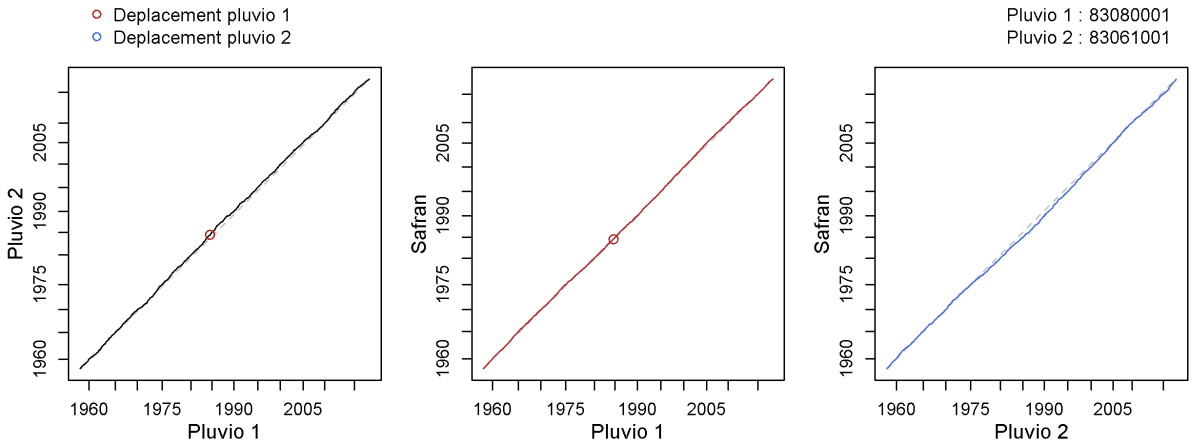


FIGURE D.15 : Courbes de doubles-cumuls des précipitations issues des pluviomètres et de SAFRAN sur le bassin de la Siagne à Callian.

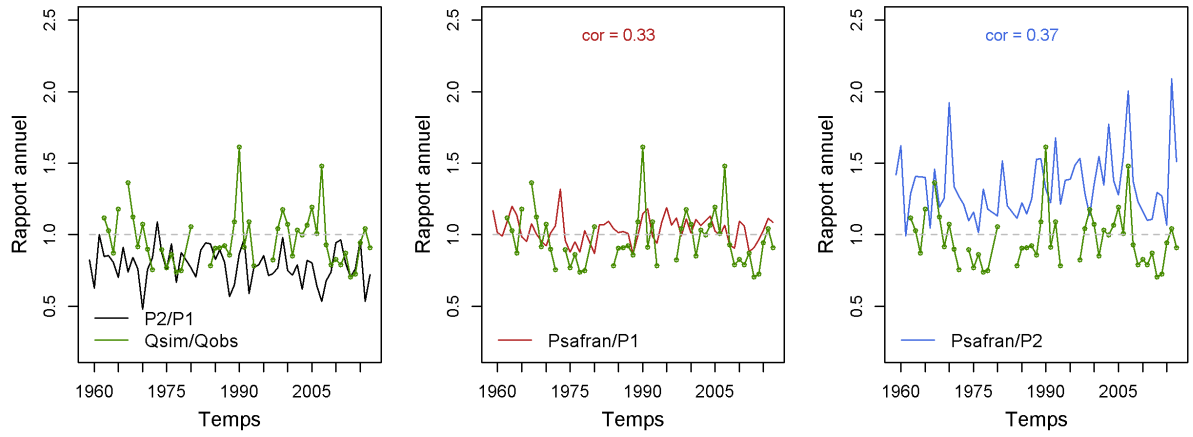


FIGURE D.16 : Comparaison des erreurs sur les volumes annuels du modèle GR4J sur le bassin de la Siagne à Callian (en vert) aux rapports entre les précipitations annuelles entre les pluviomètres (en noir) ou entre les pluviomètres et les précipitations SAFRAN (en rouge ou en bleu).

De même que sur le bassin du Saint-Nicolas, les précipitations moyennes enregistrées aux deux pluviomètres diffèrent fortement, ce qui suggère une forte hétérogénéité spatiale des précipitations (Figure D.16). Ce constat s'explique certainement par la position côtière du bassin et le relief de la région. Les rapports entre les précipitations SAFRAN et les pluviomètres montrent que les précipitations SAFRAN sont plus cohérentes avec celles du pluviomètre 1. Ce constat est contre-intuitif : ce pluviomètre est hors des frontières du bassin, dans un zone sans relief et proche de la mer. De plus, les biais de GR4J sont relativement bien corrélés aux rapports entre les précipitations SAFRAN et les précipitations du pluviomètre 2.
Dans la mesure où la confiance accordée *a priori* au pluviomètre 1 pour représenter les précipitations du bassin était nettement supérieure à celle accordée au pluviomètre 2, la qualité des données SAFRAN est relativement questionnable. Des erreurs dans les données de précipitations en forçage de GR4J sont peut-être à l'origine d'une partie des erreurs du modèle. Il convient toutefois de noter que le sous-sol du bassin est karstique, et que la modélisation hydrologique dans ce contexte est assez complexe.

Finalement, les précipitations SAFRAN de la Siagne à Callian semblent être assez douteuses, bien que cohérentes avec celles enregistrées par un pluviomètre éloigné du bassin.

Annexe E.

Courbes des biais glissants des modèles pour les bassins français de l'échantillon

E.1. Introduction

Cette annexe présente les courbes des biais glissants des modèles GR4J, TOPMO et SimHyd sur le sous-échantillon de 377 bassins versants français de la thèse. Les courbes ont été calculées pour ces modèles forcés avec les formules d'évaporation potentielle d'Oudin ou de Morton.

E.2. Mise en œuvre des courbes

La Figure E.1 illustre un exemple des courbes de biais glissants représentées.

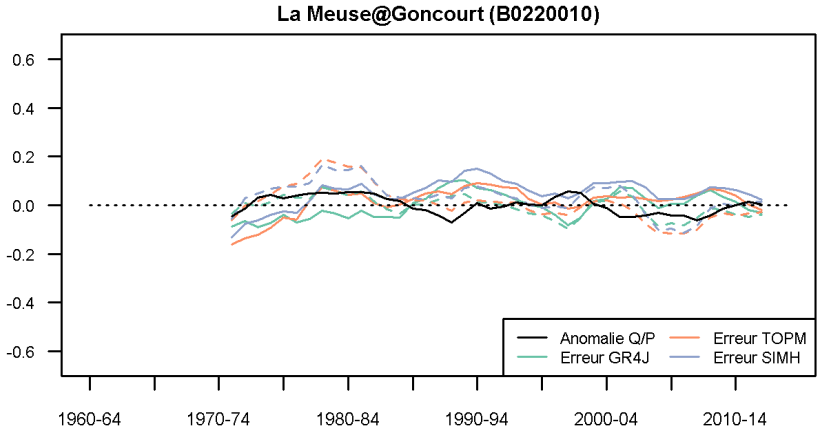
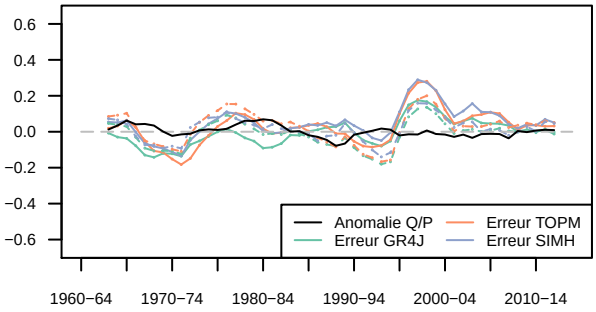


FIGURE E.1 : Courbes des biais glissants des modèles GR4J (vert), TOPMO (orange) et SimHyd (bleu), pour les formulations de l'évaporation potentielle d'Oudin (trait plein) et de Morton (trait pointillé). Les anomalies de rendement des précipitations sont représentées en noir. Les sous-périodes de 5 ans sont définies sur les années hydrologiques.

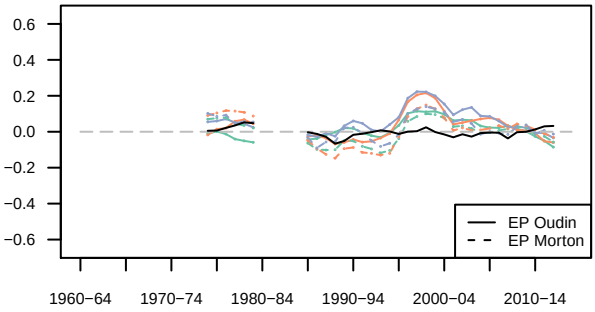
Le biais d'un modèle sur une sous-période de 5 ans a été calculé comme l'erreur moyenne du modèle sur la sous-période rapportée au débit moyen observé sur la période totale, c'est-à-dire de la même manière que dans le calcul du PMR (voir Chapitre 4). Ce n'est donc pas un biais à proprement parler, mais plutôt une erreur normalisée. Nous avons toutefois maintenu l'appellation « courbes des biais glissants ».
Les six courbes des biais glissants (trois modèles hydrologiques pour deux formulations de l'évaporation potentielle) ont été mises en perspective avec l'évolution de l'anomalie de rendement des précipitations au cours du temps, afin d'illustrer les résultats du Chapitre 5. Pour rappel, nous avons montré dans ce chapitre une anti-corrélation importante entre les biais des modèles hydrologiques et l'anomalie de rendement des précipitations. L'anomalie de rendement des précipitations sur une sous-période a été calculée comme l'écart entre le rendement moyen des précipitations sur la sous-période considérée et le rendement moyen des précipitations sur la période totale.

E.3. Courbes de biais glissants pour le sous-échantillon de bassins français

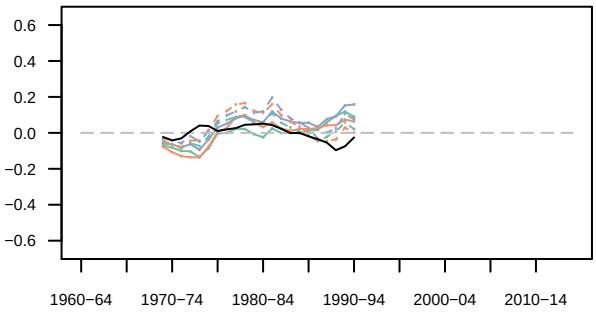
L'Ill@Altkirch (A1050310)



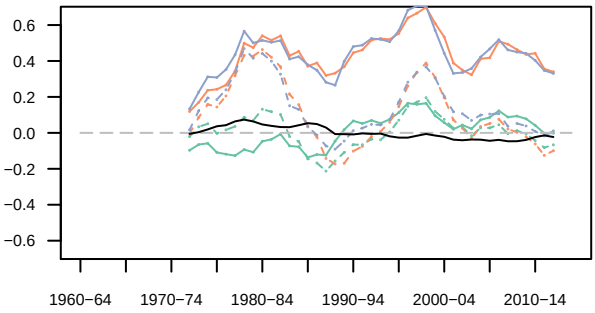
L'Ill@Didenheim (A1080330)



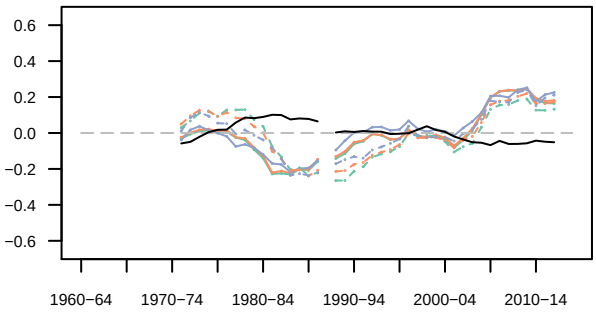
La Largue@Illfurth (A1152010)



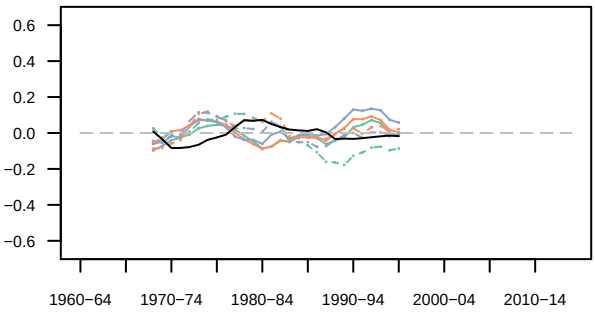
L'Ehn@Niedernai (A2612010)



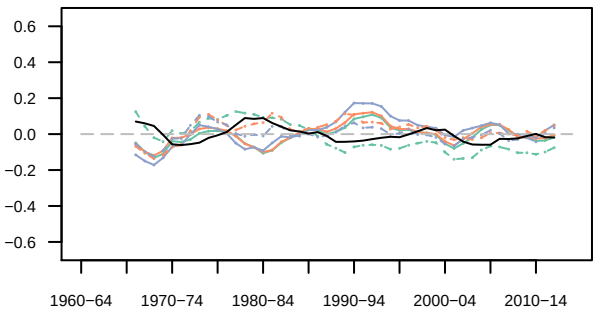
La Mossig@Soultz−les−Bains (A2842010)



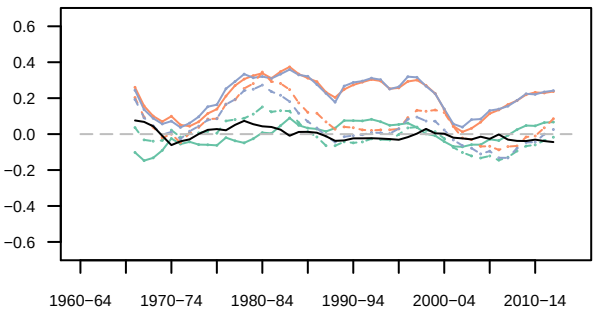
La Moder@Schweighouse−sur−Moder [amont] (A3151010)



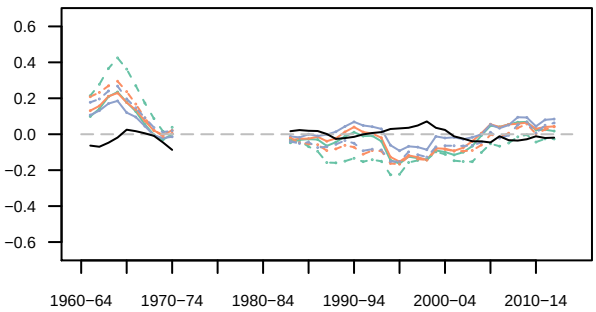
La Moder@Schweighouse−sur−Moder [aval] (A3301010)



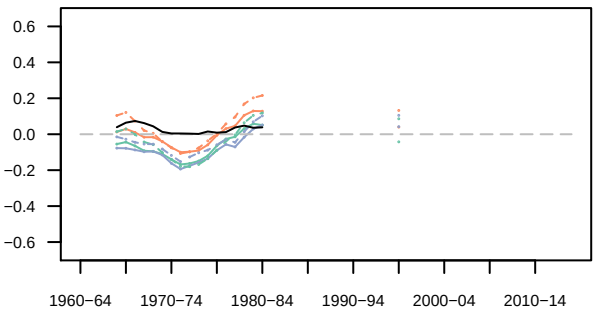
La Zorn@Saverne [Schinderthal] (A3422010)

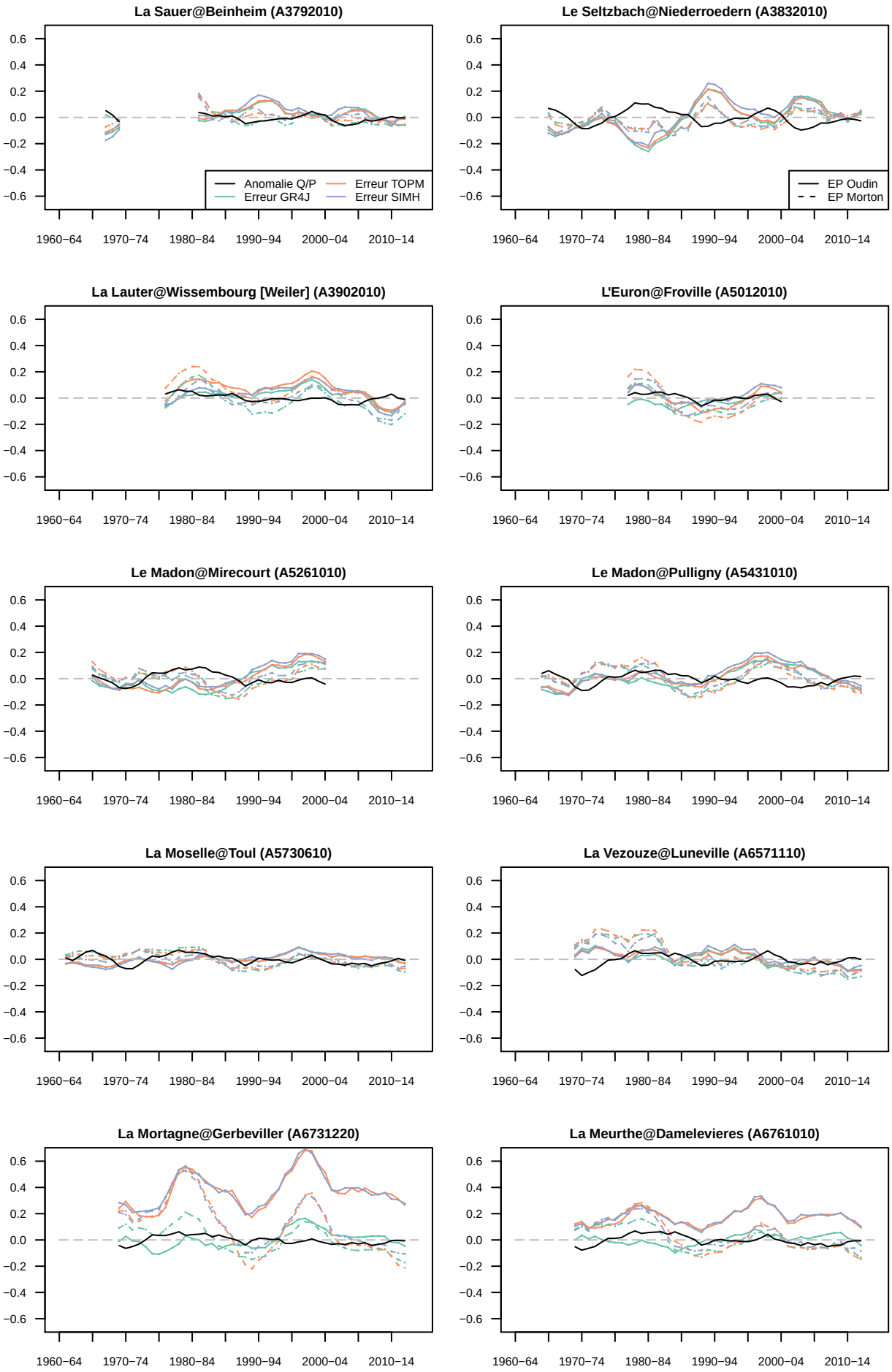


La Zorn@Waltenheim−sur−Zorn (A3472010)

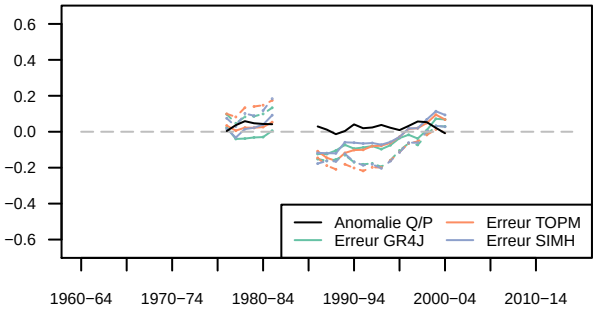


La Sauer@Goersdorf [Liebfrauenthal] (A3712010)

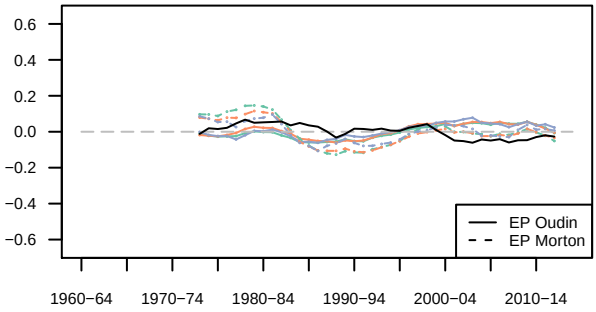



La Sauer@Beinheim (A3792010)
Le Seltzbach@Niederroedern (A3832010)
La Lauter@Wissembourg [Weiler] (A3902010)
L'Euron@Froville (A5012010)
Le Madon@Mirecourt (A5261010)
Le Madon@Pulligny (A5431010)
La Moselle@Toul (A5730610)
La Vezouze@Luneville (A6571110)
La Mortagne@Gerbeviller (A6731220)
La Meurthe@Damelevieres (A6761010)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

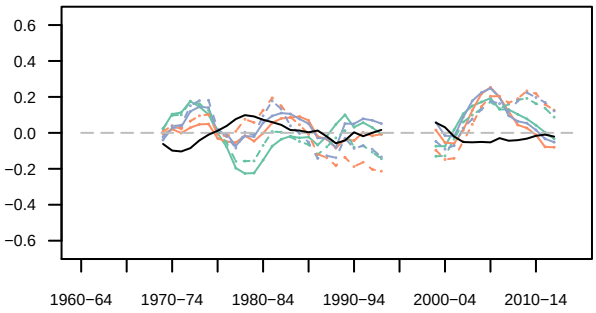
L'Amezule@Lay−Saint−Christophe (A6953010)



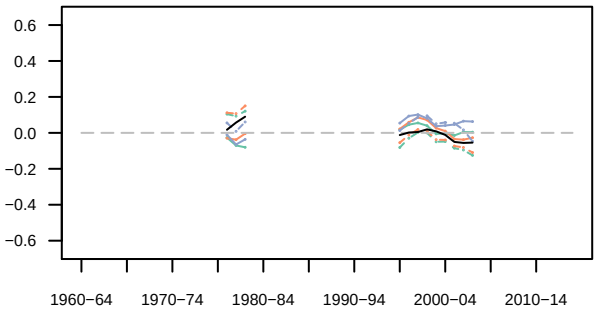
La Moselle@Custines (A7010610)



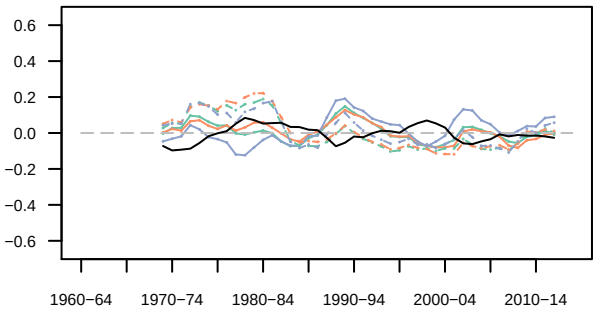
L'Esch@Jezainville (A7122010)



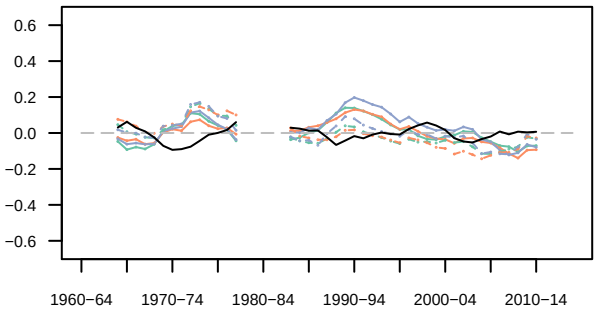
La Petite Seille@Château−Salins (A7642010)



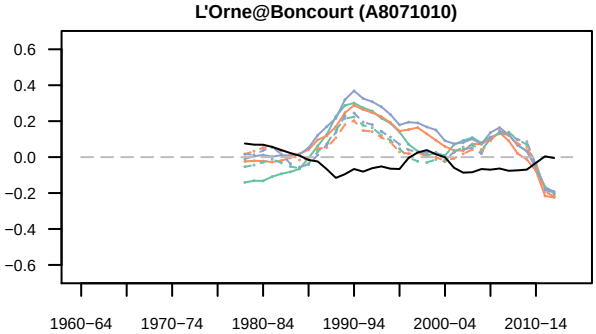
La Seille@Nomeny (A7821010)



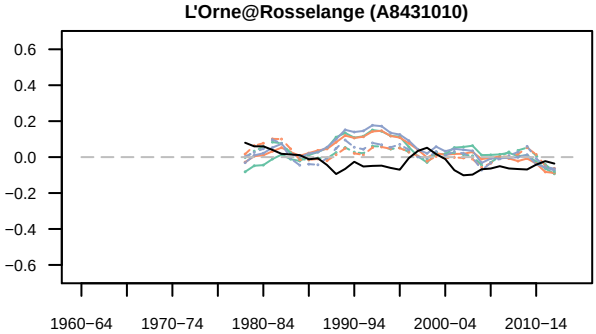
La Seille@Metz (A7881010)



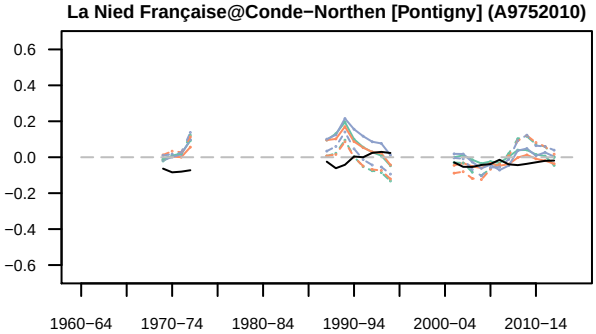
L'Orne@Boncourt (A8071010)



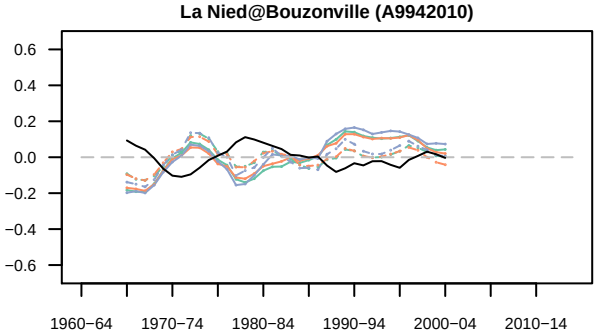
L'Orne@Rosselange (A8431010)

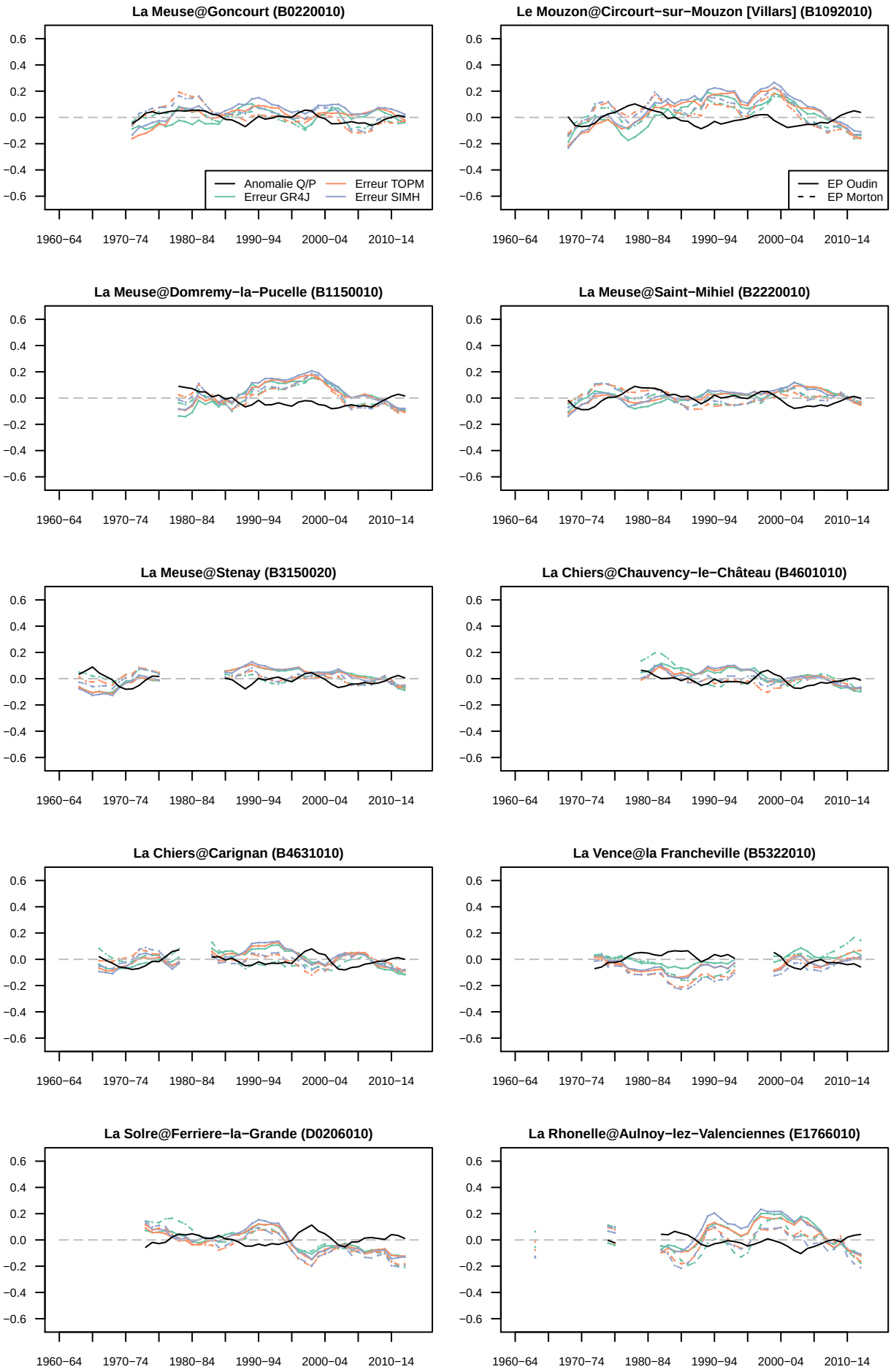


La Nied Française@Conde−Northen [Pontigny] (A9752010)

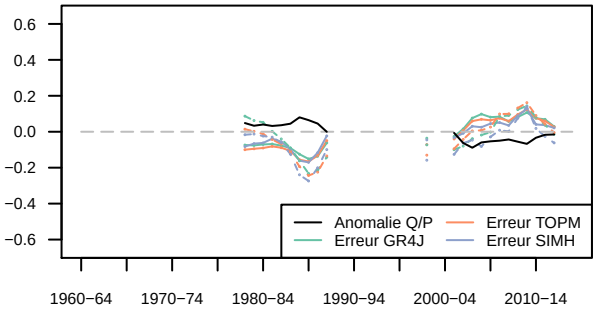


La Nied@Bouzonville (A9942010)

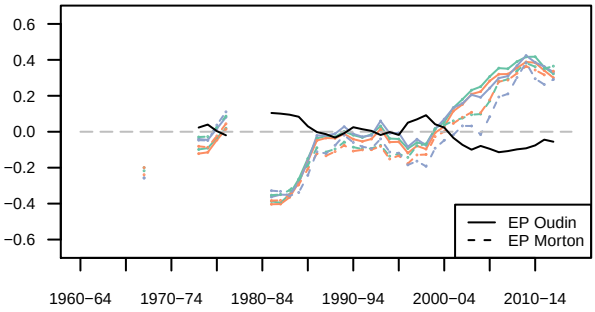



La Meuse@Goncourt (B0220010)
Le Mouzon@Circourt−sur−Mouzon [Villars] (B1092010)
La Meuse@Domremy−la−Pucelle (B1150010)
La Meuse@Saint−Mihiel (B2220010)
La Meuse@Stenay (B3150020)
La Chiers@Chauvency−le−Château (B4601010)
La Chiers@Carignan (B4631010)
La Vence@la Francheville (B5322010)
La Solre@Ferriere−la−Grande (D0206010)
La Rhonelle@Aulnoy−lez−Valenciennes (E1766010)
Anomalie Q/P
Erreur GR4J
Erreur TOPM
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

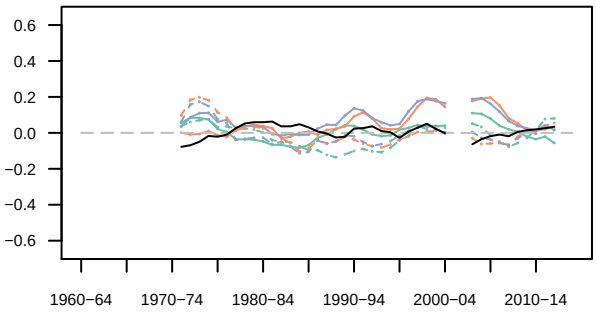
L'Hogneau@Thivencelle (E1827020)



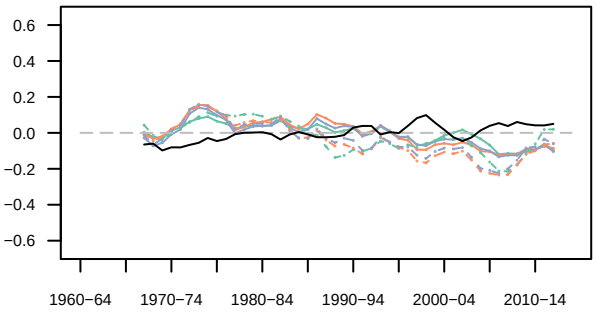
La Marque@Bouvines (E3346010)



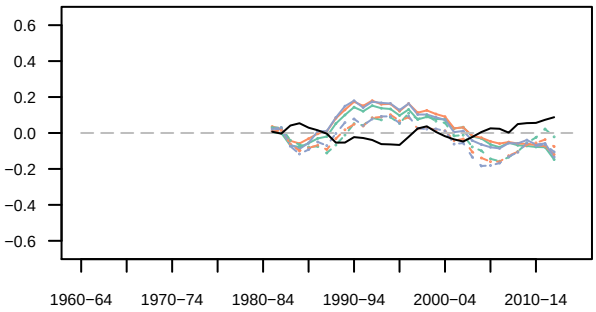
La Lys@Lugy (E3511210)



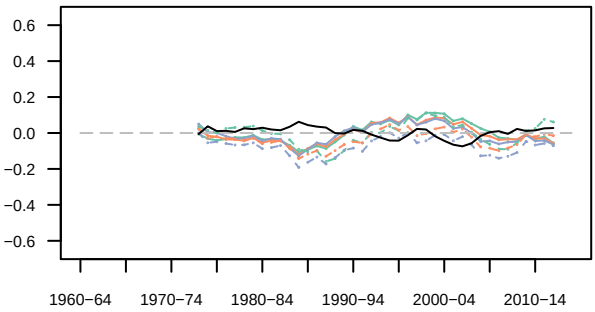
L'Aa@Wizernes (E4035710)



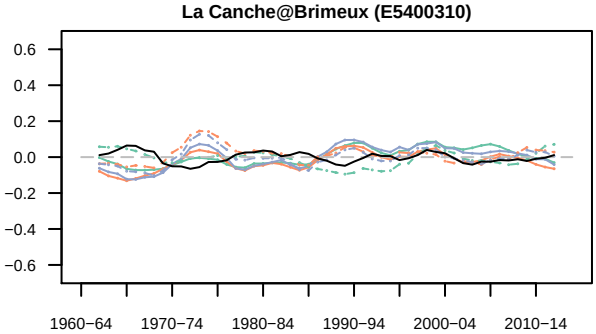
La Hem@Tournehem−sur−la−Hem [Guemy] (E4306010)



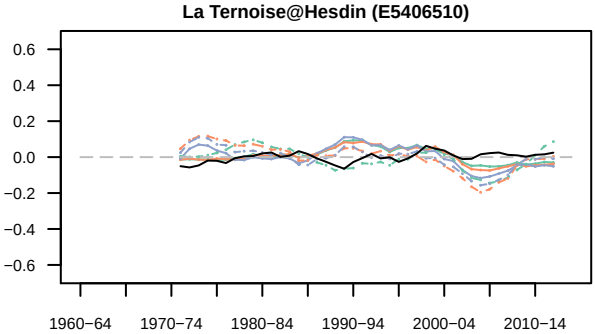
La Liane@Wirwignes (E5300210)



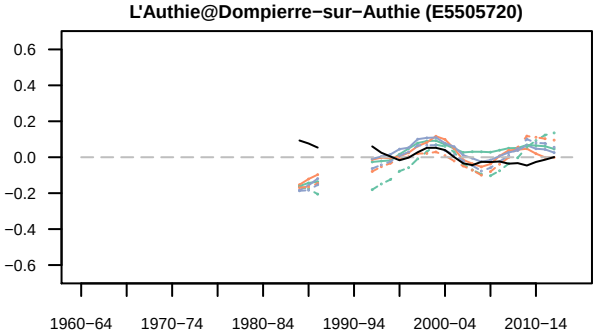
La Canche@Brimeux (E5400310)



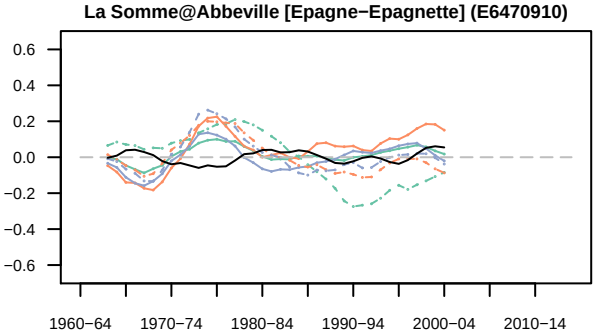
La Ternoise@Hesdin (E5406510)

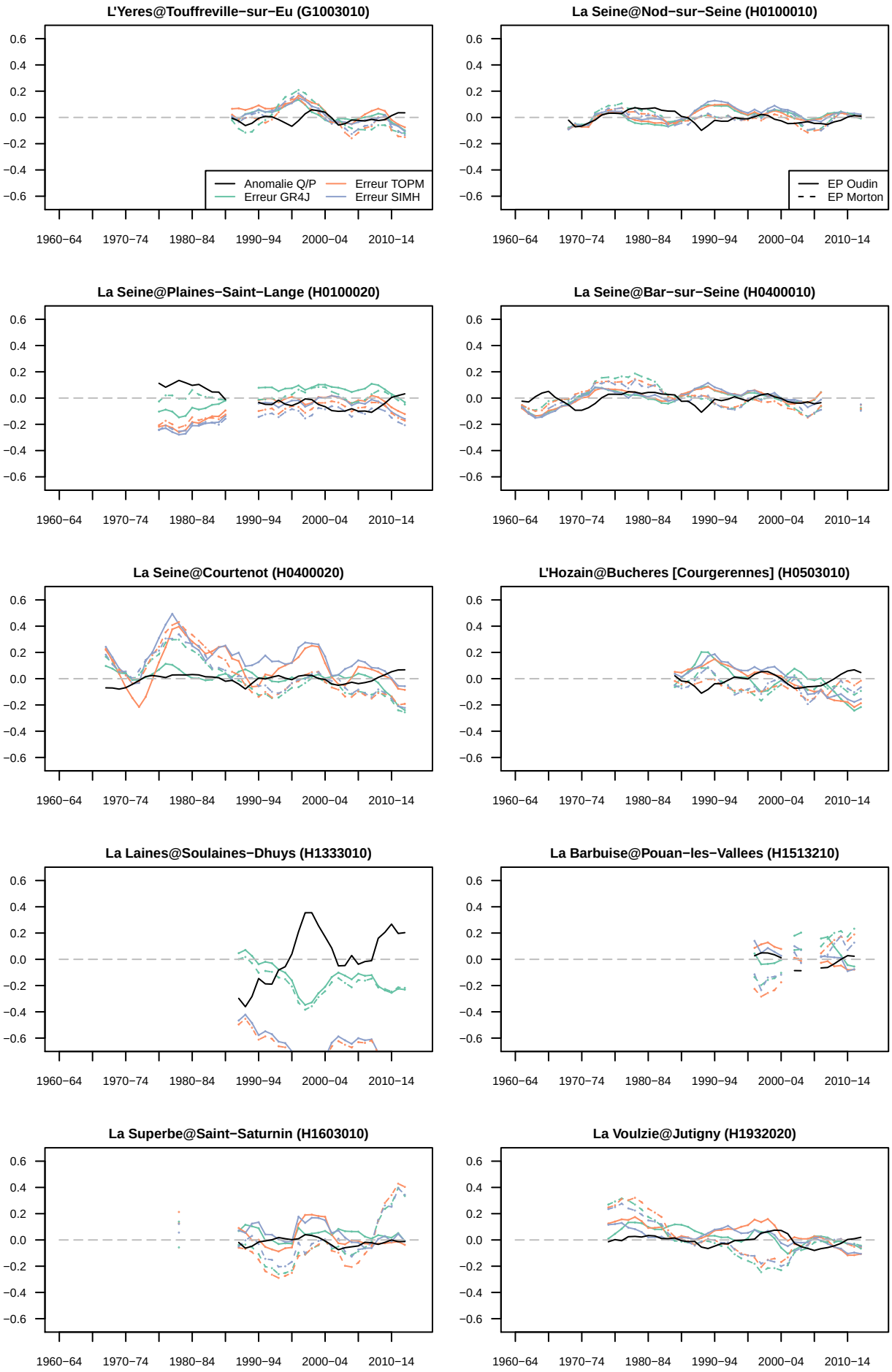


L'Authie@Dompierre−sur−Authie (E5505720)

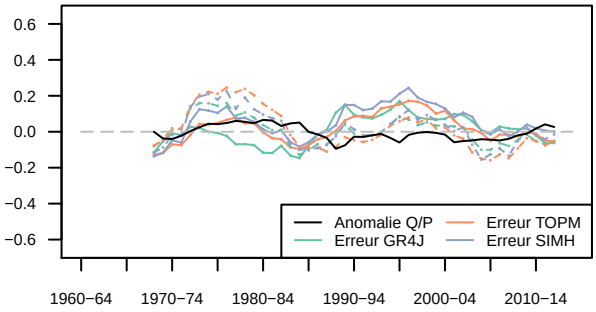


La Somme@Abbeville [Epagne−Epagnette] (E6470910)

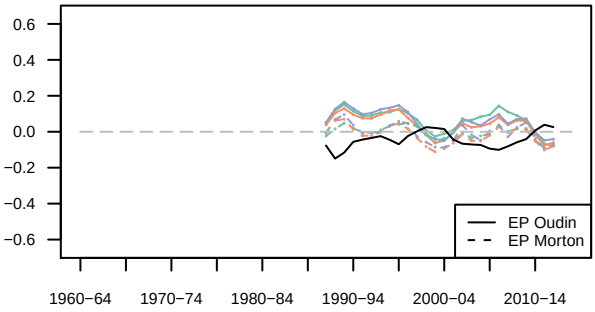



L'Yeres@Touffreville−sur−Eu (G1003010)
La Seine@Nod−sur−Seine (H0100010)
La Seine@Plaines−Saint−Lange (H0100020)
La Seine@Bar−sur−Seine (H0400010)
La Seine@Courtenot (H0400020)
L'Hozain@Bucheres [Courgerennes] (H0503010)
La Laines@Soulaines−Dhuys (H1333010)
La Barbuise@Pouan−les−Vallees (H1513210)
La Superbe@Saint−Saturnin (H1603010)
La Voulzie@Jutigny (H1932020)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

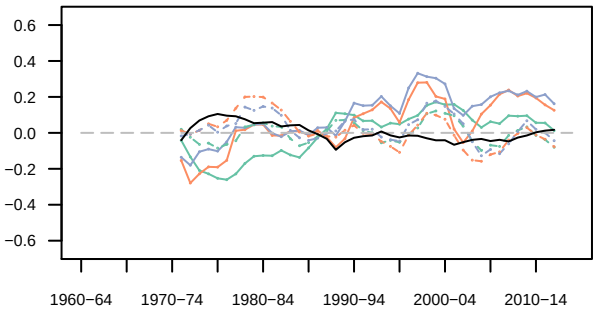
Le Beuvron@Ouagne [Champmoreau] (H2062010)



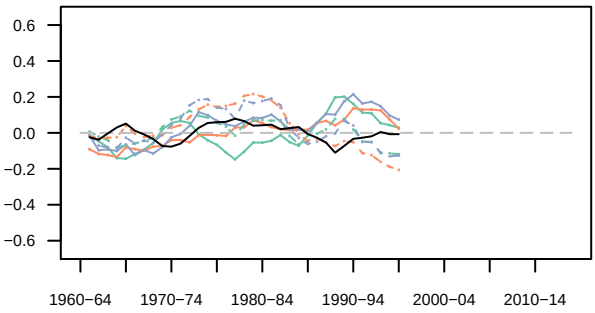
Le Sauzay@Corvol−l'Orgueilleux (H2073110)



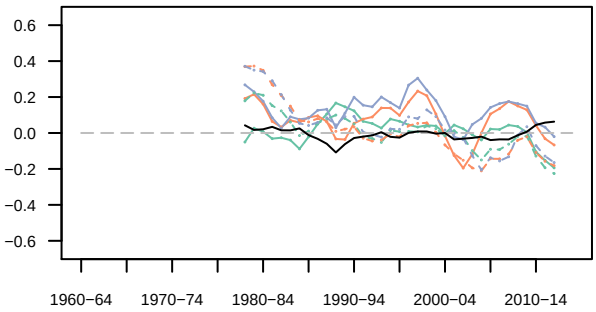
Le Serein@Bierre−les−Semur (H2322010)



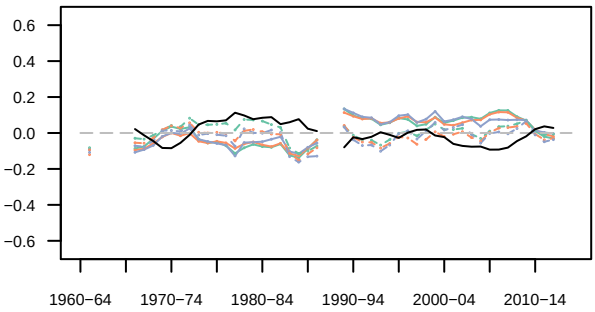
Le Serein@Chablis (H2342010)



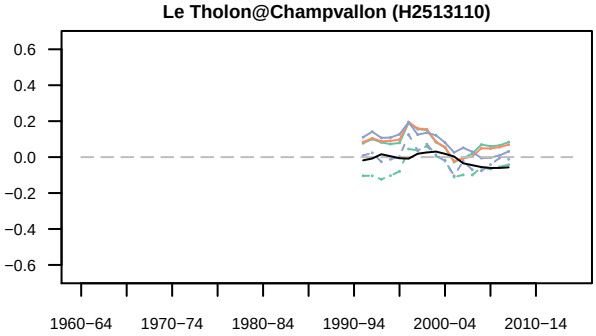
L'Armançon@Quincy−le−Vicomte (H2412010)



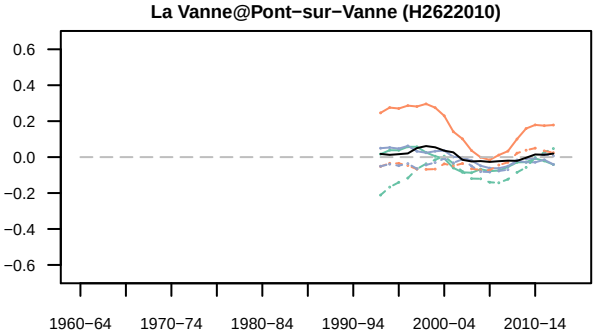
L'Armançon@Brienon−sur−Armançon (H2482010)



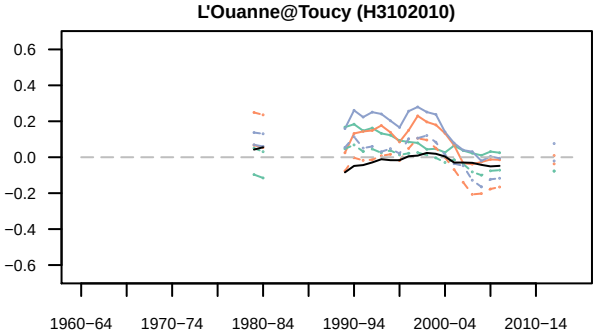
Le Tholon@Champvallon (H2513110)



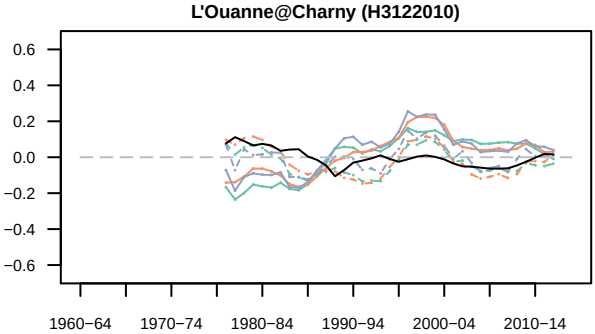
La Vanne@Pont−sur−Vanne (H2622010)

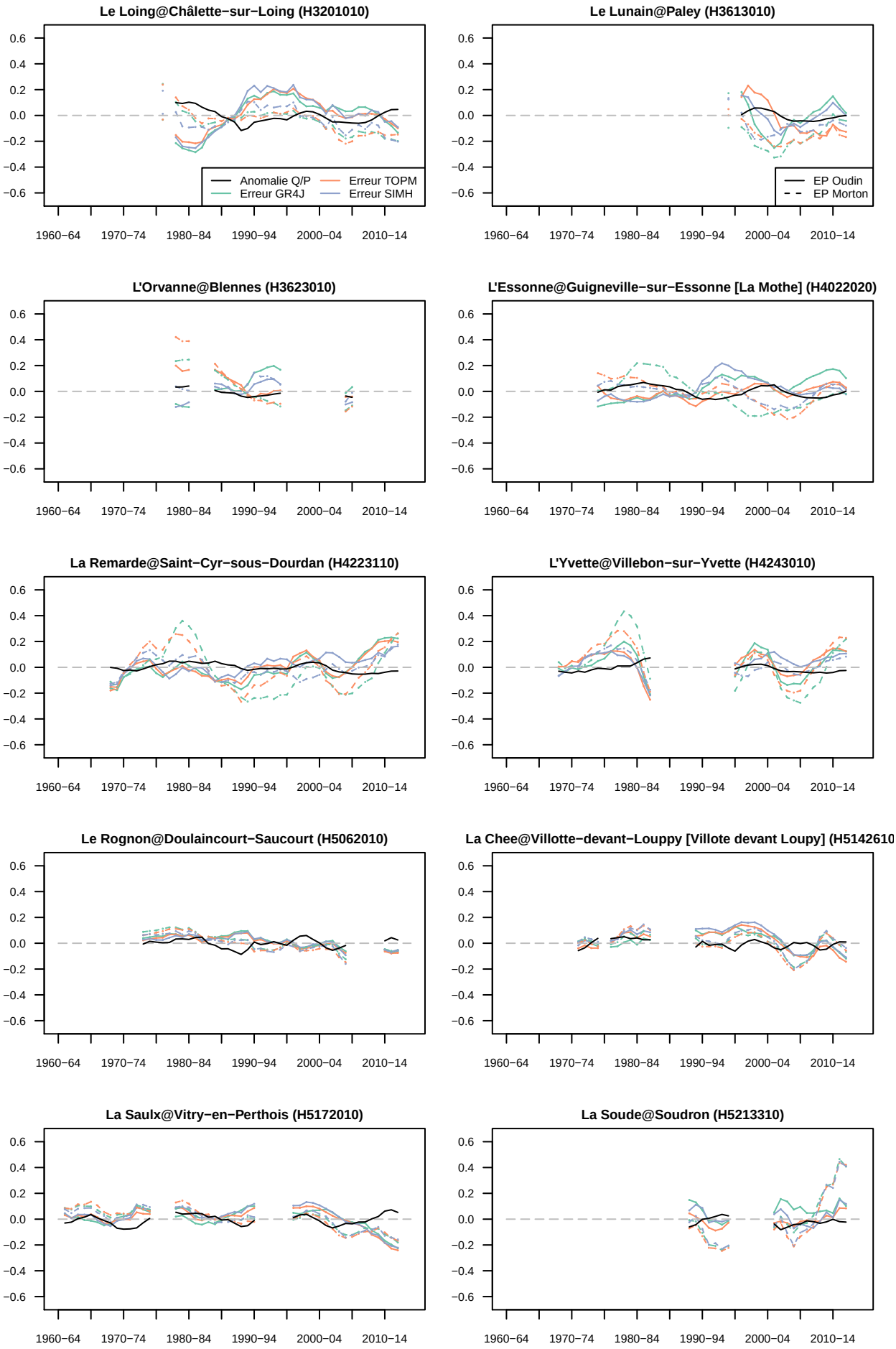


L'Ouanne@Toucy (H3102010)

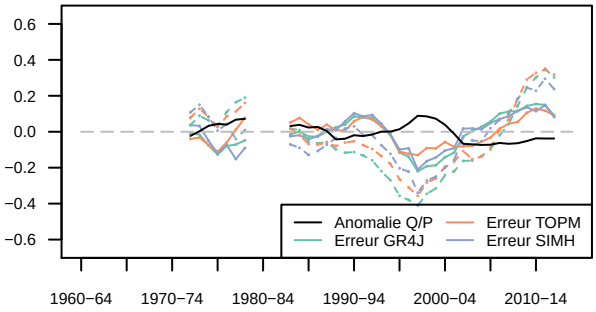


L'Ouanne@Charny (H3122010)

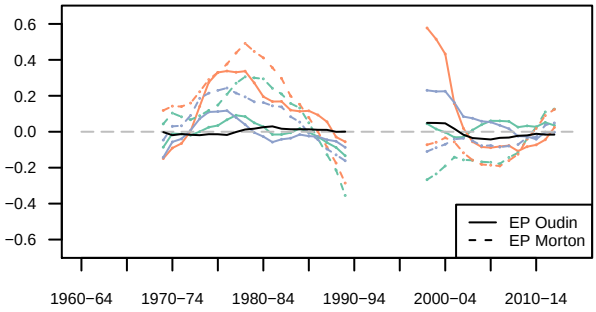



Le Loing@Châlette−sur−Loing (H3201010)
Le Lunain@Paley (H3613010)
L'Orvanne@Blennes (H3623010)
L'Essonne@Guigneville−sur−Essonne [La Mothe] (H4022020)
La Remarde@Saint−Cyr−sous−Dourdan (H4223110)
L'Yvette@Villebon−sur−Yvette (H4243010)
Le Rognon@Doulaincourt−Saucourt (H5062010)
La Chee@Villotte−devant−Louppy [Villote devant Loupy] (H5142610
La Saulx@Vitry−en−Perthois (H5172010)
La Soude@Soudron (H5213310)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

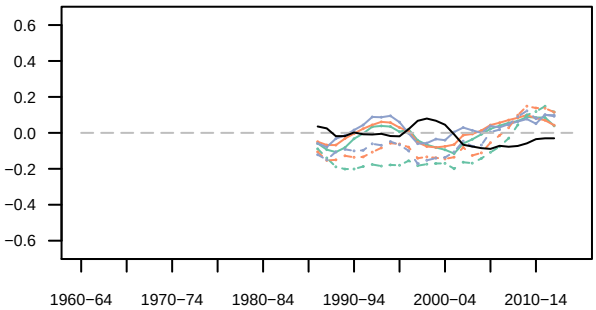
Le Petit Morin@Montmirail (H5412010)



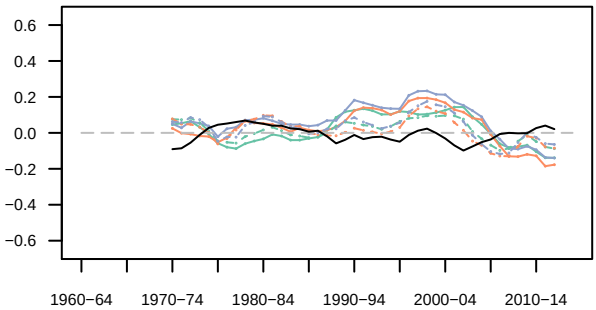
La Therouanne@Congis−sur−Therouanne [Gue@Tresmes] (H561302



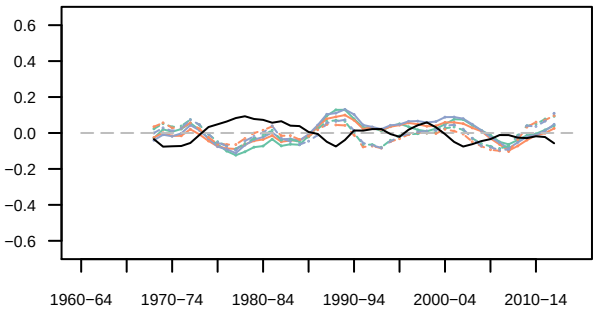
Le Grand Morin@Pommeuse (H5732010)



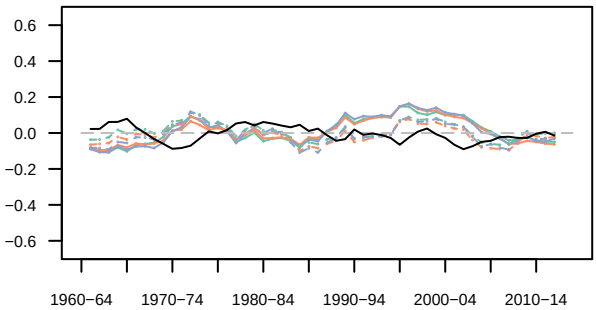
L'Aire@Beausite [Amblaincourt] (H6102010)



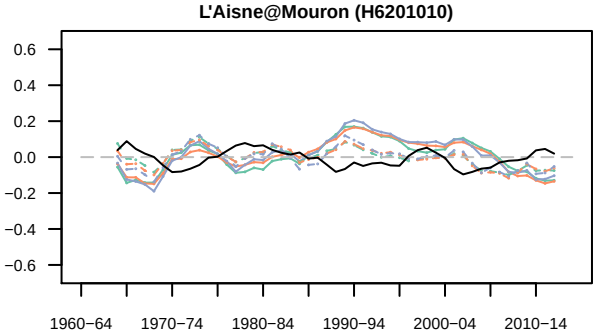
L'Aire@Varennes−en−Argonne (H6122010)



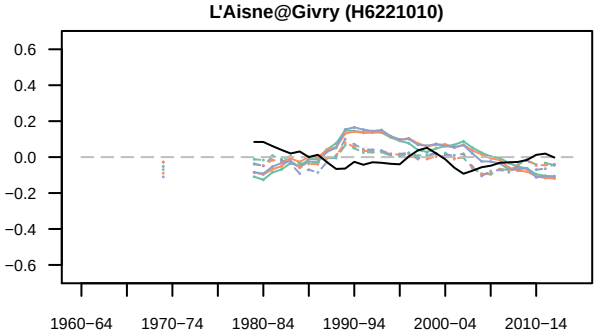
L'Aire@Chevieres (H6162010)



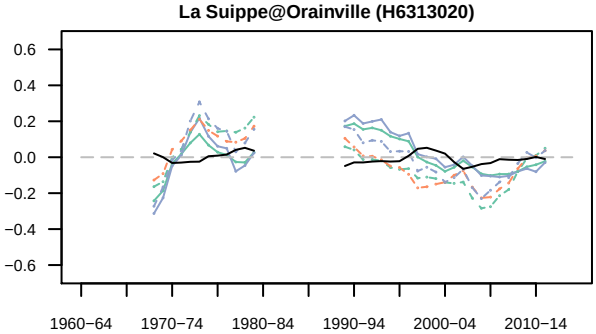
L'Aisne@Mouron (H6201010)



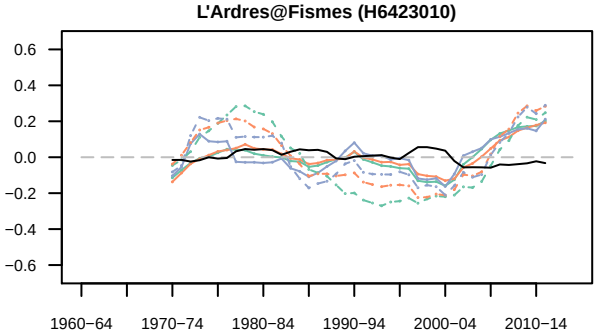
L'Aisne@Givry (H6221010)

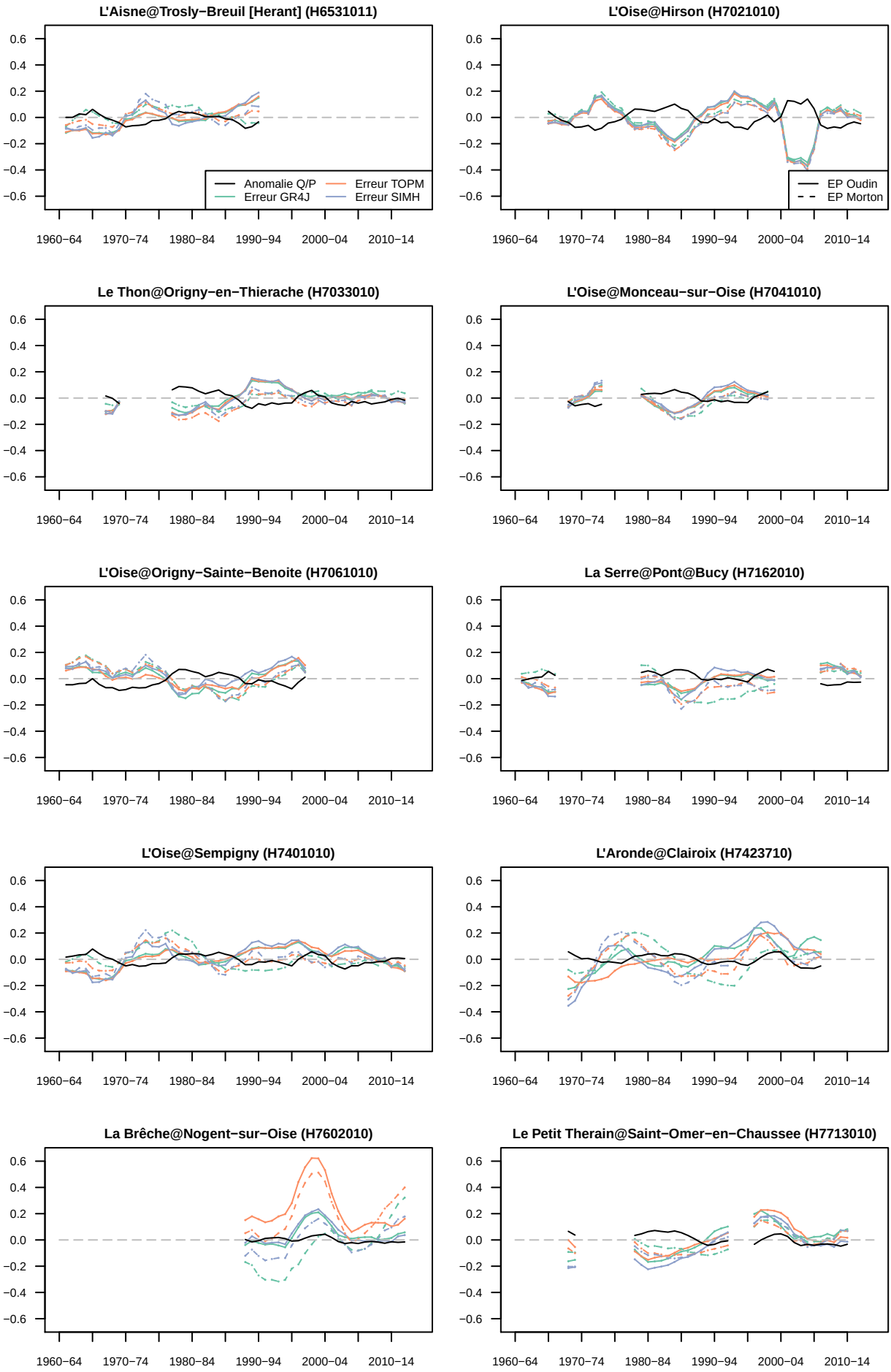


La Suippe@Orainville (H6313020)

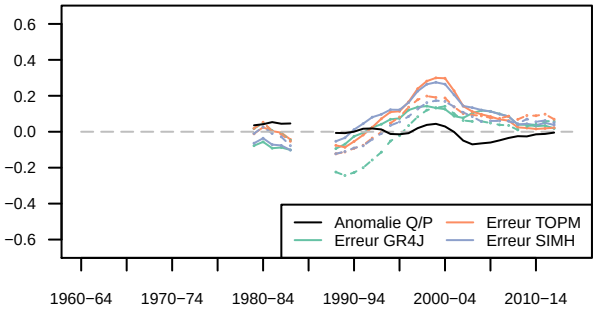


L'Ardres@Fismes (H6423010)

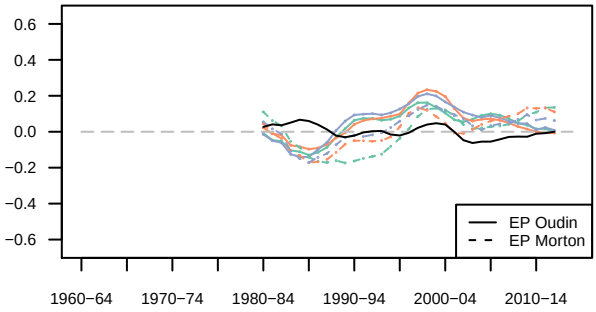



L'Aisne@Trosly−Breuil [Herant] (H6531011)
L'Oise@Hirson (H7021010)
Le Thon@Origny−en−Thierache (H7033010)
L'Oise@Monceau−sur−Oise (H7041010)
L'Oise@Origny−Sainte−Benoite (H7061010)
La Serre@Pont@Bucy (H7162010)
L'Oise@Sempigny (H7401010)
L'Aronde@Clairoix (H7423710)
La Brêche@Nogent−sur−Oise (H7602010)
Le Petit Therain@Saint−Omer−en−Chaussee (H7713010)
Anomalie Q/P
Erreur GR4J
Erreur TOPM
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

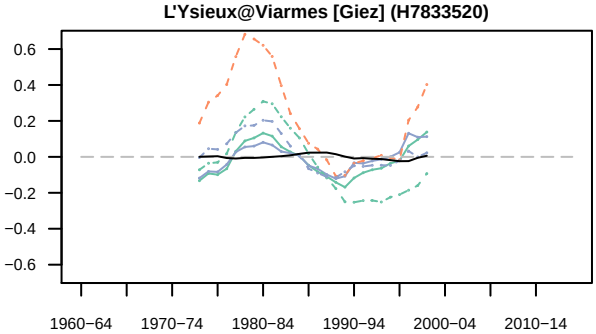
Le Therain@Beauvais (H7742010)



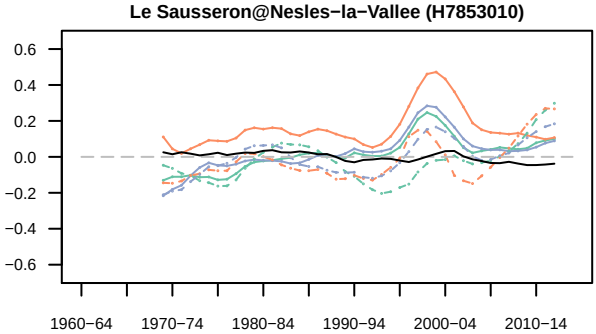
Le Therain@Maysel (H7742020)



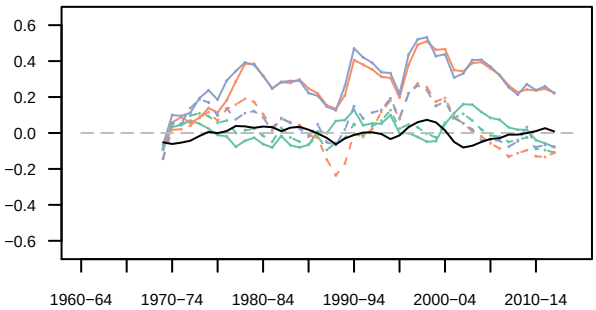
L'Ysieux@Viarmes [Giez] (H7833520)



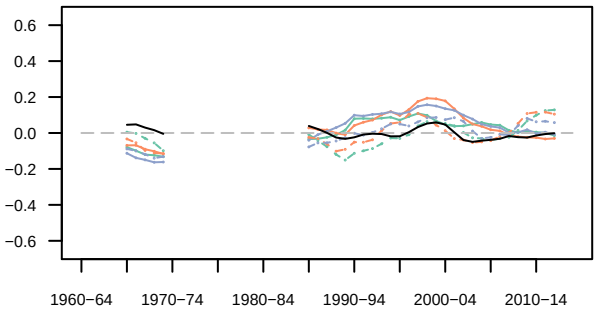
Le Sausseron@Nesles−la−Vallee (H7853010)



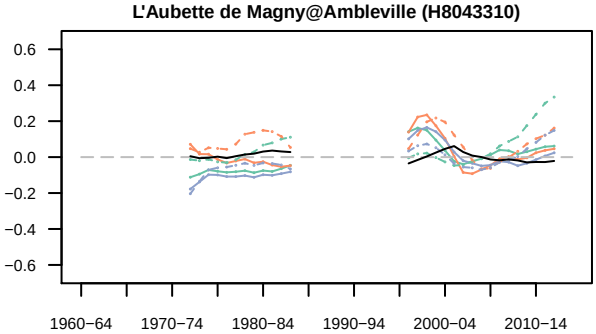
L'Epte@Gournay−en−Bray (H8012010)



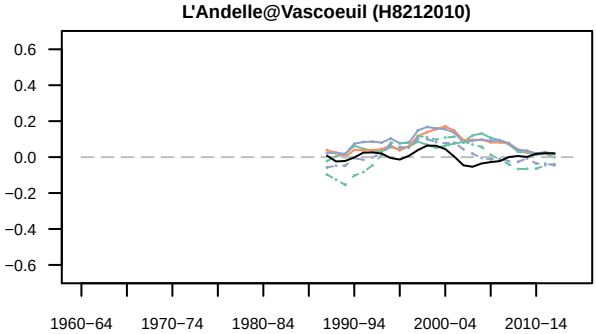
L'Epte@Fourges (H8042010)



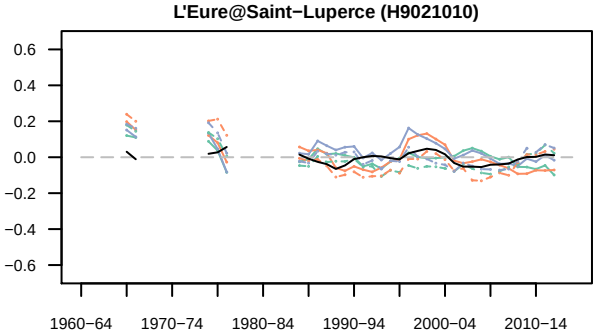
L'Aubette de Magny@Ambleville (H8043310)



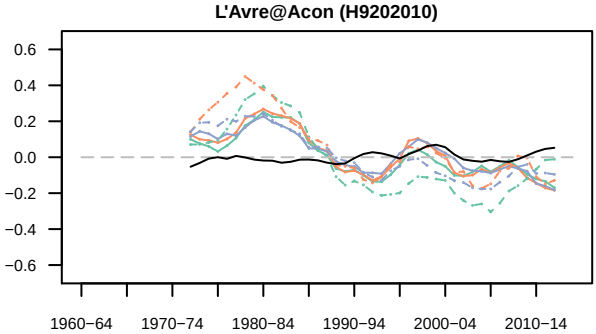
L'Andelle@Vascoeuil (H8212010)

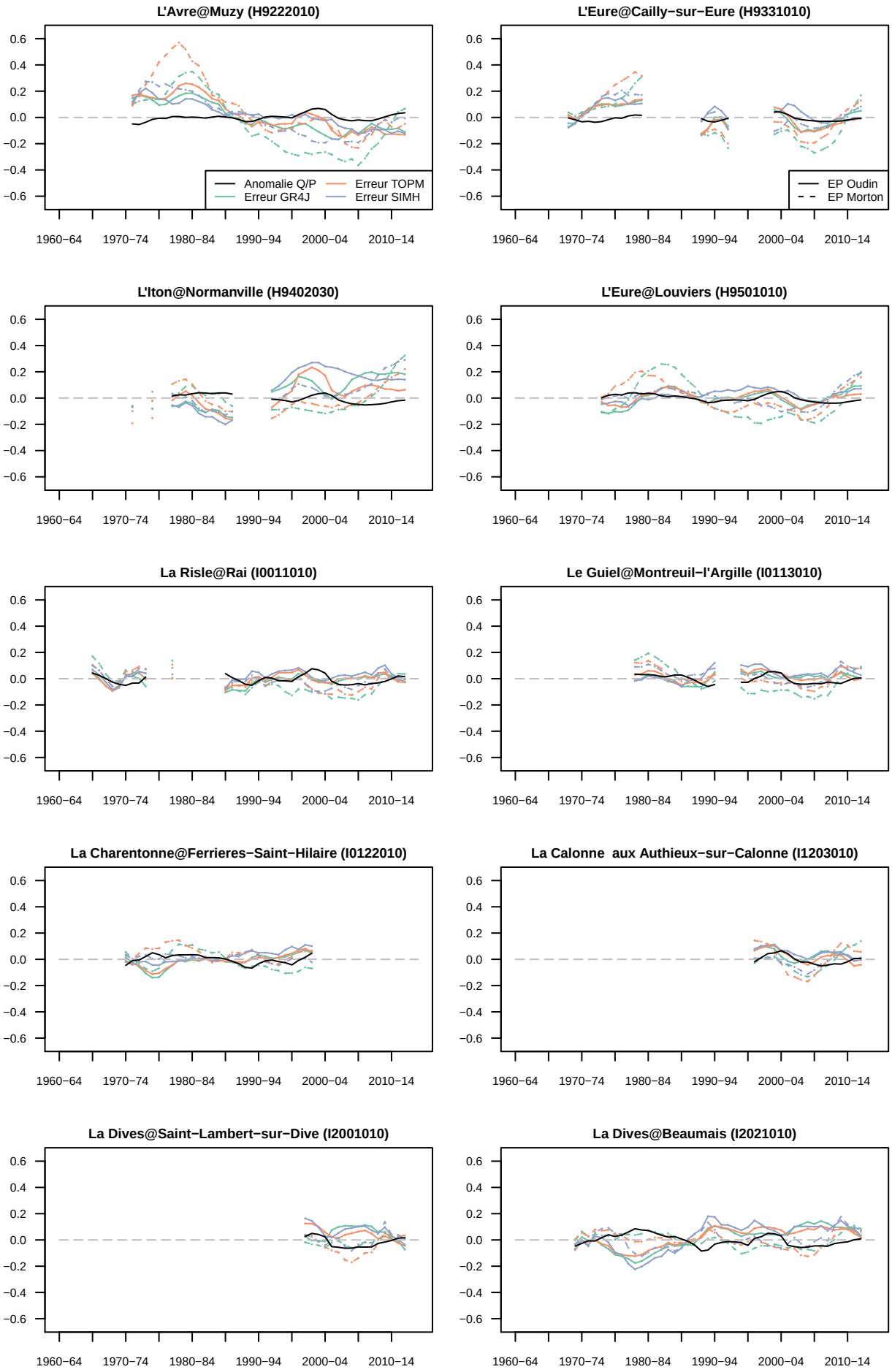


L'Eure@Saint−Luperce (H9021010)

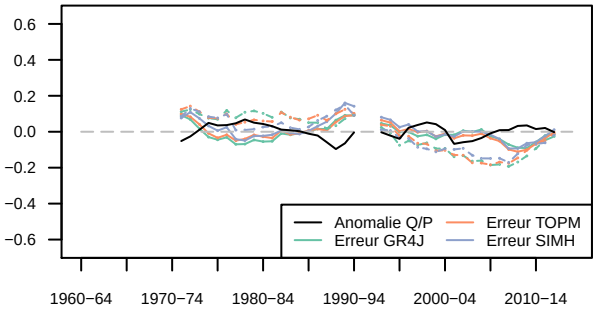


L'Avre@Acon (H9202010)

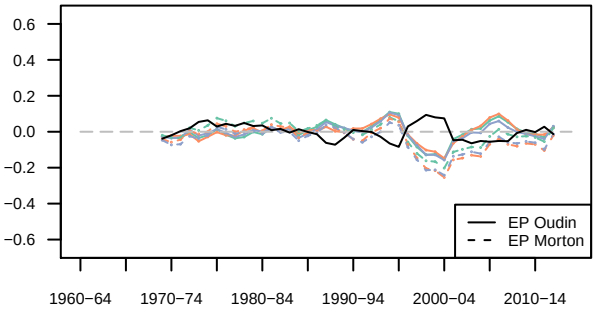



L'Avre@Muzy (H9222010)
L'Eure@Cailly−sur−Eure (H9331010)
L'Iton@Normanville (H9402030)
L'Eure@Louviers (H9501010)
La Risle@Rai (I0011010)
Le Guiel@Montreuil−l'Argille (I0113010)
La Charentonne@Ferrieres−Saint−Hilaire (I0122010)
La Calonne aux Authieux−sur−Calonne (I1203010)
La Dives@Saint−Lambert−sur−Dive (I2001010)
La Dives@Beaumais (I2021010)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

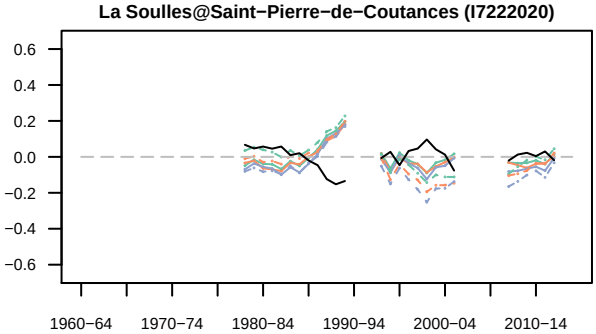
La Seulles@Tierceville (I4032010)



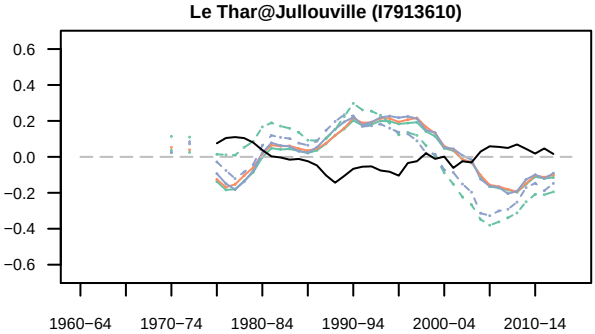
La Souleuvre@Carville (I5053010)



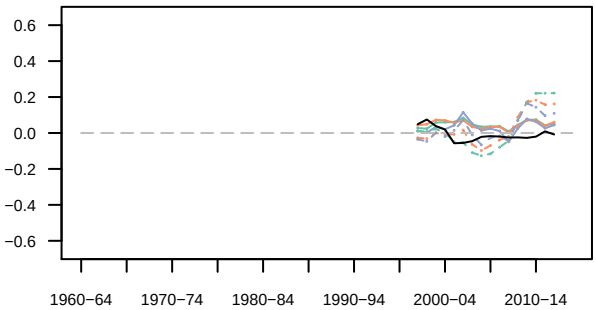
La Soulles@Saint−Pierre−de−Coutances (I7222020)



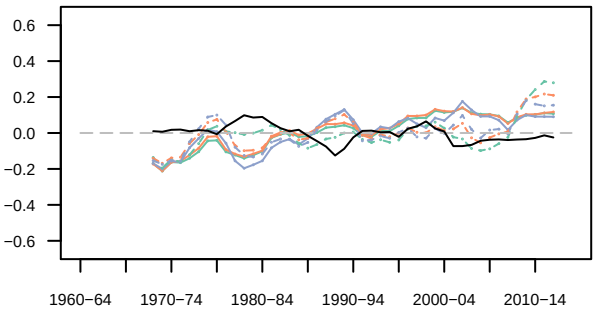
Le Thar@Jullouville (I7913610)



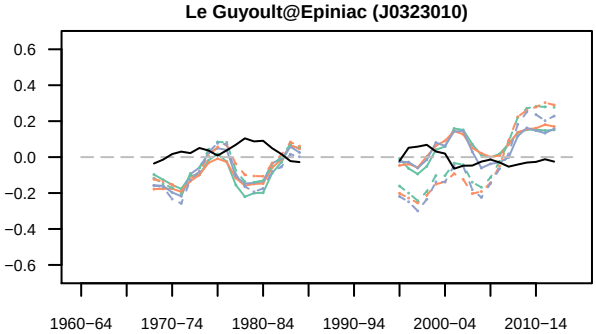
Le Nançon@Lecousse [Pont aux Anes] (J0014010)



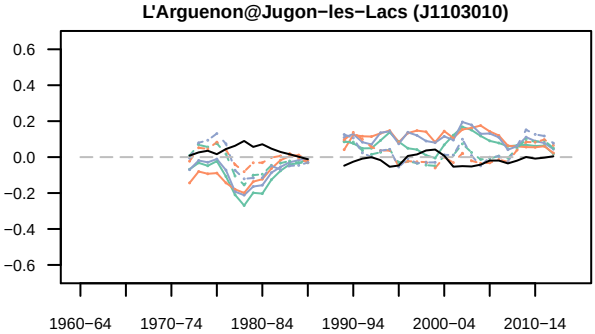
La Loysance@Saint−Ouen−la−Rouërie (J0144010)



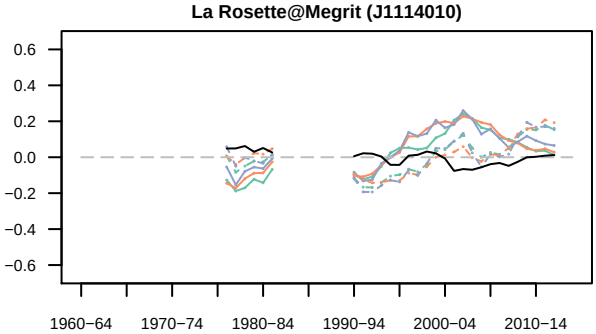
Le Guyoult@Epiniac (J0323010)



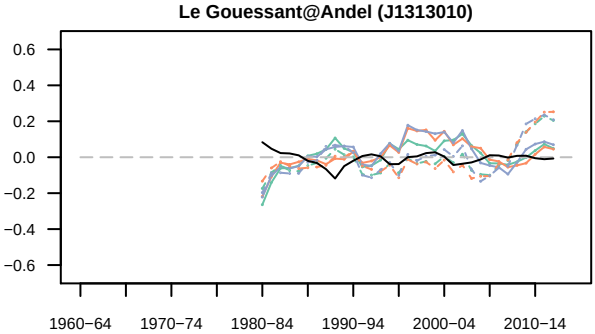
L'Arguenon@Jugon−les−Lacs (J1103010)

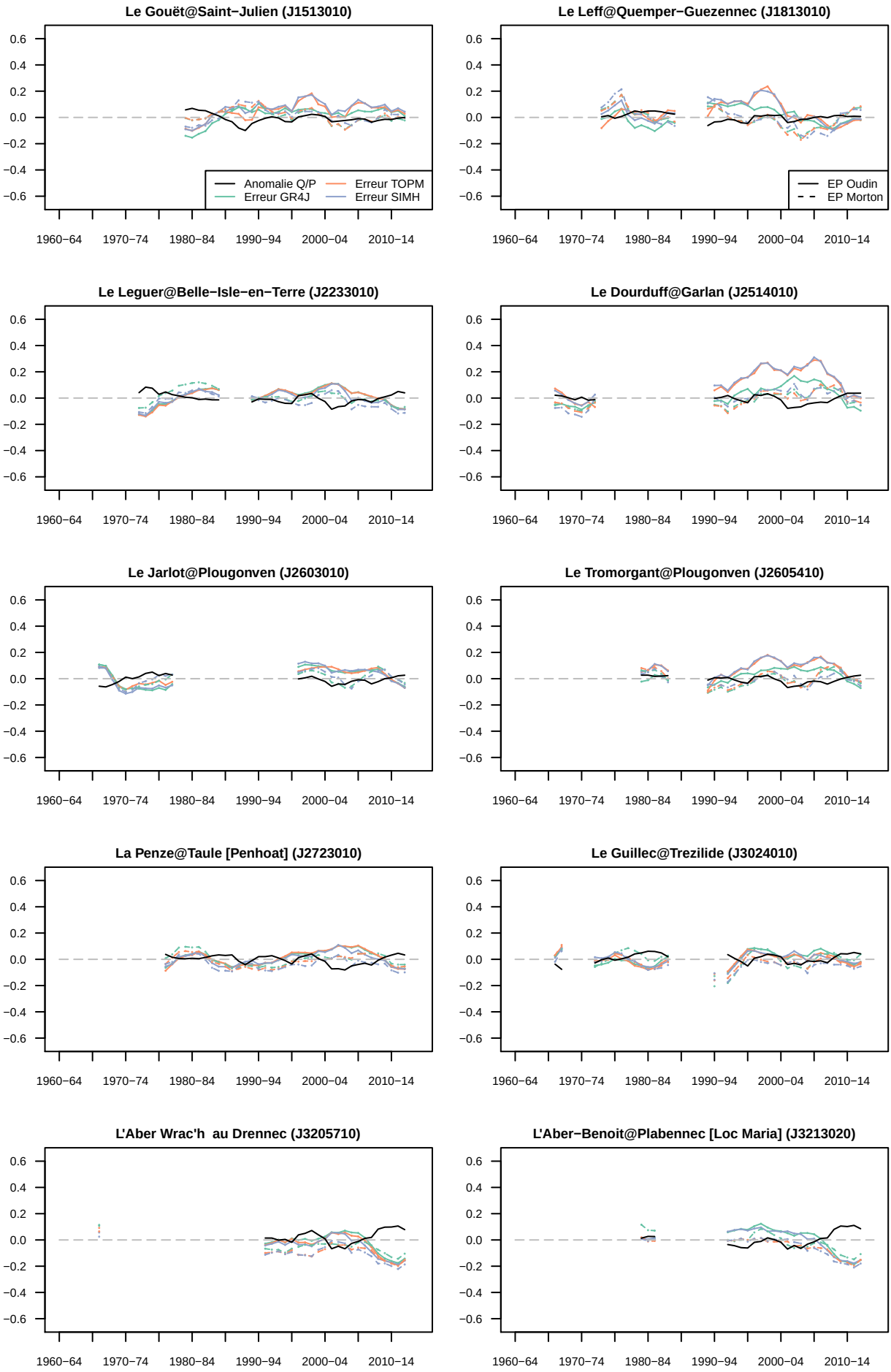


La Rosette@Megrit (J1114010)

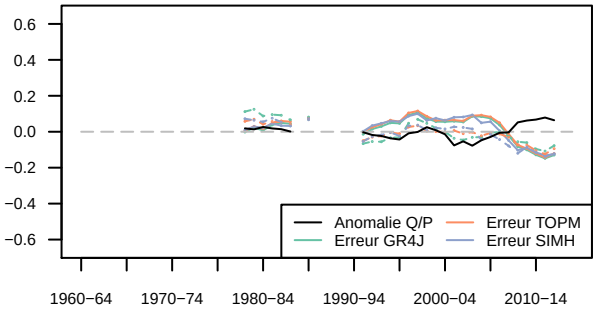


Le Gouessant@Andel (J1313010)

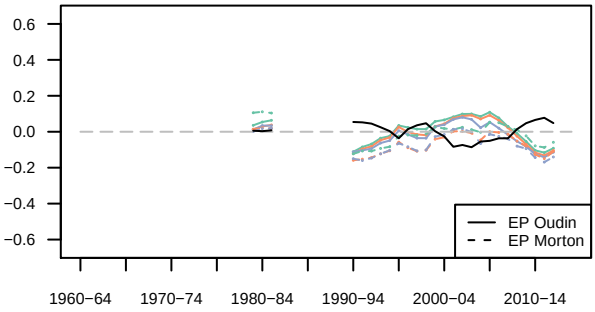



Le Gouët@Saint−Julien (J1513010)
Le Leff@Quemper−Guezennec (J1813010)
Le Leguer@Belle−Isle−en−Terre (J2233010)
Le Dourduff@Garlan (J2514010)
Le Jarlot@Plougonven (J2603010)
Le Tromorgant@Plougonven (J2605410)
La Penze@Taule [Penhoat] (J2723010)
Le Guillec@Trezilide (J3024010)
L'Aber Wrac'h au Drennec (J3205710)
L'Aber−Benoit@Plabennec [Loc Maria] (J3213020)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

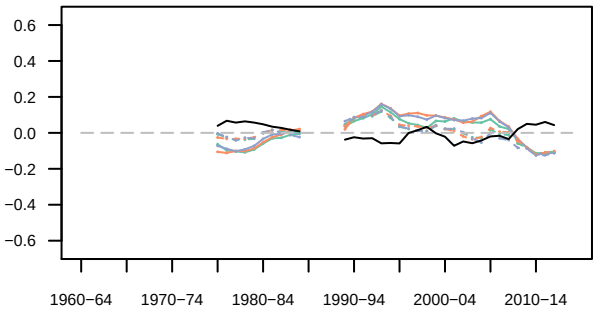
L'Aber Ildut@Breles [Keringar] (J3323020)



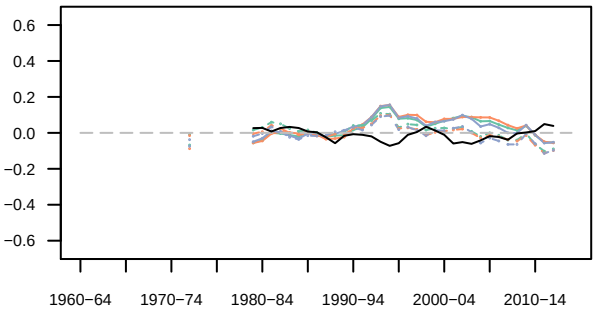
La Mignonne@Irvillac [Pont Mell] (J3514010)



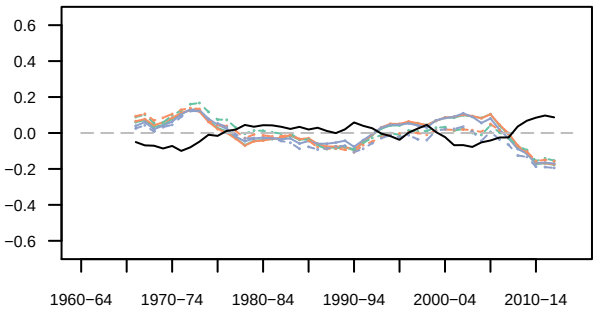
L'Aulne@Scrignac [Le Goask] (J3601810)



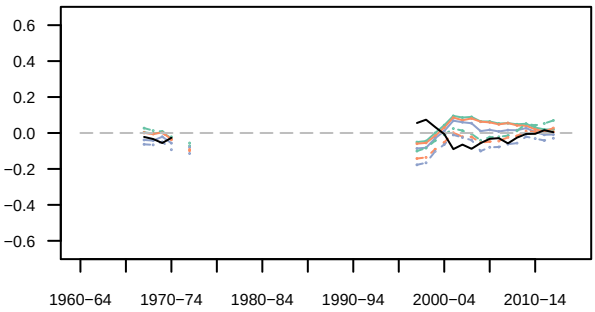
L'Hyeres@Trebrivan [Pont Neuf] (J3713010)



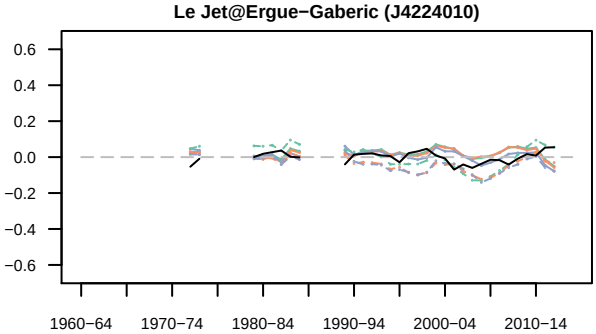
La Douffine@Saint−Segal [Kerbriant] (J3834010)



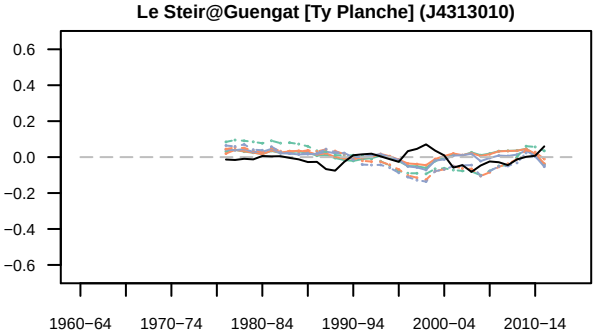
Le Langelin@Briec [Pont D 72] (J4214510)



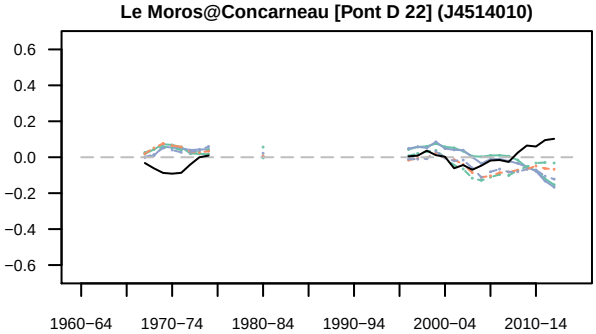
Le Jet@Ergue−Gaberic (J4224010)



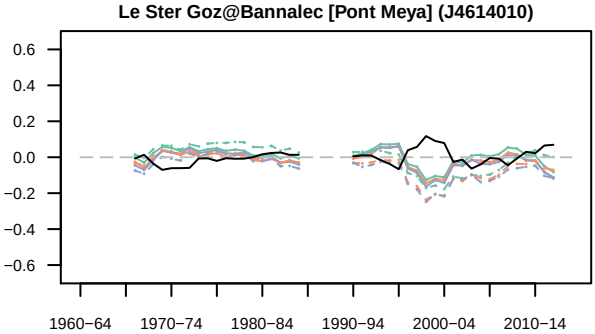
Le Steir@Guengat [Ty Planche] (J4313010)

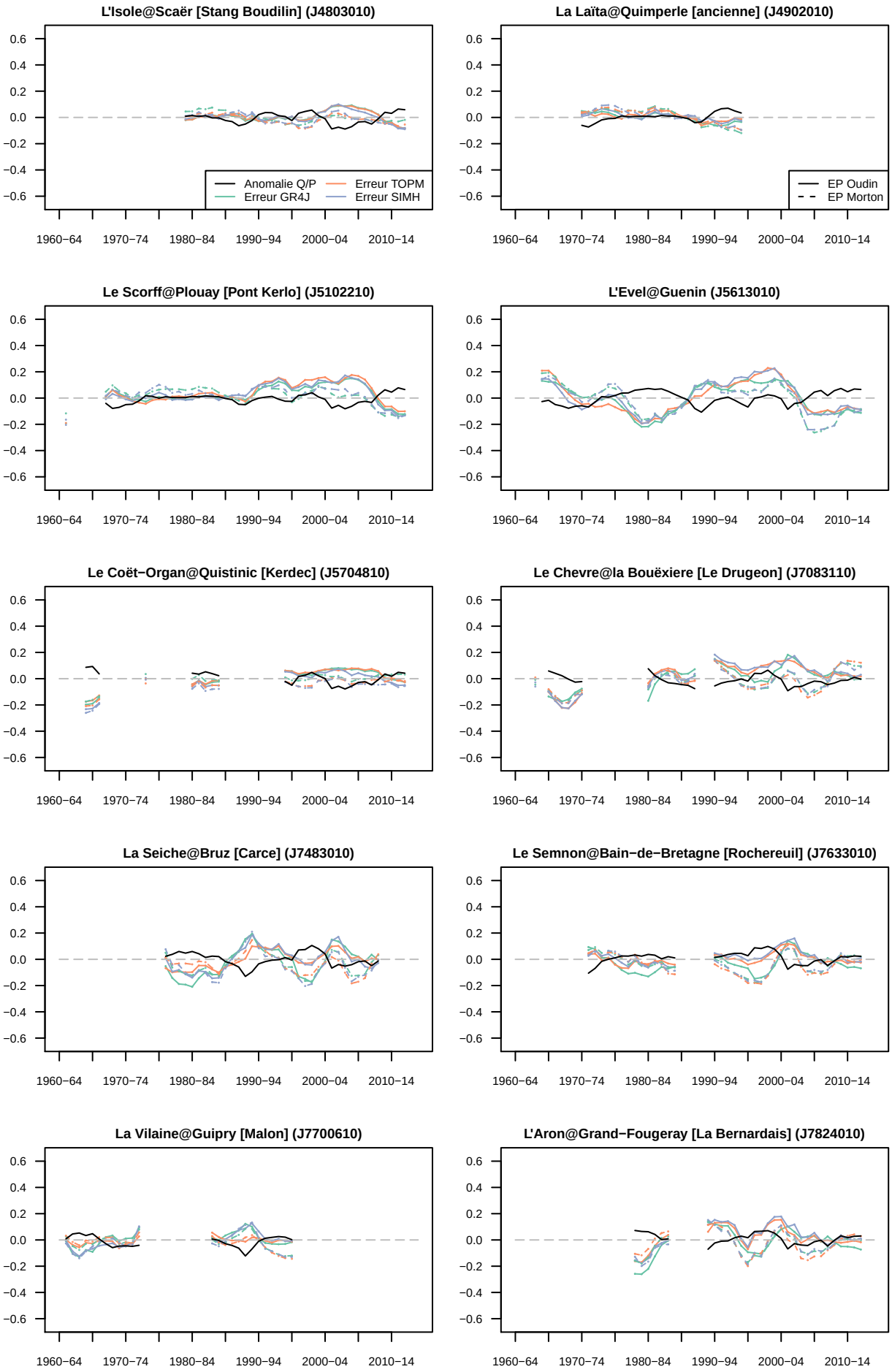


Le Moros@Concarneau [Pont D 22] (J4514010)

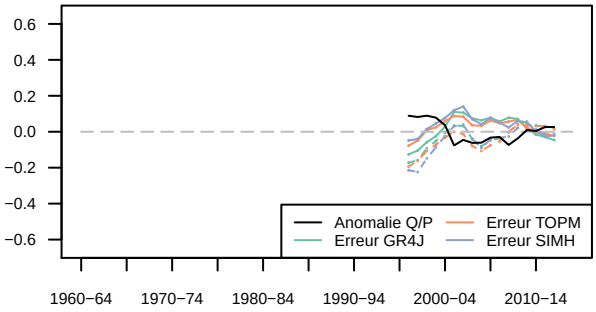


Le Ster Goz@Bannalec [Pont Meya] (J4614010)

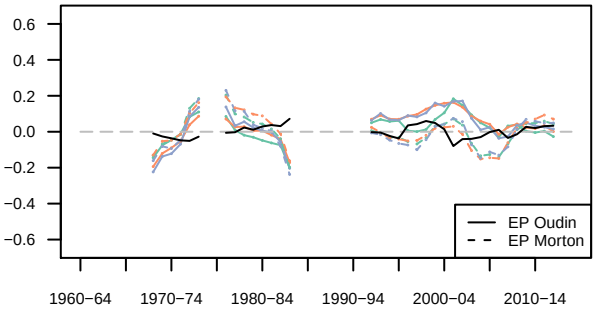



L'Isole@Scaër [Stang Boudilin] (J4803010)
La Laïta@Quimperle [ancienne] (J4902010)
Le Scorff@Plouay [Pont Kerlo] (J5102210)
L'Evel@Guenin (J5613010)
Le Coët−Organ@Quistinic [Kerdec] (J5704810)
Le Chevre@la Bouëxiere [Le Drugeon] (J7083110)
La Seiche@Bruz [Carce] (J7483010)
Le Semnon@Bain−de−Bretagne [Rochereuil] (J7633010)
La Vilaine@Guipry [Malon] (J7700610)
L'Aron@Grand−Fougeray [La Bernardais] (J7824010)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

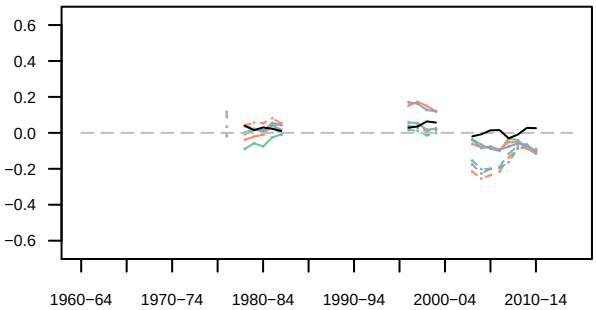
Le Canut Sud@Saint−Just [La riviere Colombel] (J7973010)



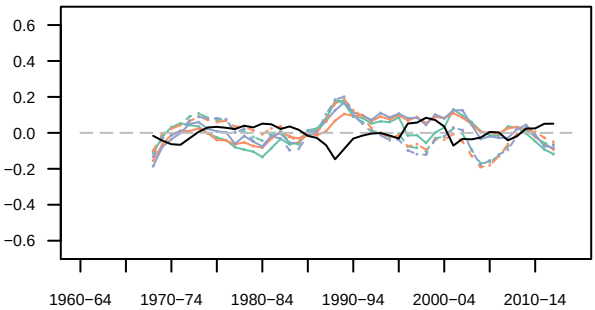
L'Yvel@Loyat [Pont D 129] (J8363110)



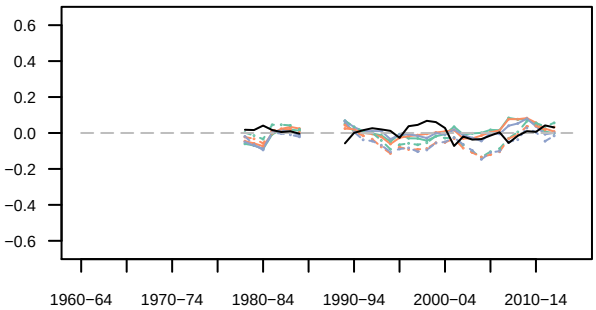
L'Aff@Paimpont [Pont du Secret] (J8602410)



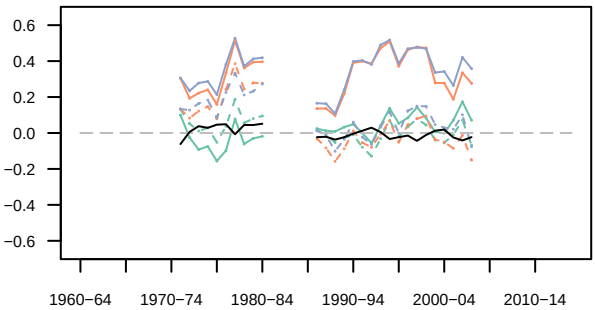
L'Aff@Quelneuc [La riviere] (J8632410)



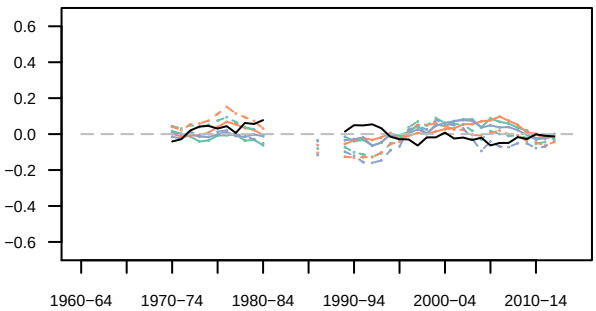
L'Arz@Molac [Le Qinquizio] (J8813010)



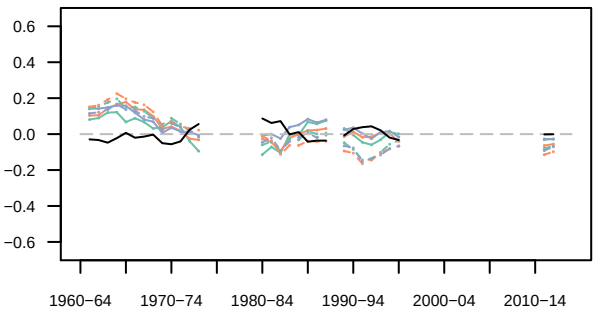
Le Bonson@Saint−Marcellin−en−Forez [Le Bled] (K0624510)



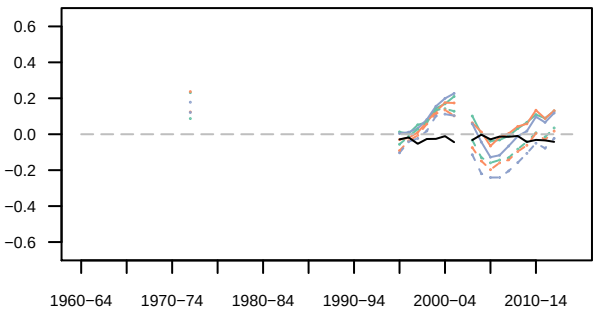
La Coise@Larajasse [Le Nezel] (K0663310)



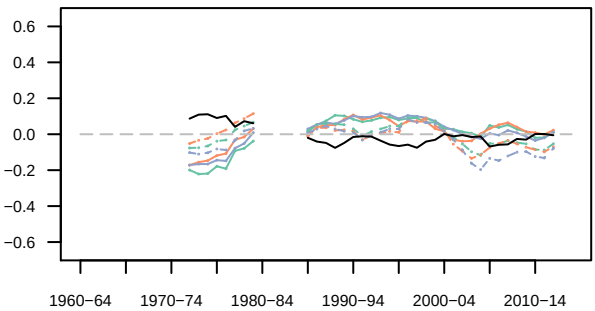
La Coise@Saint−Medard−en−Forez [Moulin BrÃ»le] (K0673310)

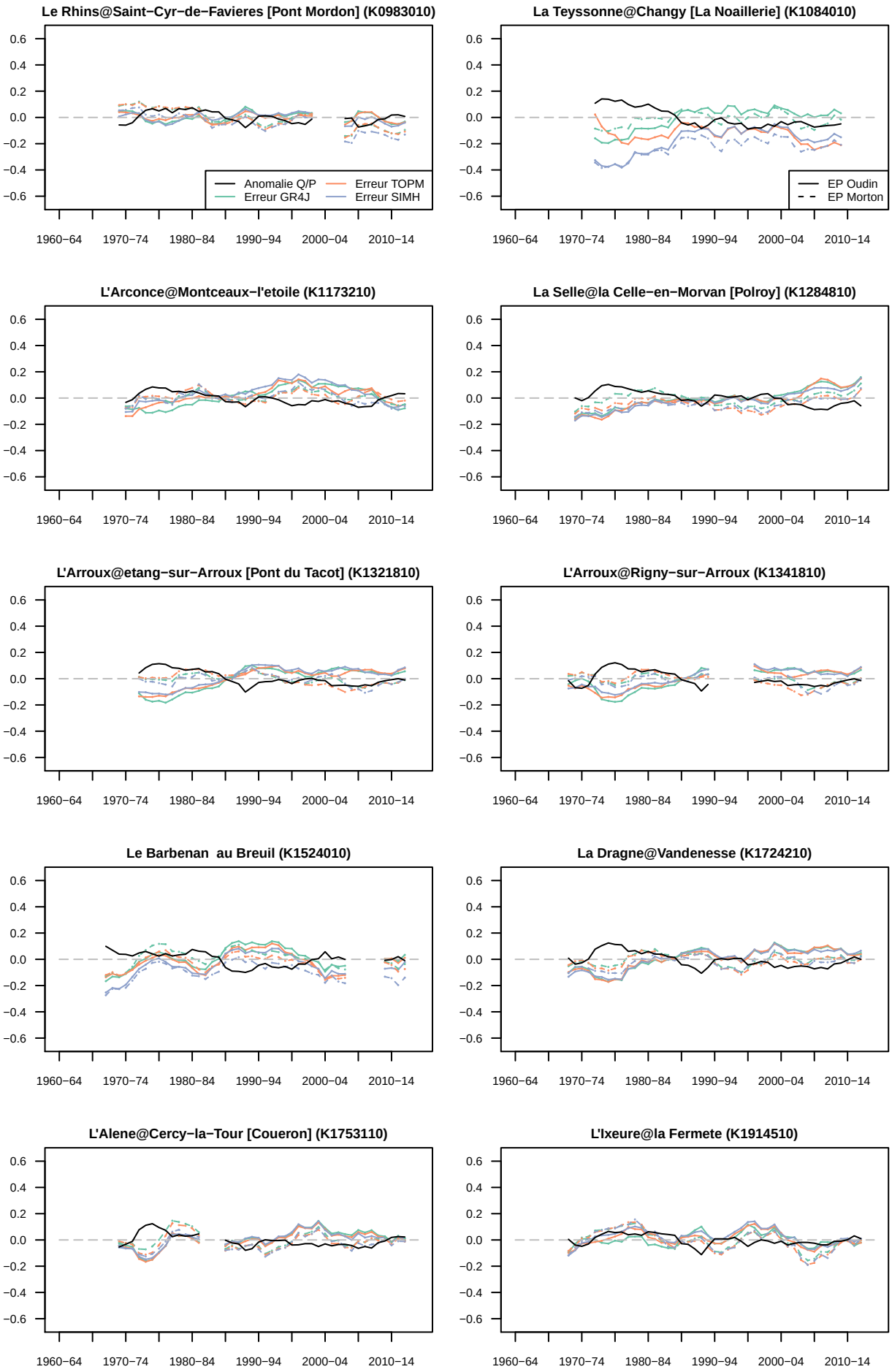


Le Chanasson@Civens [La riviere] (K0724510)

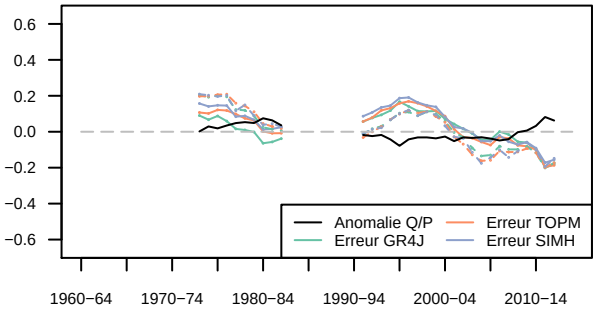


Le Gand@Neaux (K0974010)

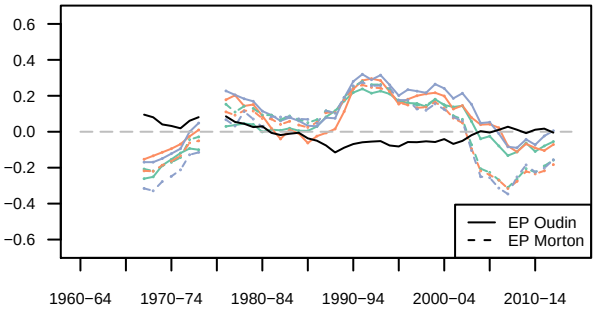



Le Rhins@Saint−Cyr−de−Favieres [Pont Mordon] (K0983010)
La Teyssonne@Changy [La Noaillerie] (K1084010)
L'Arconce@Montceaux−l'etoile (K1173210)
La Selle@la Celle−en−Morvan [Polroy] (K1284810)
L'Arroux@etang−sur−Arroux [Pont du Tacot] (K1321810)
L'Arroux@Rigny−sur−Arroux (K1341810)
Le Barbenan au Breuil (K1524010)
La Dragne@Vandenesse (K1724210)
L'Alene@Cercy−la−Tour [Coueron] (K1753110)
L'Ixeure@la Fermete (K1914510)
Anomalie Q/P
Erreur GR4J
Erreur TOPM
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

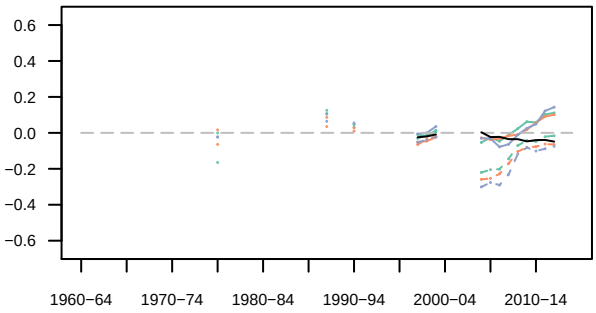
La Nievre d'Arzembouy@Poiseux [Poisson] (K1954010)



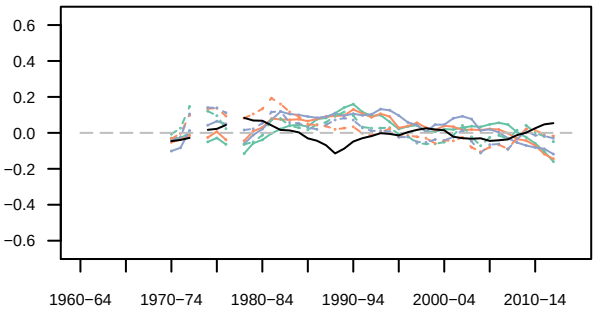
La Saunade@Pontaumur (K3264010)



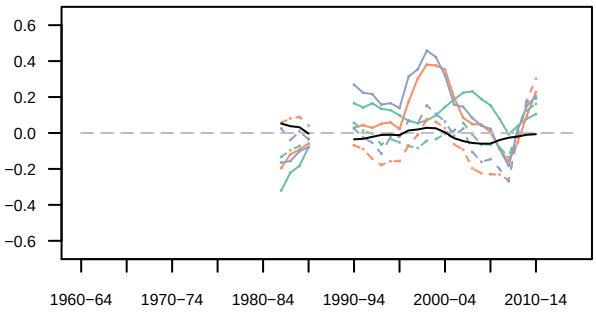
La Bouble@Chareil−Cintrat (K3373010)



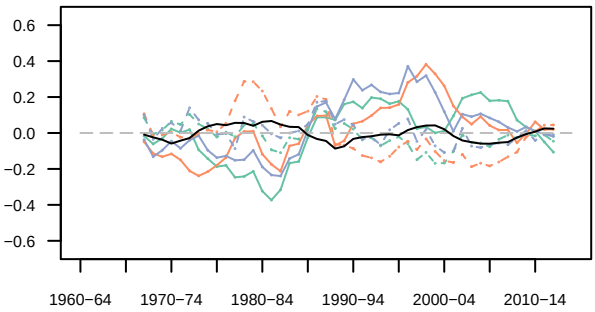
Le Nohain@Saint−Martin−sur−Nohain [Villiers] (K4094010)



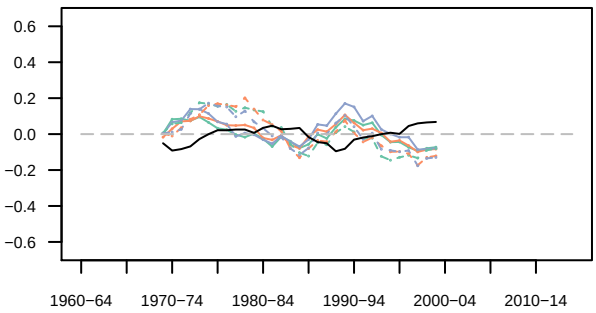
Le Dhuy@Saint−Cyr−en−Val [Gobson] (K4383110)



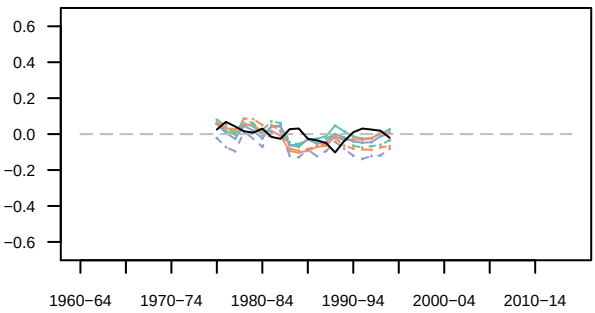
L'Ardoux@Lailly−en−Val (K4443010)



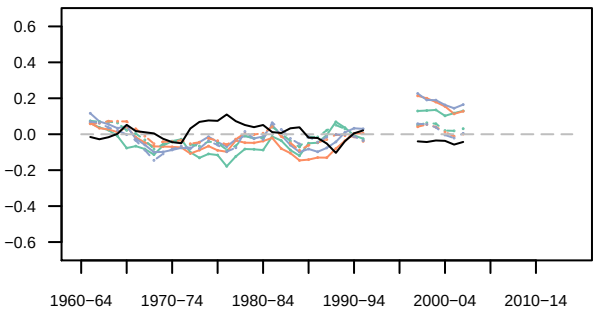
La Brenne@Villedômer [Bas−Villaumay] (K4873110)



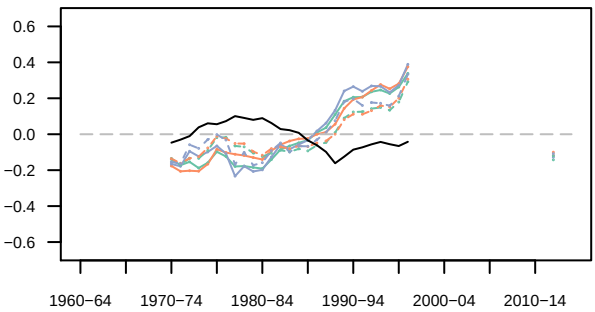
Le Cher@Chambonchard (K5090910)

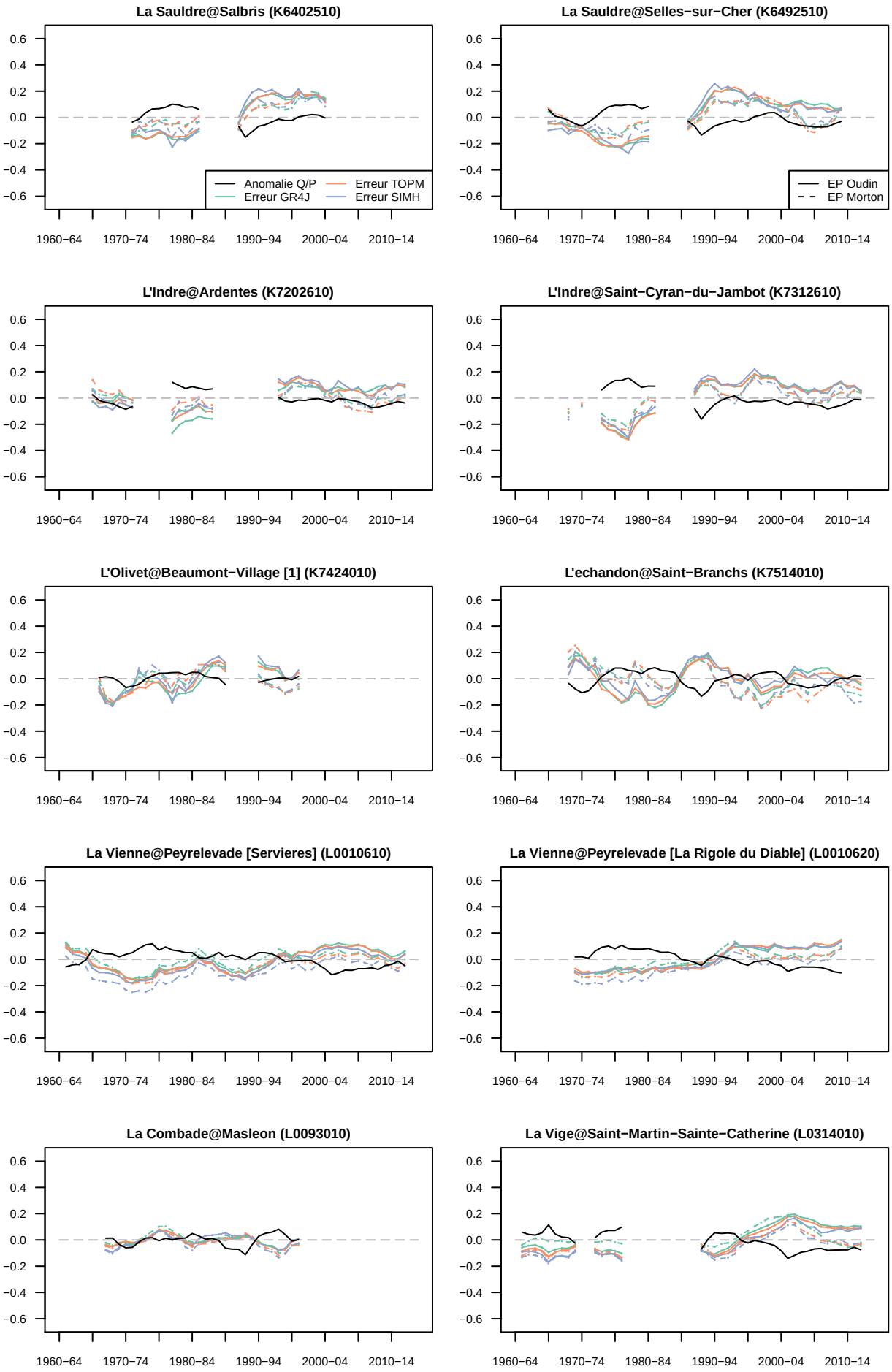


La Tardes@evaux−les−Bains (K5183010)

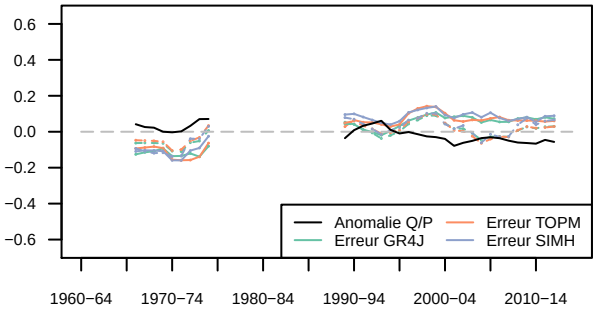


La Nere@Aubigny−sur−Nere (K6334010)

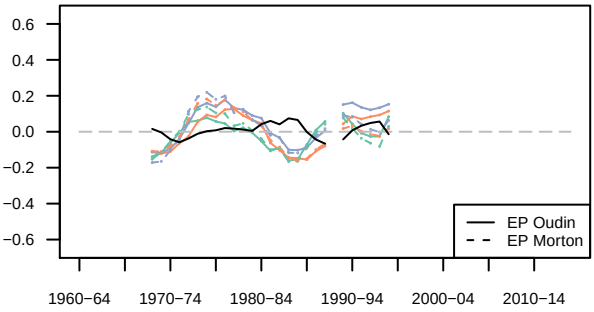



La Sauldre@Salbris (K6402510)
La Sauldre@Selles−sur−Cher (K6492510)
L'Indre@Ardentes (K7202610)
L'Indre@Saint−Cyran−du−Jambot (K7312610)
L'Olivet@Beaumont−Village [1] (K7424010)
L'echandon@Saint−Branchs (K7514010)
La Vienne@Peyrelevade [Servieres] (L0010610)
La Vienne@Peyrelevade [La Rigole du Diable] (L0010620)
La Combade@Masleon (L0093010)
La Vige@Saint−Martin−Sainte−Catherine (L0314010)
Anomalie Q/P
Erreur GR4J
Erreur TOPM
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

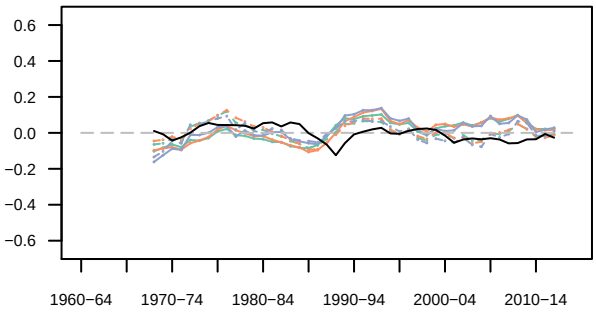
La Briance@Condat−sur−Vienne [Chambon Veyrinas] (L0563010)



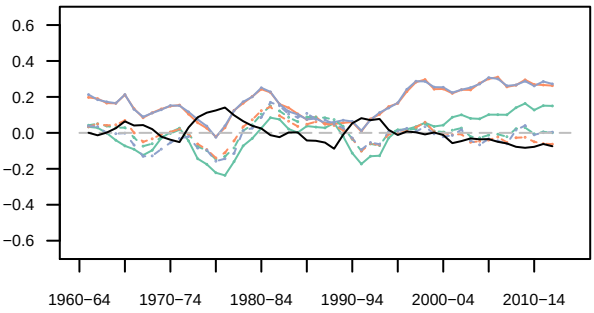
L'Aixette@Aixe−sur−Vienne (L0624010)



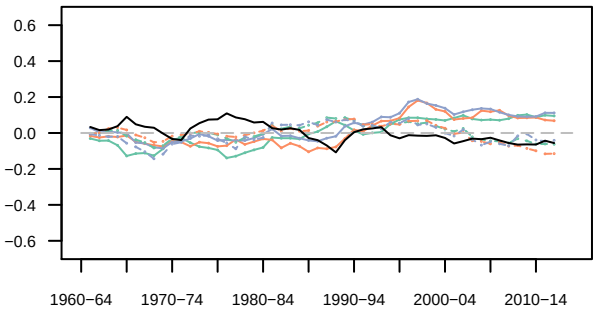
La Glane@Saint−Junien [Le Derot] (L0813010)



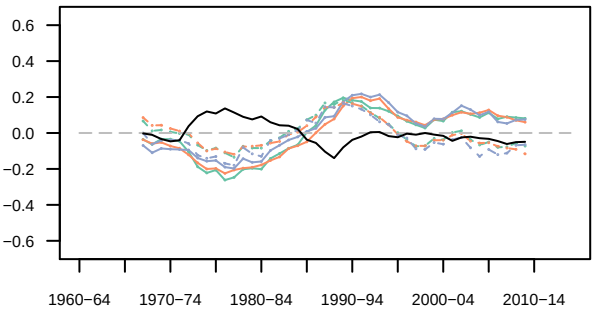
La Rozeille@Moutier−Rozeille [Aubusson] (L4033010)



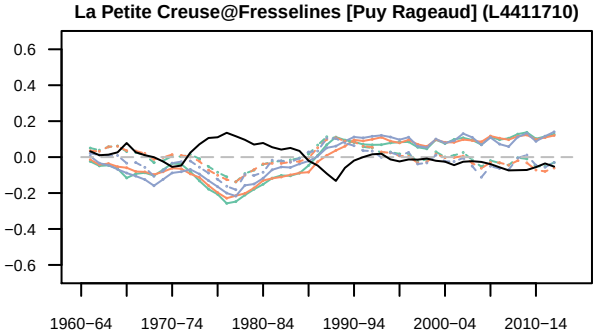
La Creuse@Fresselines (L4220710)



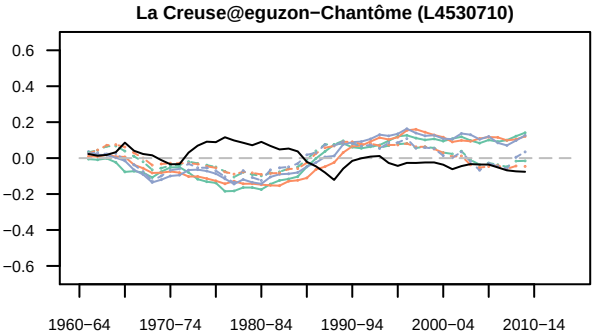
La Petite Creuse@Genouillac (L4321710)



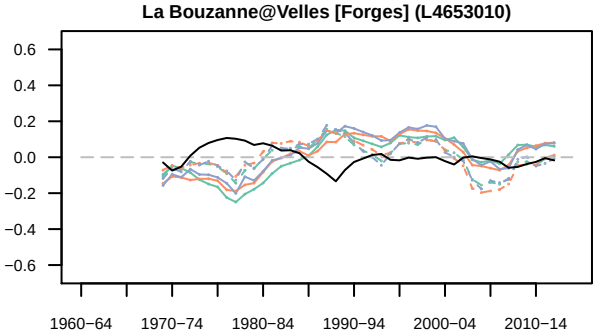
La Petite Creuse@Fresselines [Puy Rageaud] (L4411710)



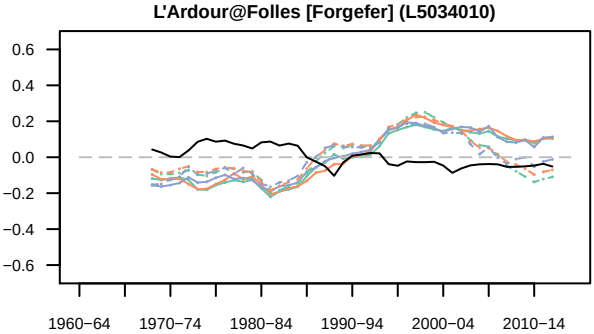
La Creuse@eguzon−Chantôme (L4530710)

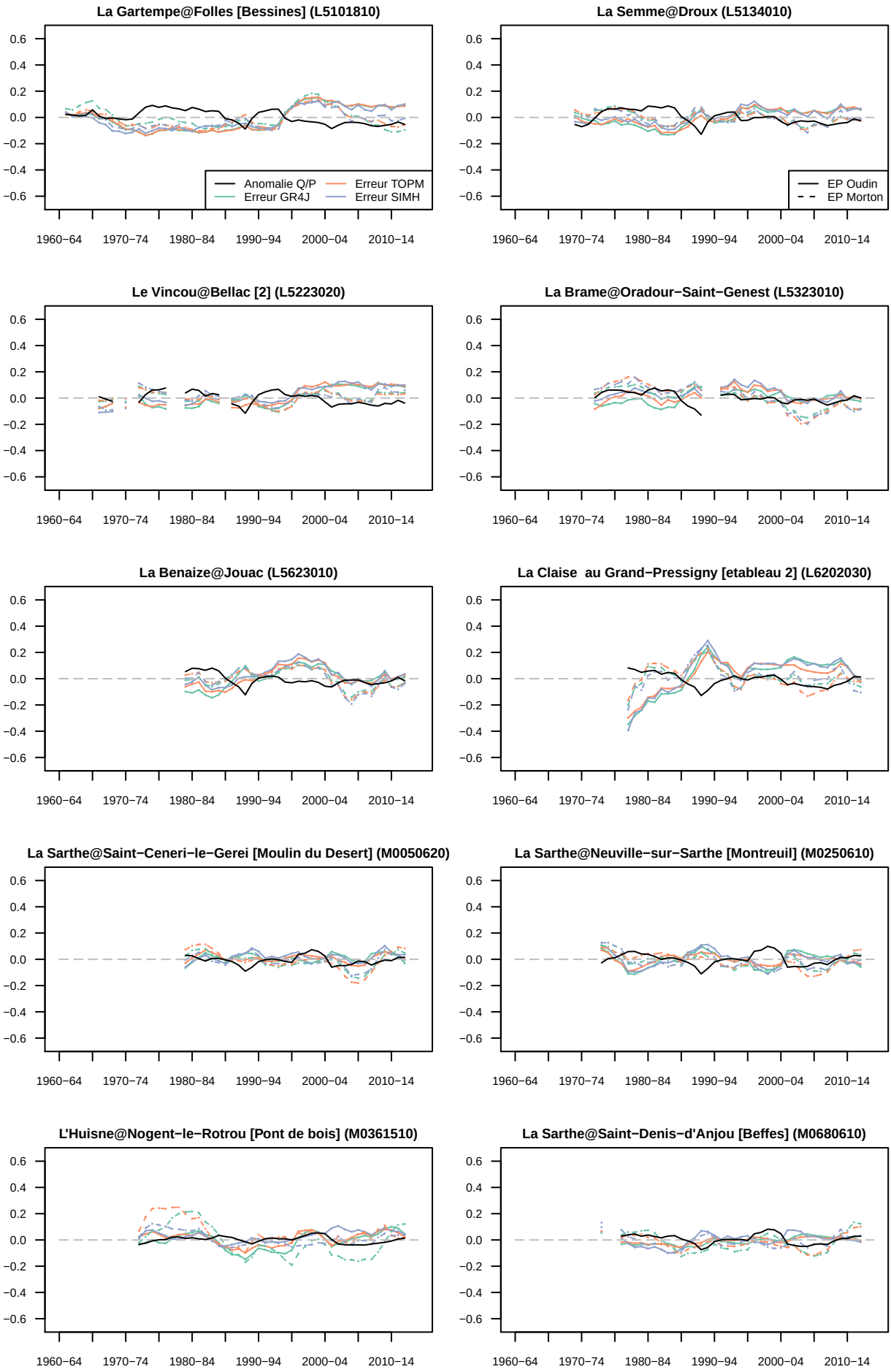


La Bouzanne@Velles [Forges] (L4653010)

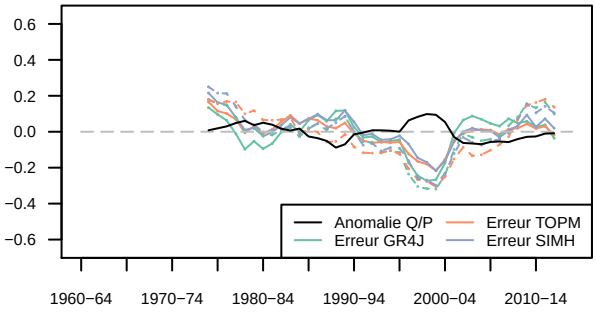


L'Ardour@Folles [Forgefer] (L5034010)

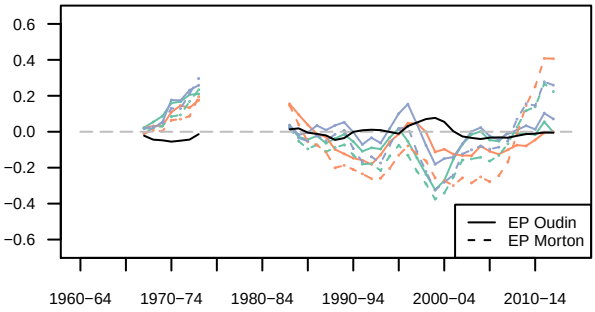



La Gartempe@Folles [Bessines] (L5101810)
La Semme@Droux (L5134010)
Le Vincou@Bellac [2] (L5223020)
La Brame@Oradour−Saint−Genest (L5323010)
La Benaize@Jouac (L5623010)
La Claise au Grand−Pressigny [etableau 2] (L6202030)
La Sarthe@Saint−Ceneri−le−Gerei [Moulin du Desert] (M0050620)
La Sarthe@Neuville−sur−Sarthe [Montreuil] (M0250610)
L'Huisne@Nogent−le−Rotrou [Pont de bois] (M0361510)
La Sarthe@Saint−Denis−d'Anjou [Beffes] (M0680610)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

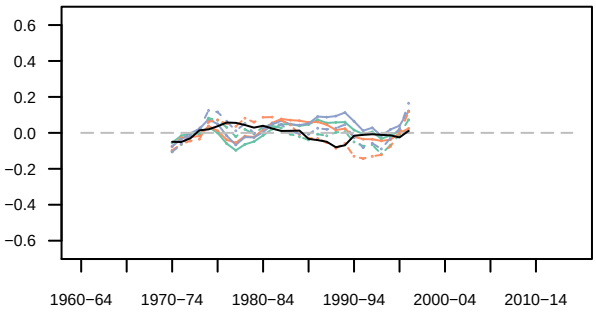
L'Ozanne@Trizay−les−Bonneval [Premoteux] (M1034020)



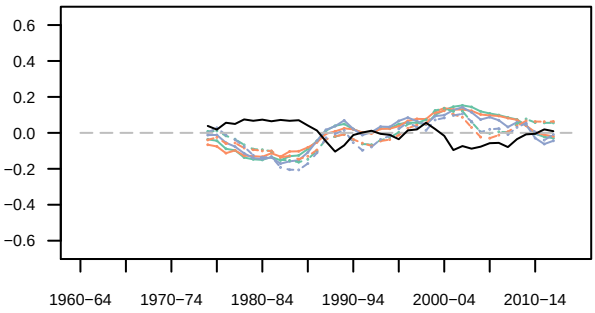
Le Loir@Saint−Maur−sur−le−Loir (M1041610)



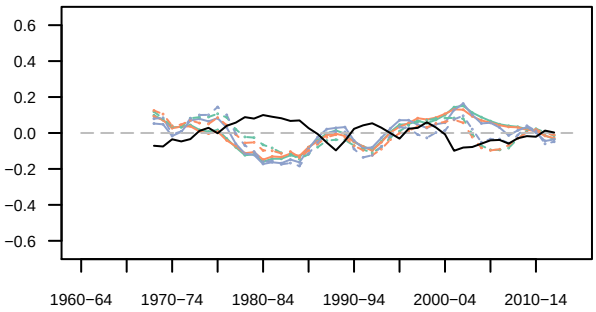
Le Couëtron@Souday [Glatigny] (M1214010)



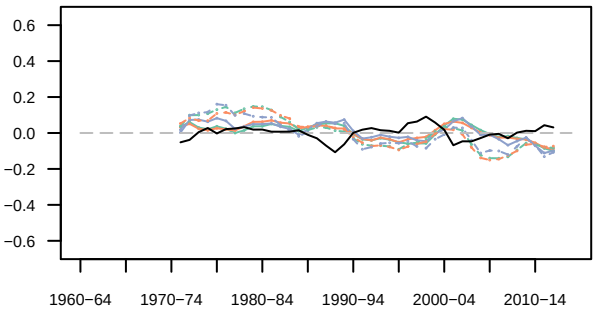
L'Aron@Moulay (M3253110)



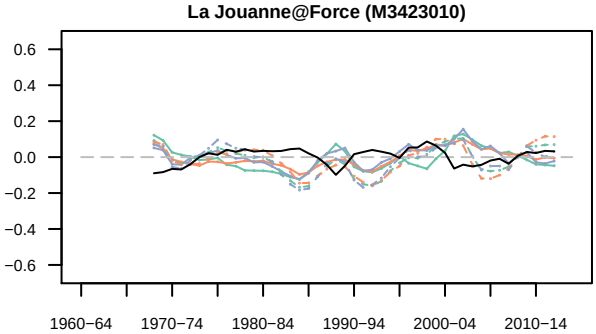
L'Ernee@Andouille [Les Vaugeois] (M3323010)



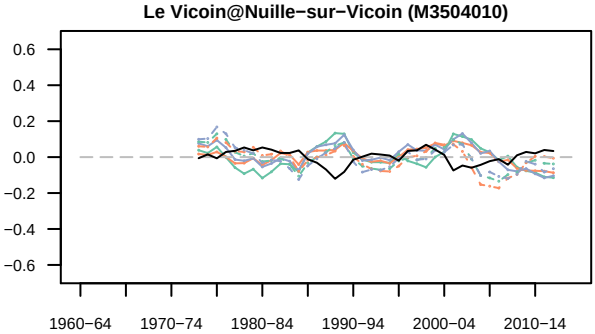
La Mayenne@l'Huisserie [Bonne] (M3340910)



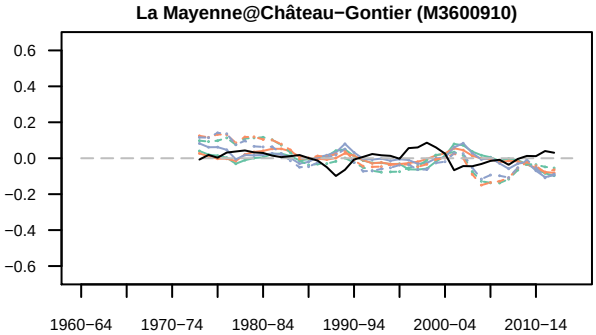
La Jouanne@Force (M3423010)



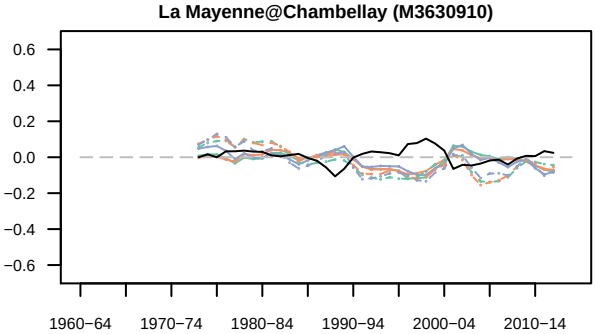
Le Vicoin@Nuille−sur−Vicoin (M3504010)

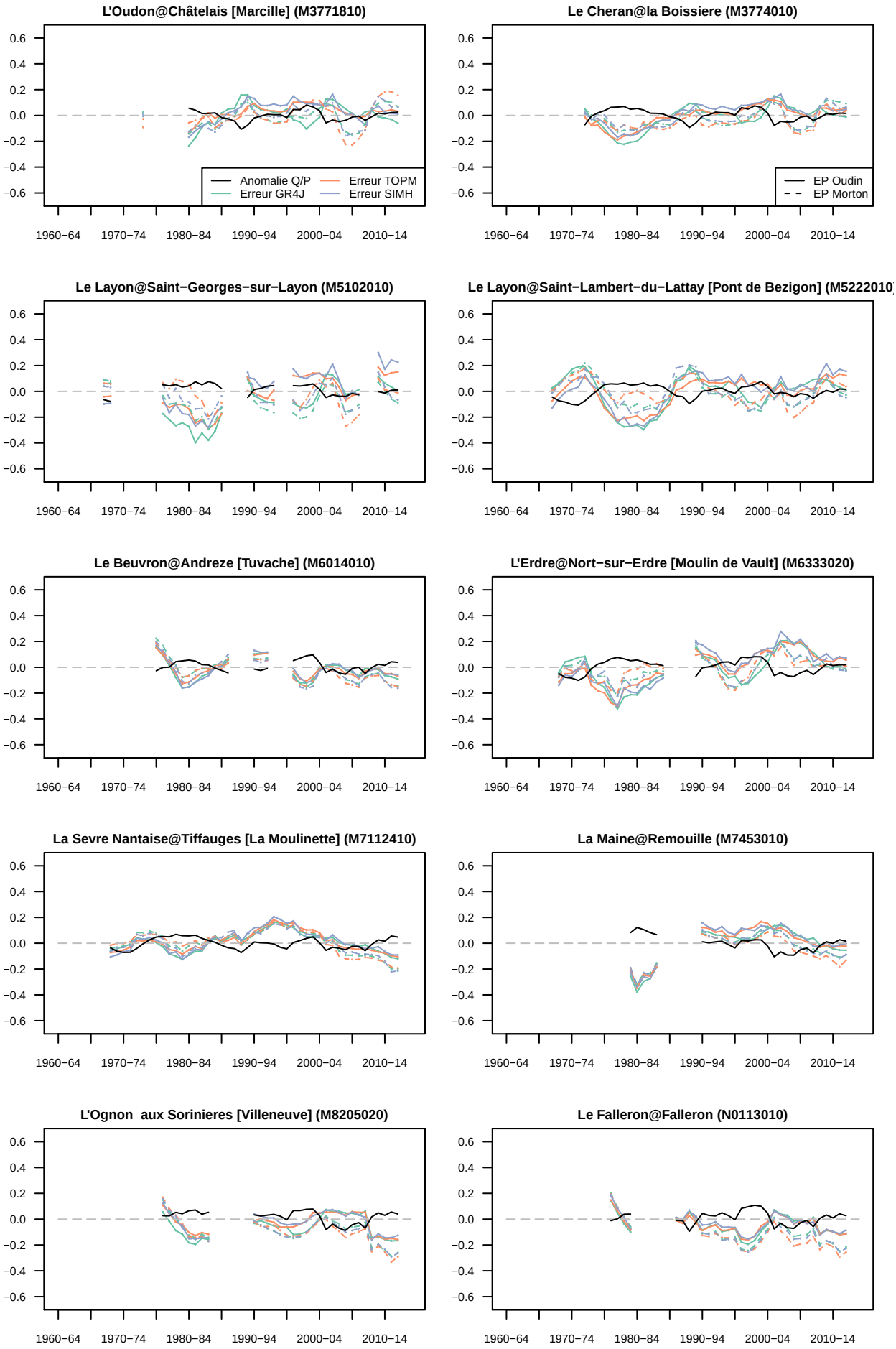


La Mayenne@Château−Gontier (M3600910)

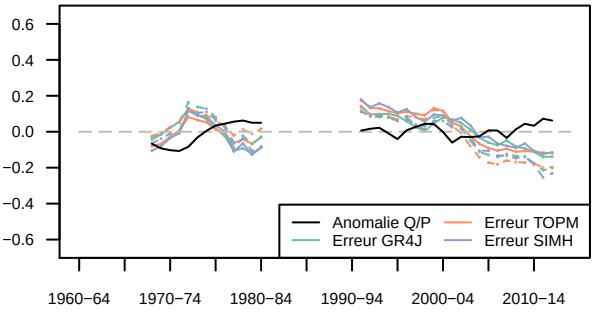


La Mayenne@Chambellay (M3630910)

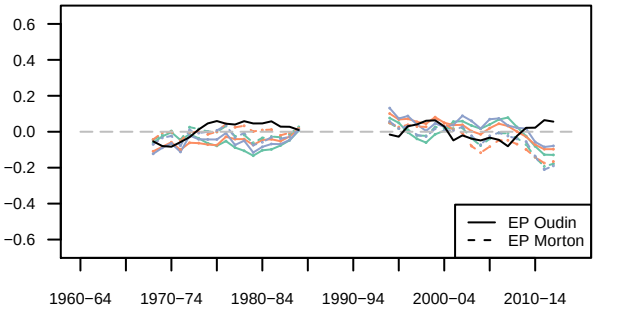



L'Oudon@Châtelais [Marcille] (M3771810)
Le Cheran@la Boissiere (M3774010)
Le Layon@Saint−Georges−sur−Layon (M5102010)
Le Layon@Saint−Lambert−du−Lattay [Pont de Bezigon] (M5222010)
Le Beuvron@Andreze [Tuvache] (M6014010)
L'Erdre@Nort−sur−Erdre [Moulin de Vault] (M6333020)
La Sevre Nantaise@Tiffauges [La Moulinette] (M7112410)
La Maine@Remouille (M7453010)
L'Ognon aux Sorinieres [Villeneuve] (M8205020)
Le Falleron@Falleron (N0113010)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

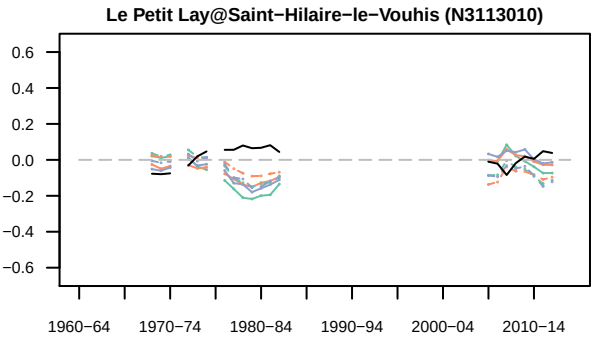
Le Grand Lay@Saint−Prouant [Monsireigne] (N3001610)



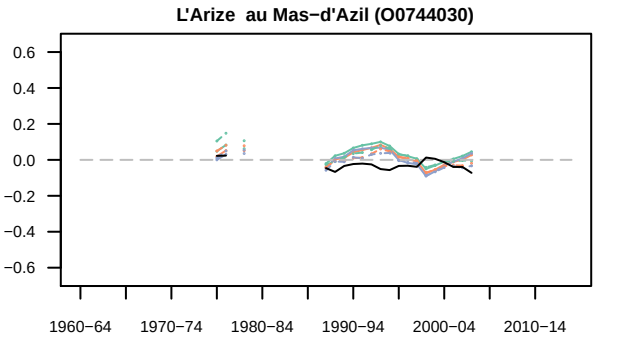
Le Louing@Chantonnay [St−Philbert du Pont Charrault] (N3024010)



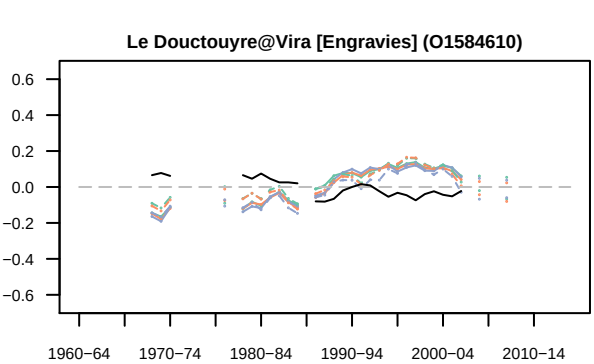
Le Petit Lay@Saint−Hilaire−le−Vouhis (N3113010)



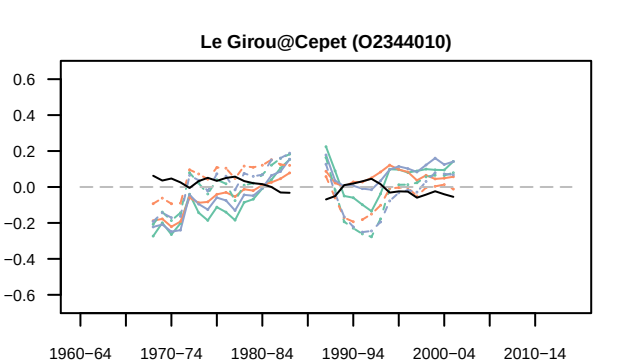
L'Arize au Mas−d'Azil (O0744030)



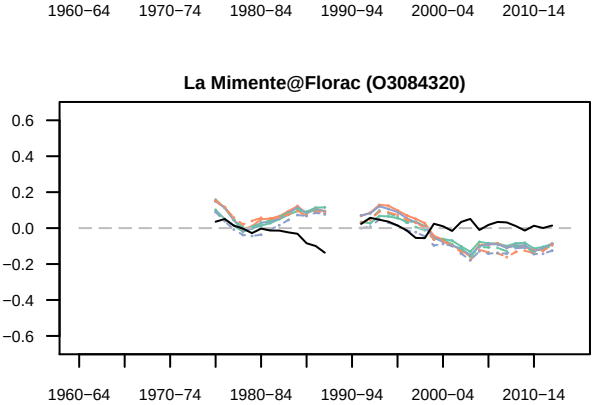
Le Douctouyre@Vira [Engravies] (O1584610)



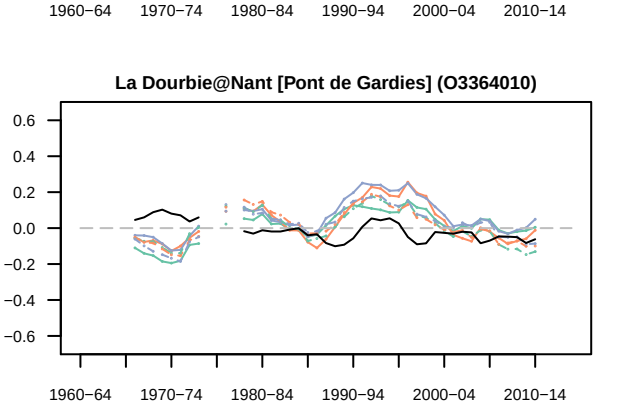
Le Girou@Cepet (O2344010)



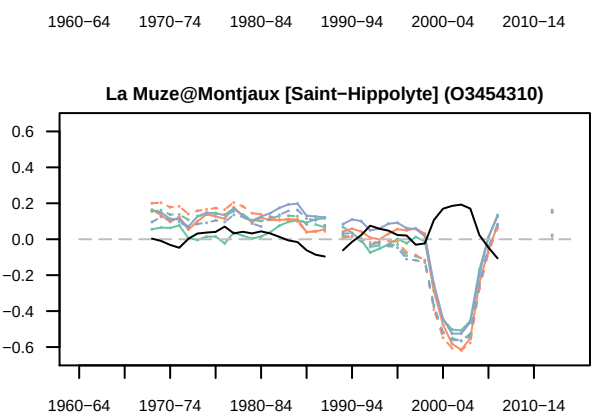
La Mimente@Florac (O3084320)



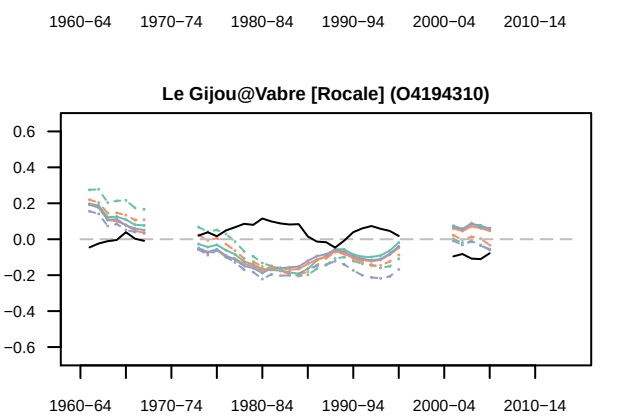
La Dourbie@Nant [Pont de Gardies] (O3364010)

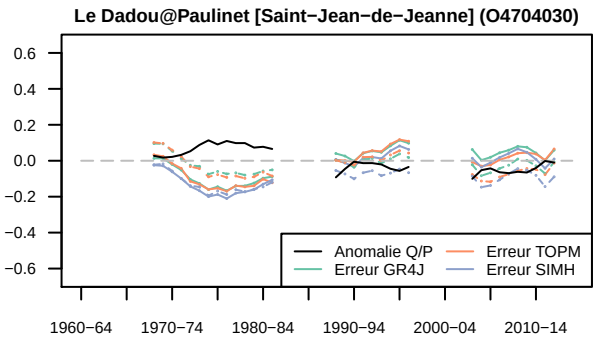


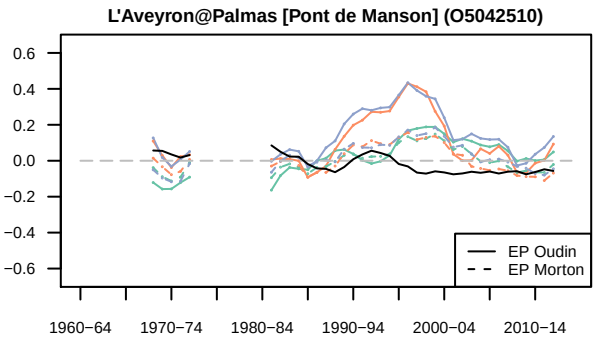
La Muze@Montjaux [Saint−Hippolyte] (O3454310)

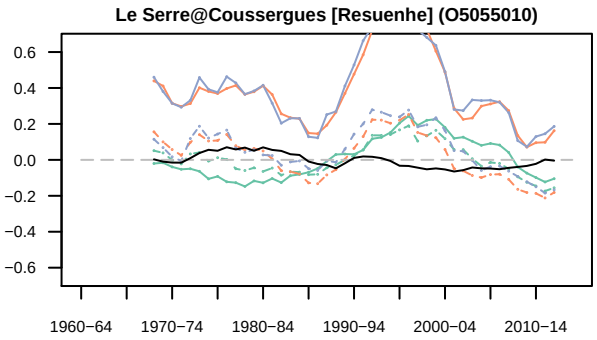


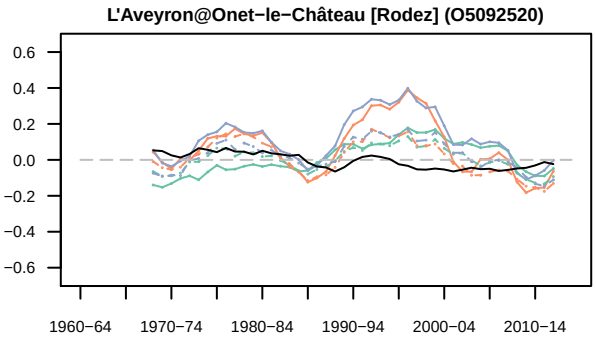
Le Gijou@Vabre [Rocale] (O4194310)

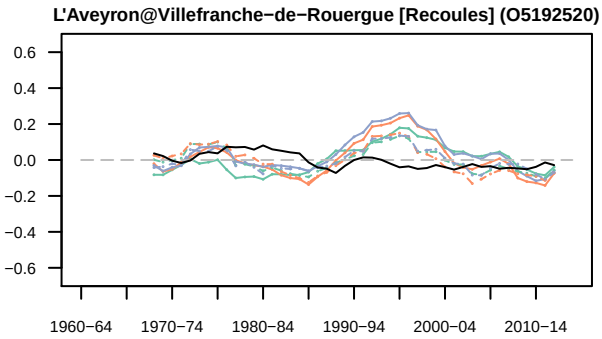


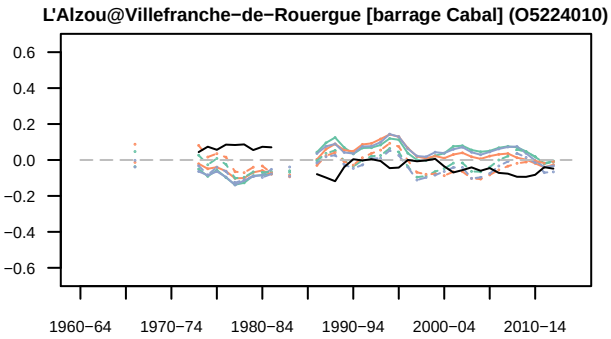

Le Dadou@Paulinet [Saint−Jean−de−Jeanne] (O4704030)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

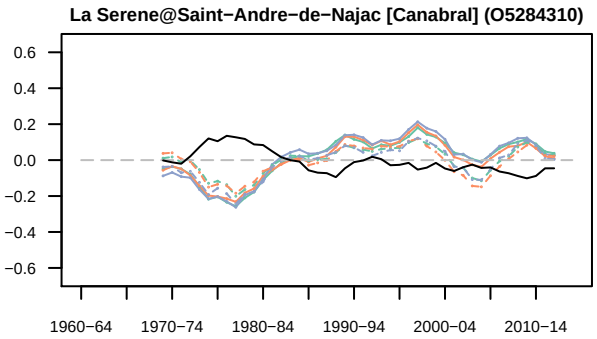

L'Aveyron@Palmas [Pont de Manson] (O5042510)
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

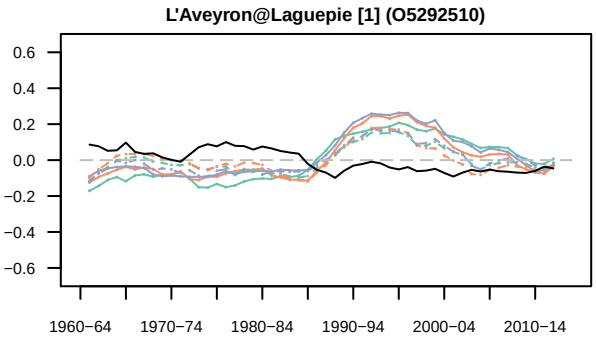

Le Serre@Coussergues [Resuenhe] (O5055010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

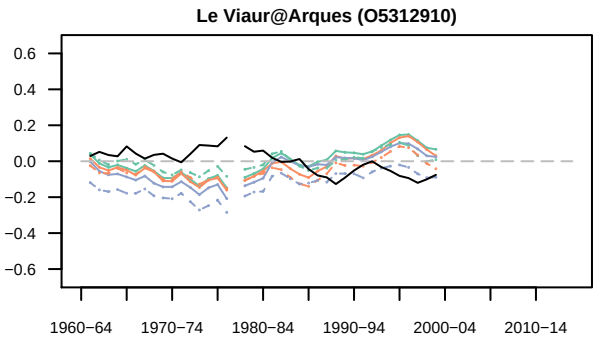

L'Aveyron@Onet−le−Château [Rodez] (O5092520)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

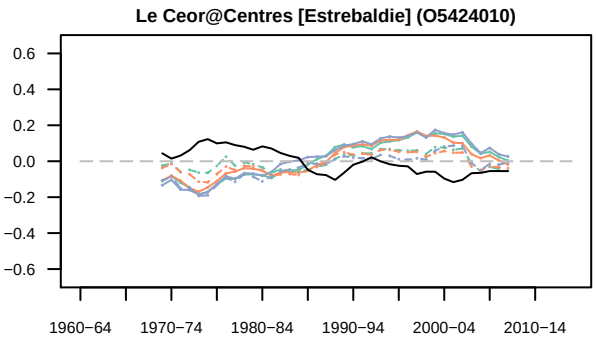

L'Aveyron@Villefranche−de−Rouergue [Recoules] (O5192520)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


L'Alzou@Villefranche−de−Rouergue [barrage Cabal] (O5224010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

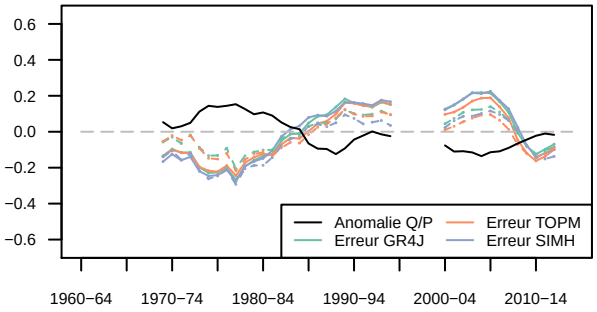

La Serene@Saint−Andre−de−Najac [Canabral] (O5284310)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


L'Aveyron@Laguepie [1] (O5292510)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

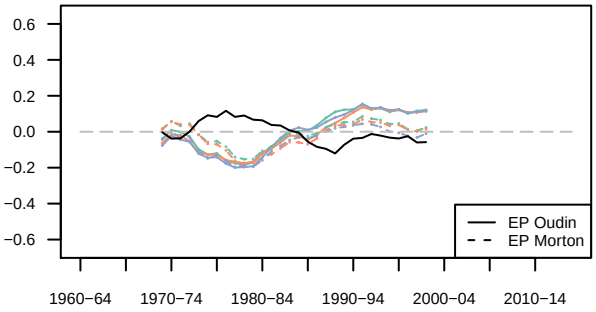

Le Viaur@Arques (O5312910)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


Le Ceor@Centres [Estrebaldie] (O5424010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

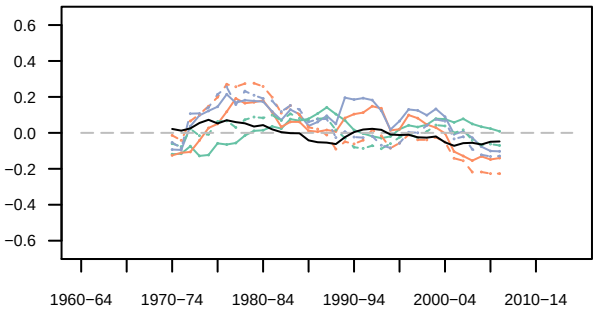
Le Giffou@Saint−Just−sur−Viaur [La Fabreguerie] (O5464310)



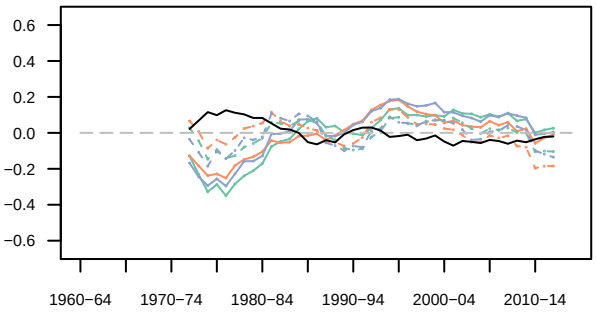
Le Lezert@Saint−Julien−du−Puy [Port de la Besse] (O5534010)



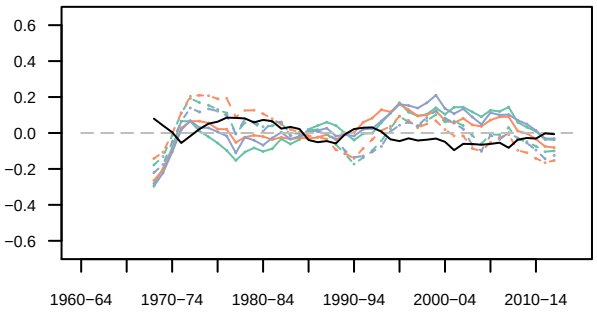
La Bonnette@Saint−Antonin−Noble−Val (O5685010)



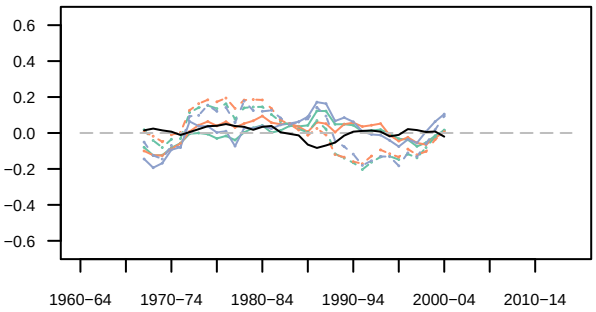
La Vere@Bruniquel [La Gauterie] (O5754020)



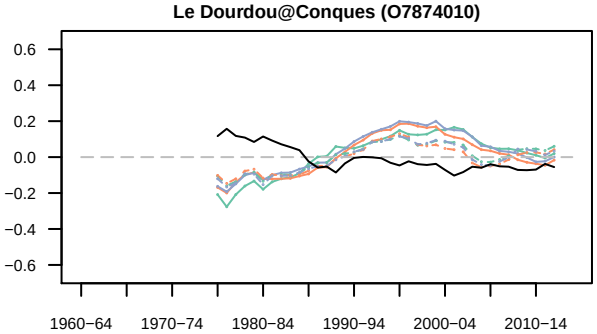
La Barguelonne@Valence [Fourquet] (O6134010)



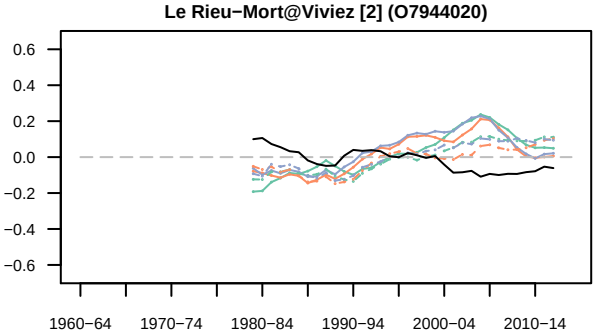
La Gelise@Mezin [Courbian] (O6793310)



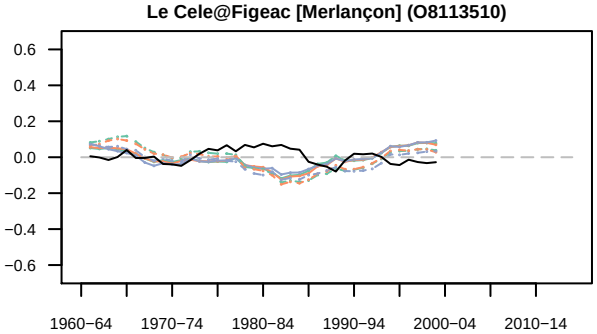
Le Dourdou@Conques (O7874010)



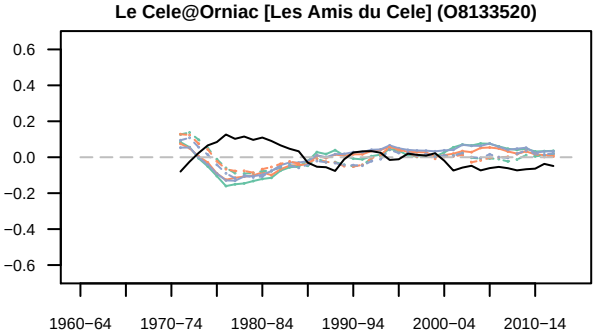
Le Rieu−Mort@Viviez [2] (O7944020)

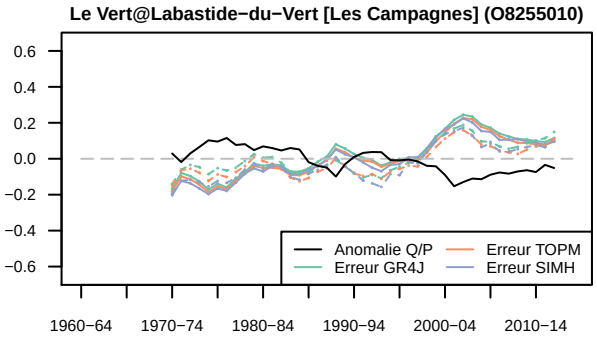


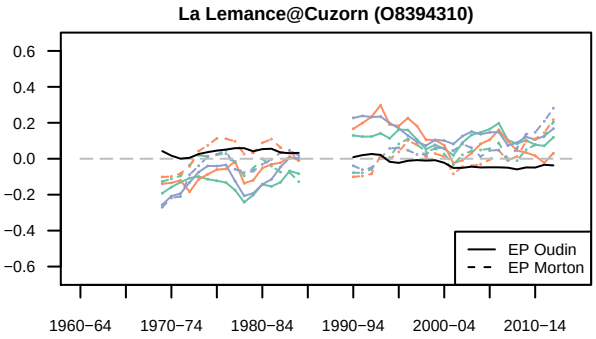
Le Cele@Figeac [Merlançon] (O8113510)

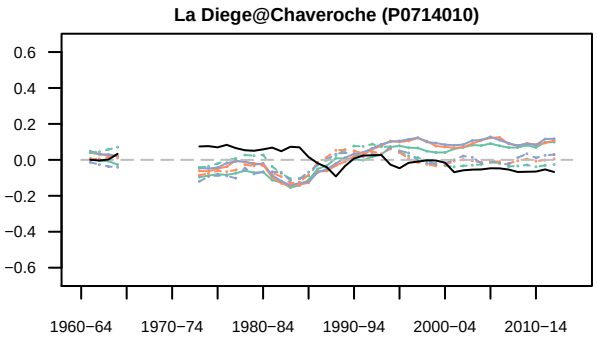


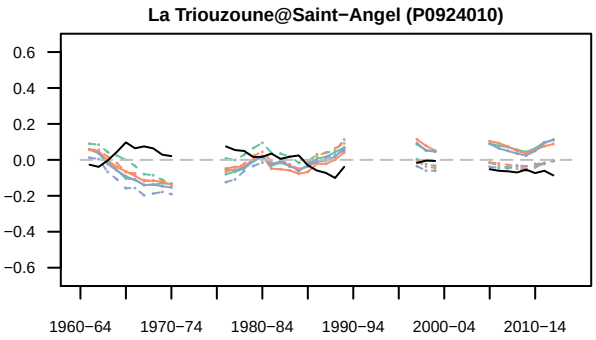
Le Cele@Orniac [Les Amis du Cele] (O8133520)

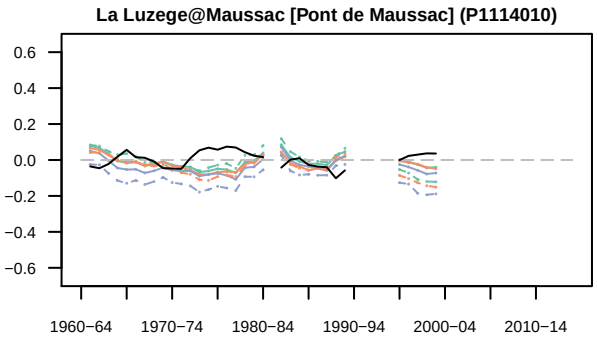


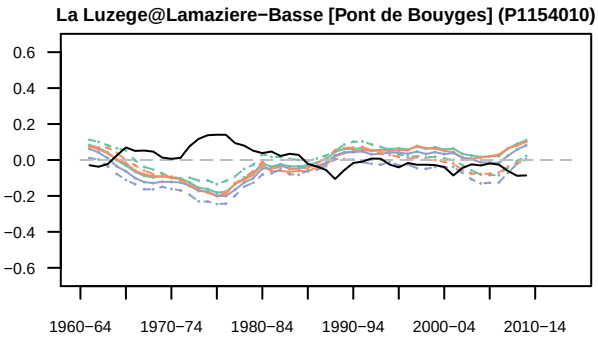

Le Vert@Labastide−du−Vert [Les Campagnes] (O8255010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
Anomalie Q/P
Erreur GR4J
Erreur TOPM
Erreur SIMH
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

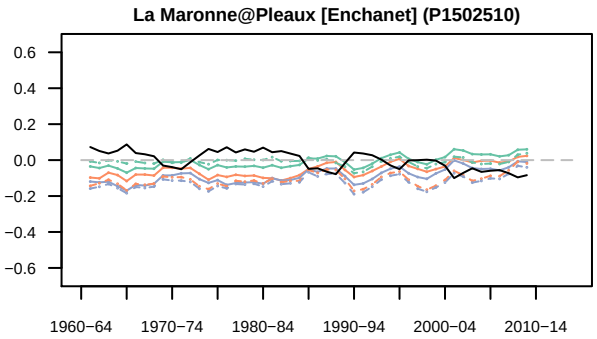

La Lemance@Cuzorn (O8394310)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
EP Oudin
EP Morton
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

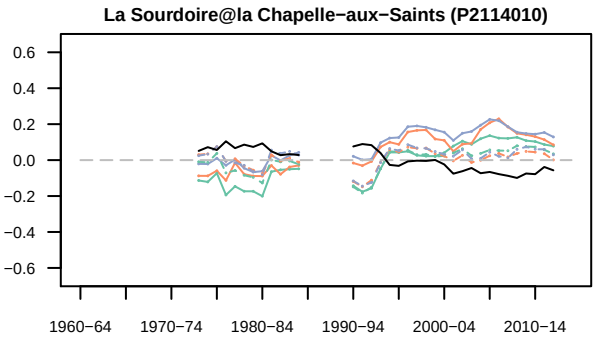

La Diege@Chaveroche (P0714010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

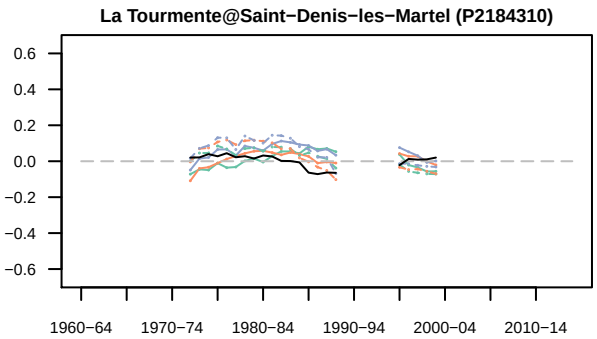

La Triouzoune@Saint−Angel (P0924010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

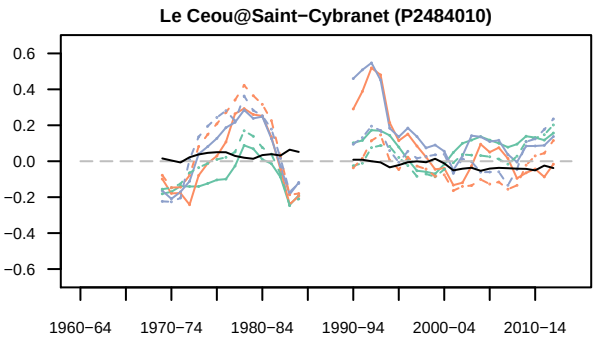

La Luzege@Maussac [Pont de Maussac] (P1114010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


La Luzege@Lamaziere−Basse [Pont de Bouyges] (P1154010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

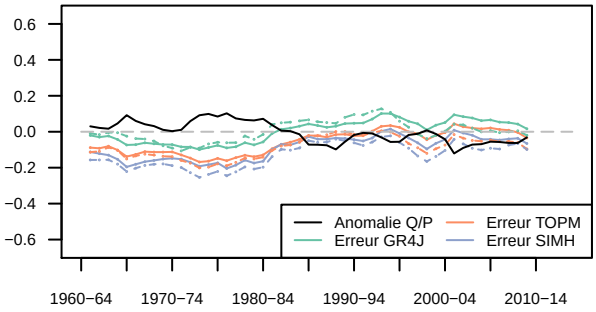

La Maronne@Pleaux [Enchanet] (P1502510)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


La Sourdoire@la Chapelle−aux−Saints (P2114010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

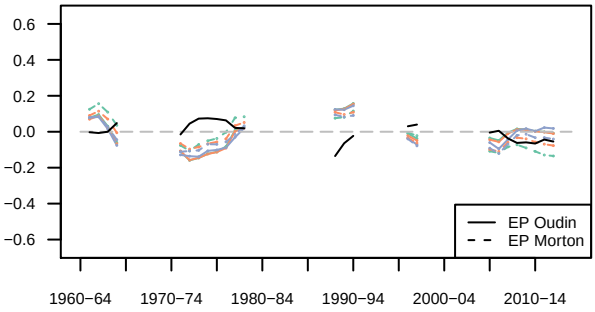

La Tourmente@Saint−Denis−les−Martel (P2184310)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


Le Ceou@Saint−Cybranet (P2484010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

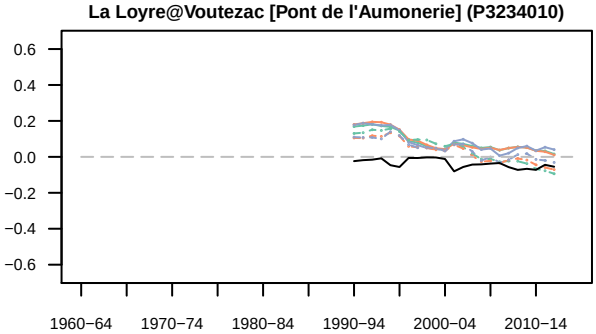
La Vezere@Bugeat (P3021010)



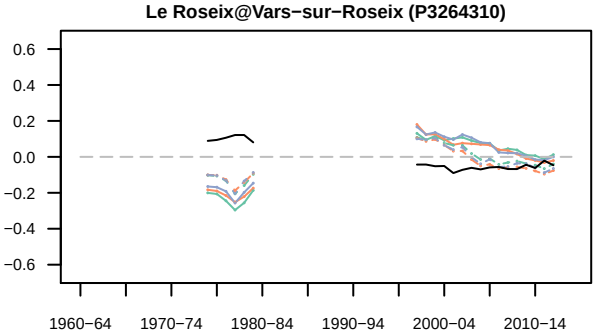
La Brezou@Vigeois [Pont de Bleygeat] (P3194310)



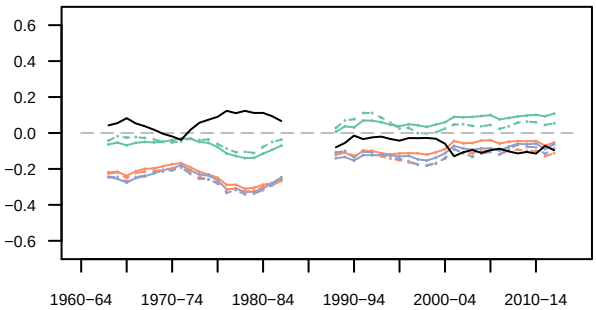
La Loyre@Voutezac [Pont de l'Aumonerie] (P3234010)



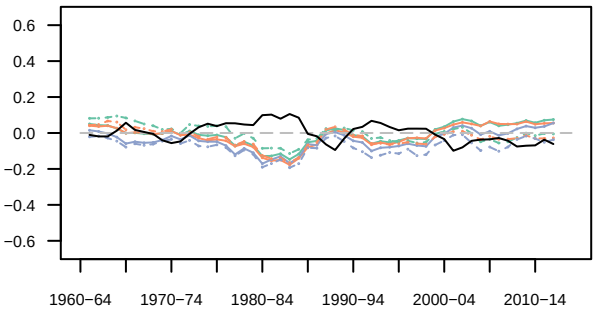
Le Roseix@Vars−sur−Roseix (P3264310)



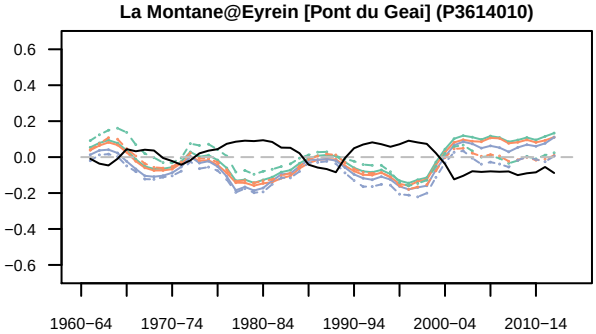
La Correze@Saint−Yrieix−le−Dejalat [Pont de Lanour] (P3322510)



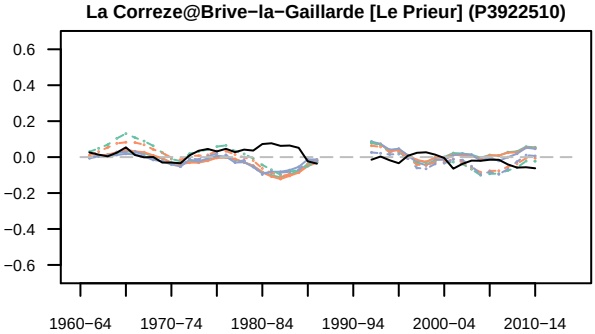
La Correze@Tulle [Pont des soldats] (P3502510)



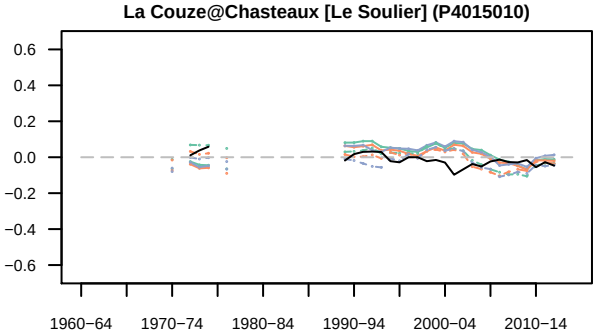
La Montane@Eyrein [Pont du Geai] (P3614010)



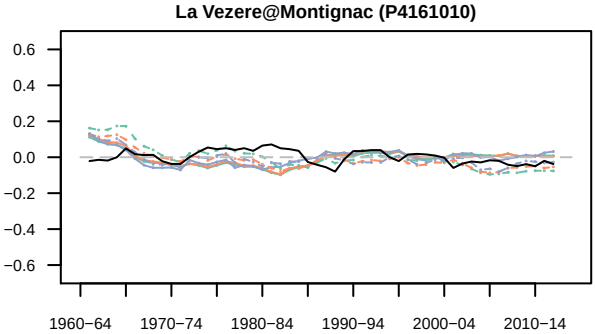
La Correze@Brive−la−Gaillarde [Le Prieur] (P3922510)

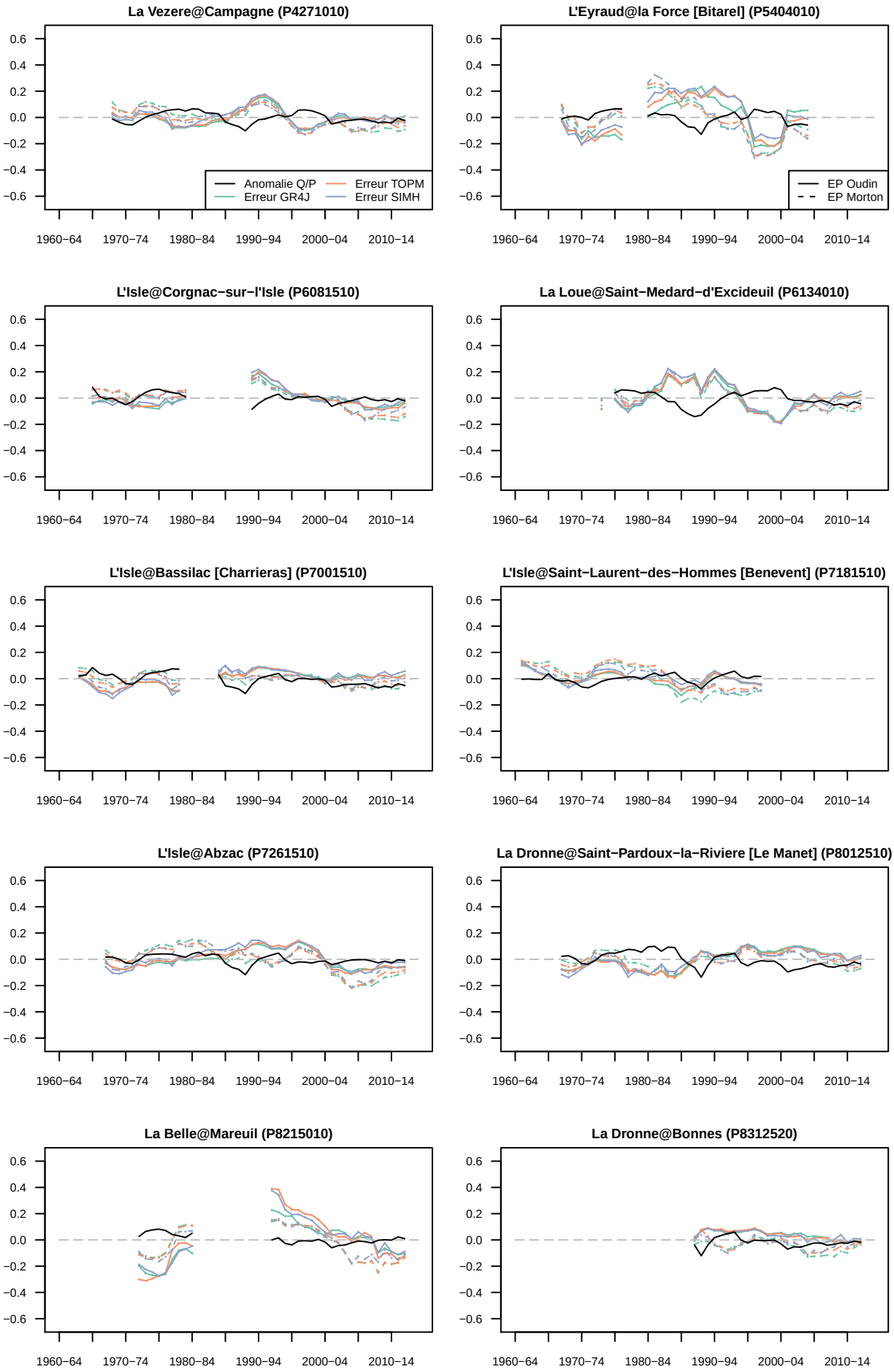


La Couze@Chasteaux [Le Soulier] (P4015010)

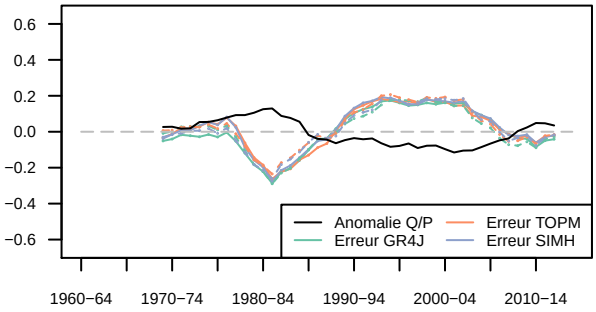


La Vezere@Montignac (P4161010)

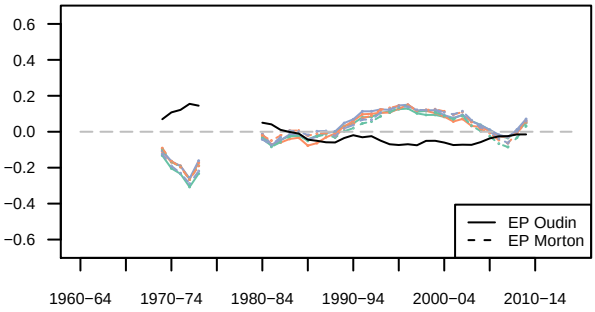



La Vezere@Campagne (P4271010)
L'Eyraud@la Force [Bitarel] (P5404010)
L'Isle@Corgnac−sur−l'Isle (P6081510)
La Loue@Saint−Medard−d'Excideuil (P6134010)
L'Isle@Bassilac [Charrieras] (P7001510)
L'Isle@Saint−Laurent−des−Hommes [Benevent] (P7181510)
L'Isle@Abzac (P7261510)
La Dronne@Saint−Pardoux−la−Riviere [Le Manet] (P8012510)
La Belle@Mareuil (P8215010)
La Dronne@Bonnes (P8312520)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14
Anomalie Q/P
Erreur GR4J
Erreur TOPM
Erreur SIMH
EP Oudin
EP Morton

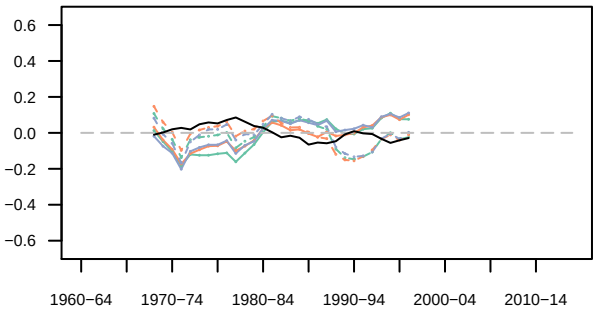
L'echez@Louey (Q0214010)



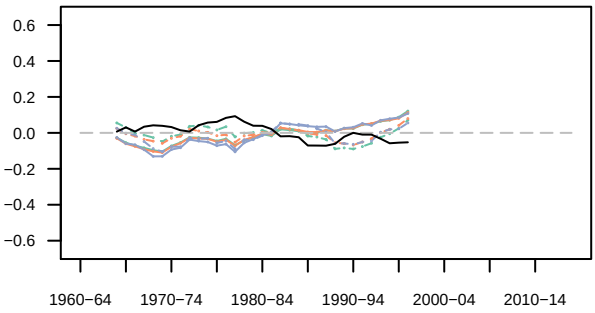
L'Arros@Gourgue (Q0522520)



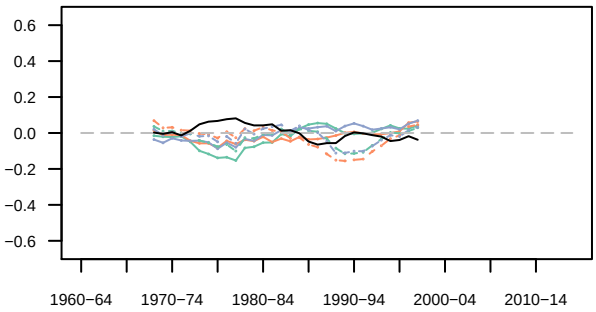
Le Boues@Juillac (Q0664010)



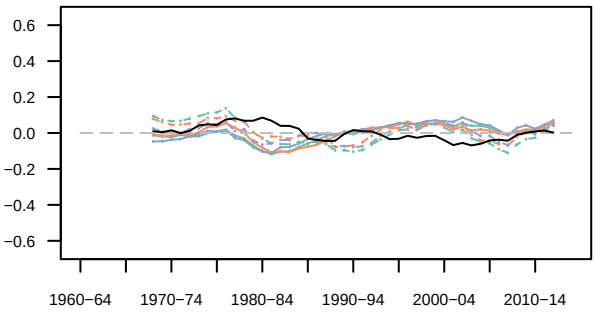
L'Adour@Corneillan (Q0740010)



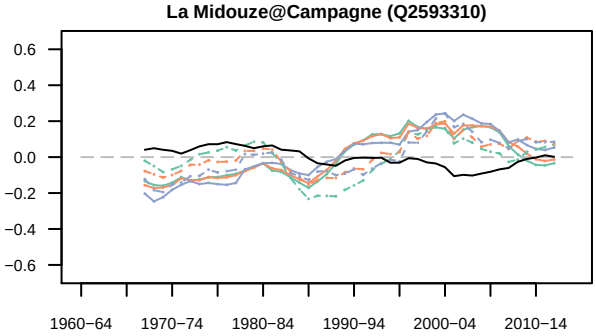
Le Larcis@Lannux (Q1094010)



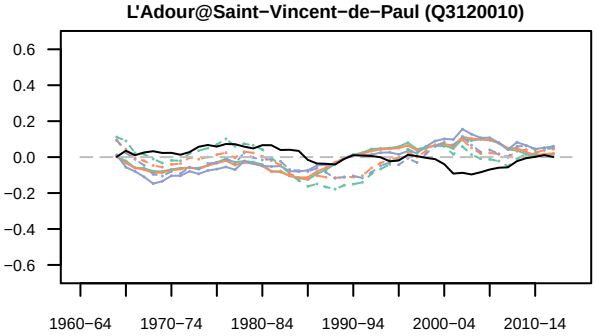
L'Adour@Aire−sur−l'Adour [2] (Q1100010)



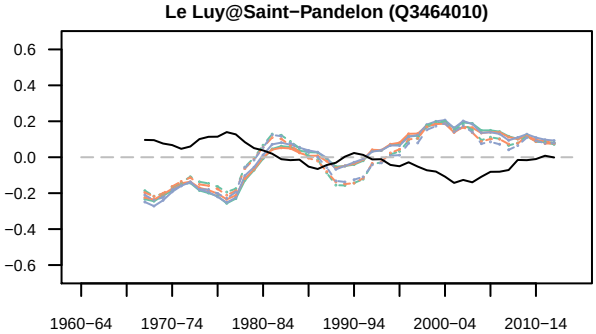
La Midouze@Campagne (Q2593310)



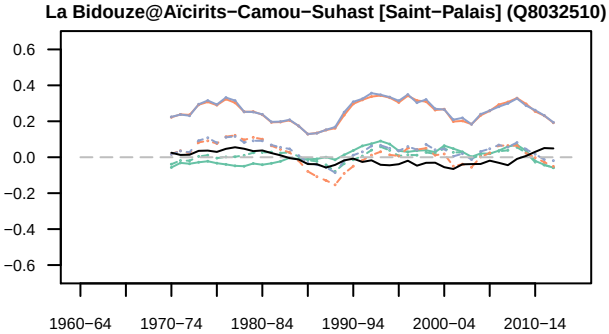
L'Adour@Saint−Vincent−de−Paul (Q3120010)

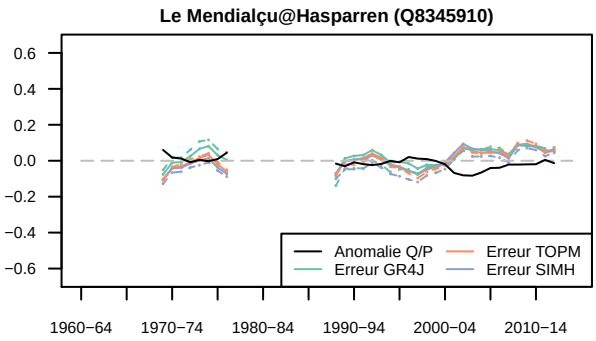


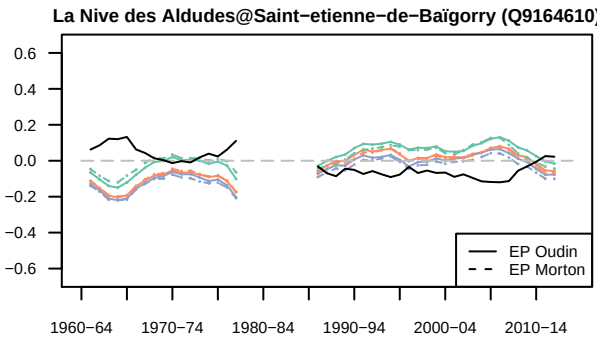
Le Luy@Saint−Pandelon (Q3464010)

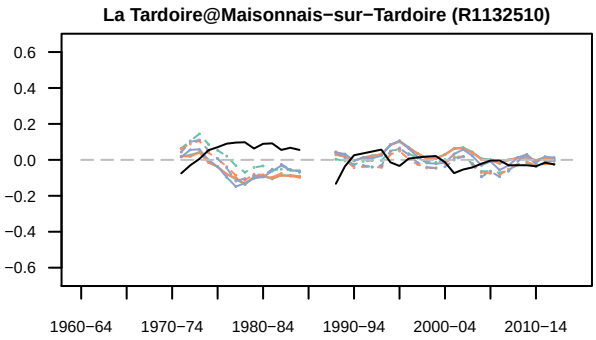


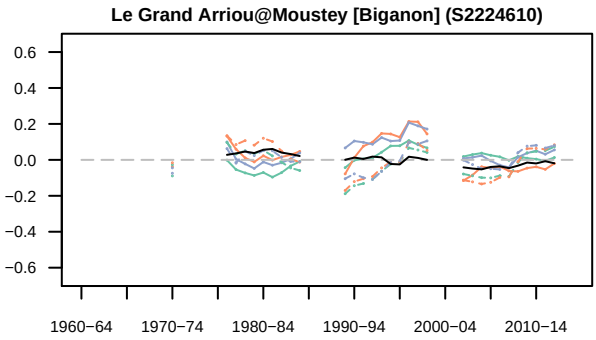
La Bidouze@Aïcirits−Camou−Suhast [Saint−Palais] (Q8032510)

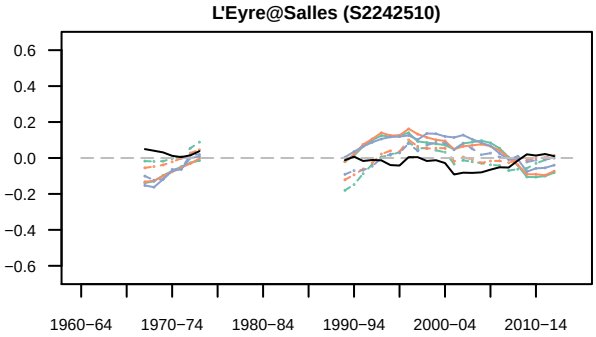


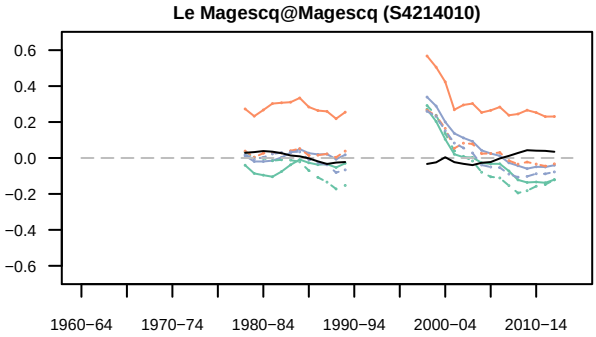

Le Mendialçu@Hasparren (Q8345910)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

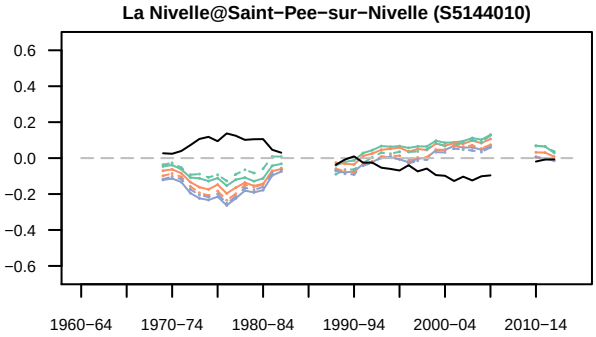

La Nive des Aldudes@Saint−etienne−de−Baïgorry (Q9164610)
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

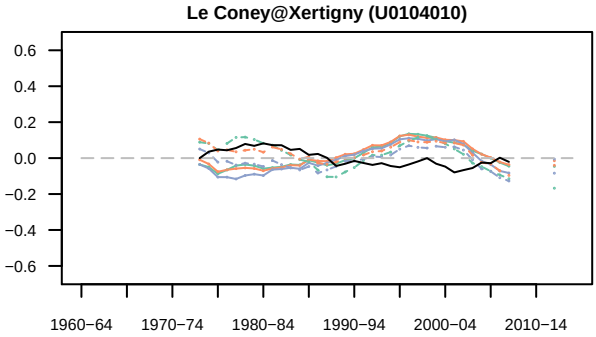

La Tardoire@Maisonnais−sur−Tardoire (R1132510)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

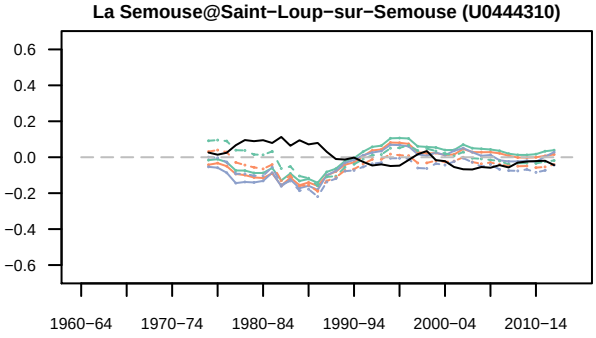

Le Grand Arriou@Moustey [Biganon] (S2224610)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

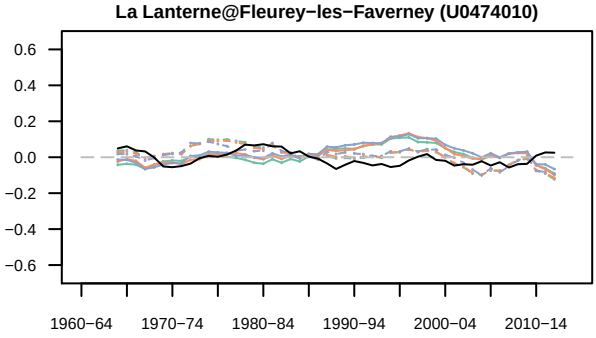

L'Eyre@Salles (S2242510)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


Le Magescq@Magescq (S4214010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

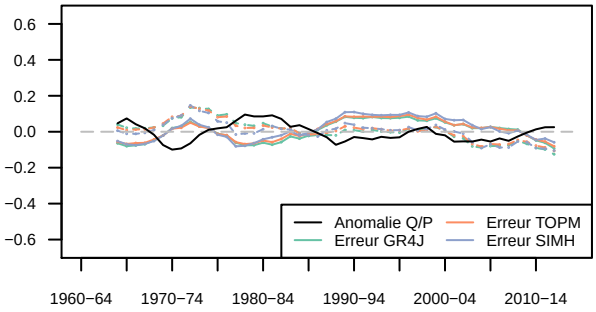

La Nivelle@Saint−Pee−sur−Nivelle (S5144010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


Le Coney@Xertigny (U0104010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

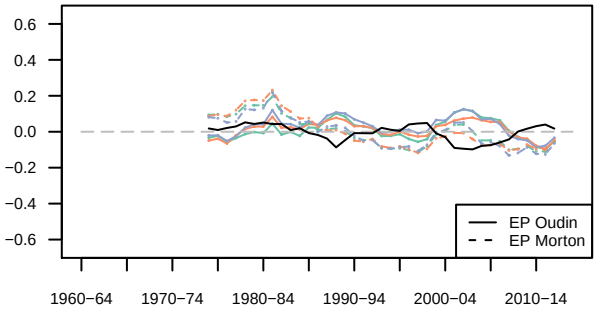

La Semouse@Saint−Loup−sur−Semouse (U0444310)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


La Lanterne@Fleurey−les−Faverney (U0474010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

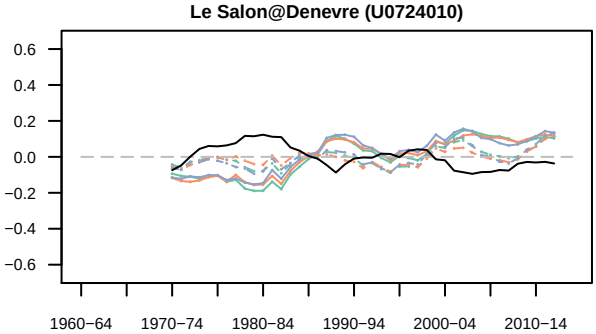
La Saône@Ray−sur−Saône (U0610010)



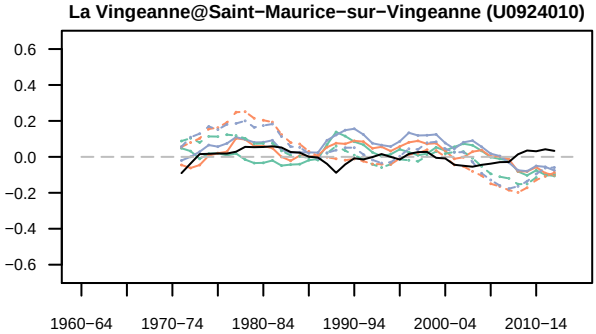
La Gourgeonne@Tincey−et−Pontrebeau (U0635010)



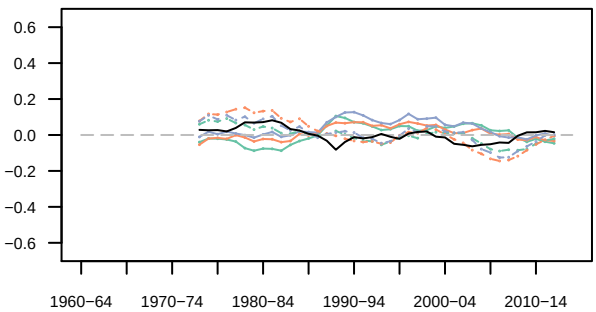
Le Salon@Denevre (U0724010)



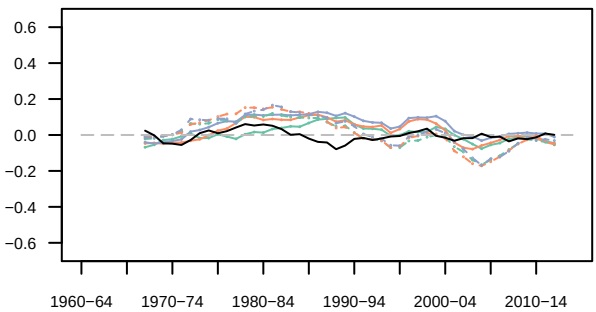
La Vingeanne@Saint−Maurice−sur−Vingeanne (U0924010)



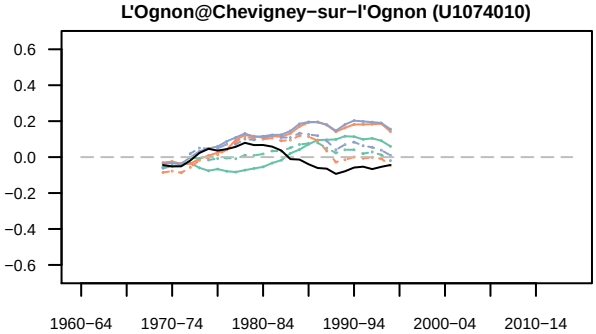
La Vingeanne@Oisilly (U0924020)



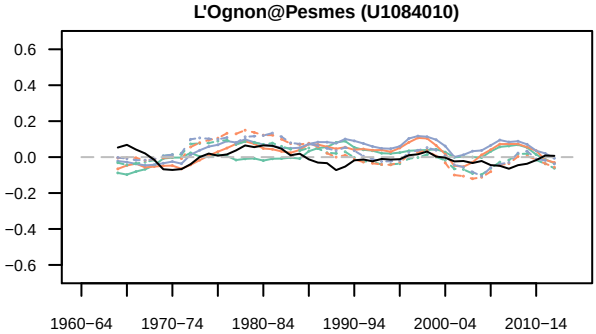
L'Ognon@Beaumotte−Aubertans (U1054010)



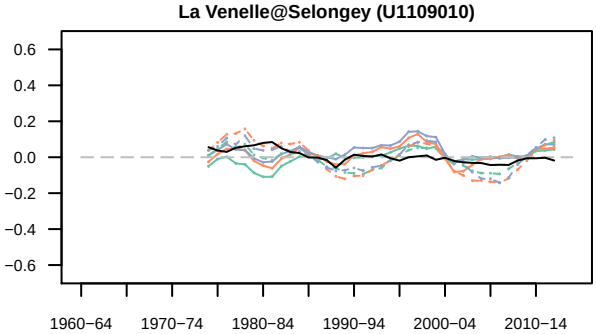
L'Ognon@Chevigney−sur−l'Ognon (U1074010)



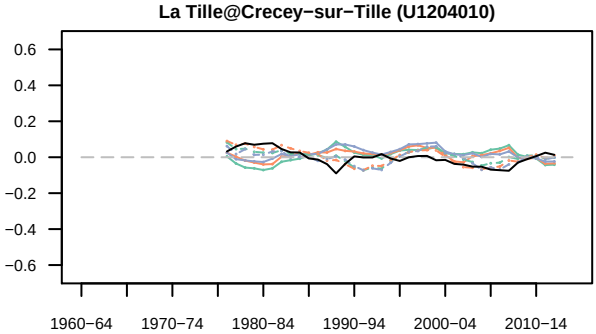
L'Ognon@Pesmes (U1084010)

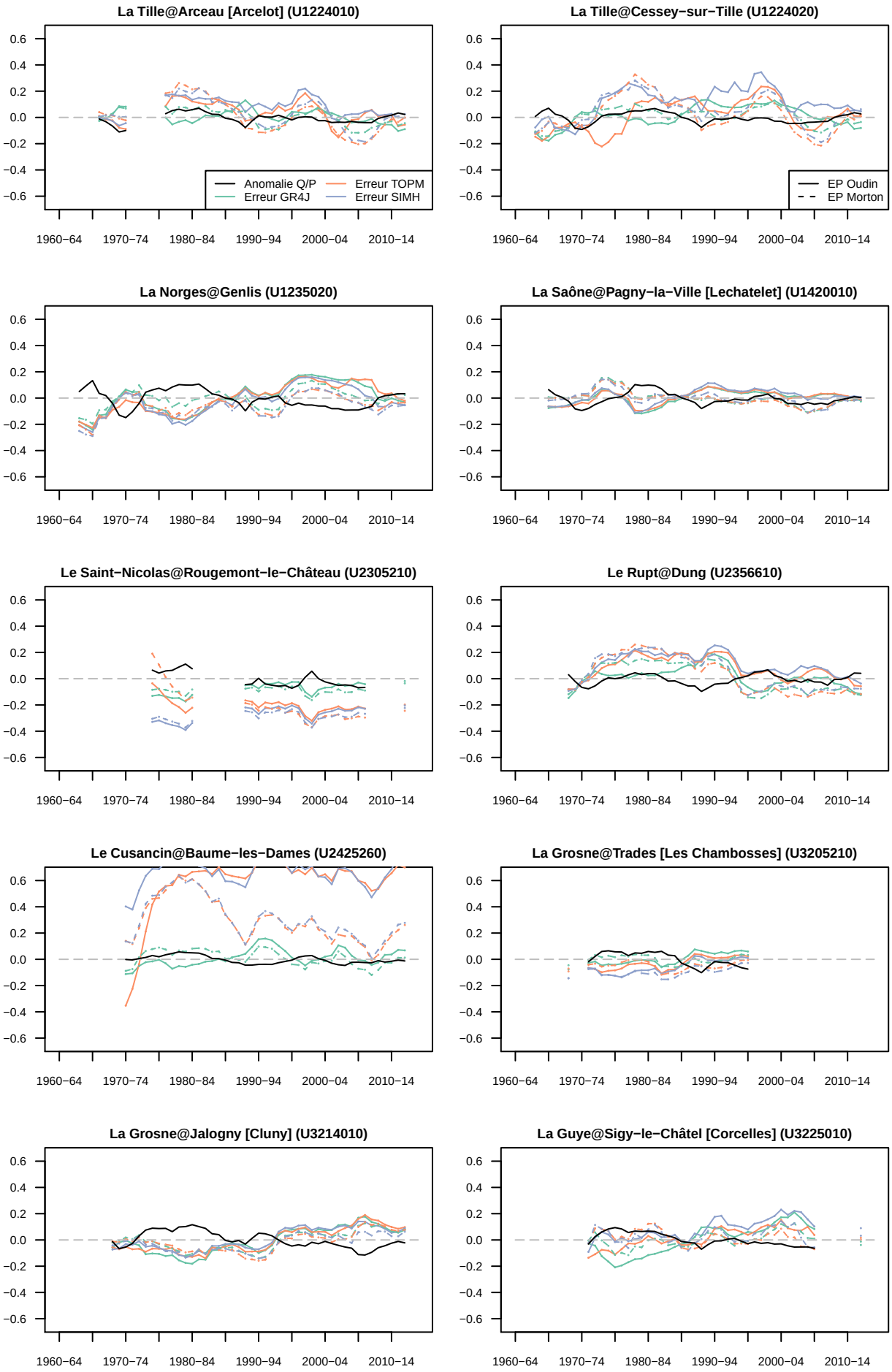


La Venelle@Selongey (U1109010)

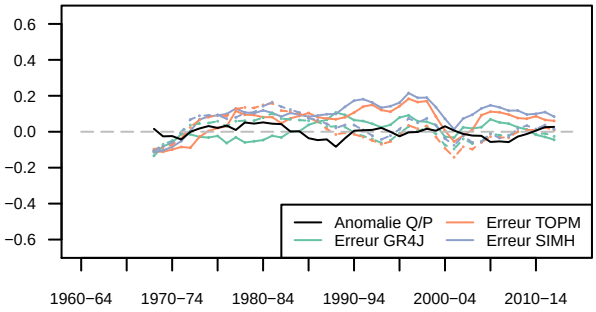


La Tille@Crecey−sur−Tille (U1204010)

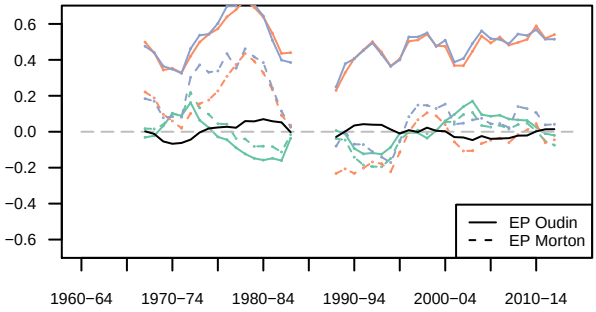



La Tille@Arceau [Arcelot] (U1224010)
La Tille@Cessey−sur−Tille (U1224020)
La Norges@Genlis (U1235020)
La Saône@Pagny−la−Ville [Lechatelet] (U1420010)
Le Saint−Nicolas@Rougemont−le−Château (U2305210)
Le Rupt@Dung (U2356610)
Le Cusancin@Baume−les−Dames (U2425260)
La Grosne@Trades [Les Chambosses] (U3205210)
La Grosne@Jalogny [Cluny] (U3214010)
La Guye@Sigy−le−Châtel [Corcelles] (U3225010)
Anomalie Q/P
Erreur GR4J
Erreur TOPM
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

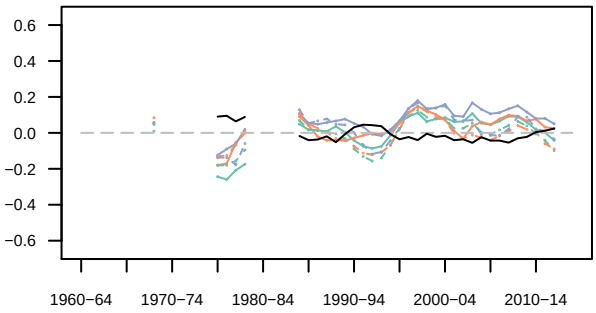
La Seille@Saint−Usuge (U3424010)



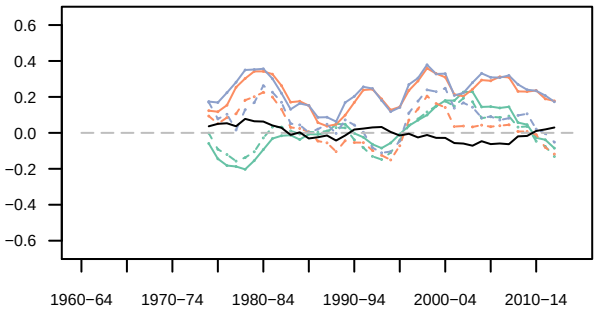
La Reyssouze@Montagnat (U4014010)



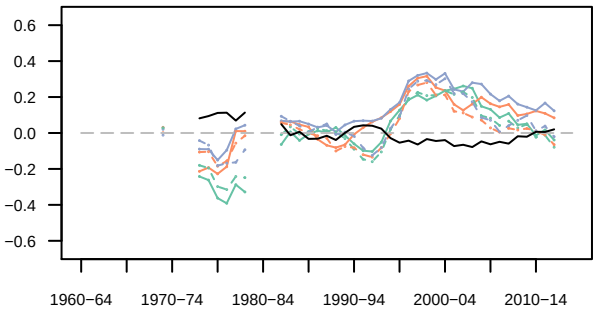
La Veyle@Lent (U4204010)



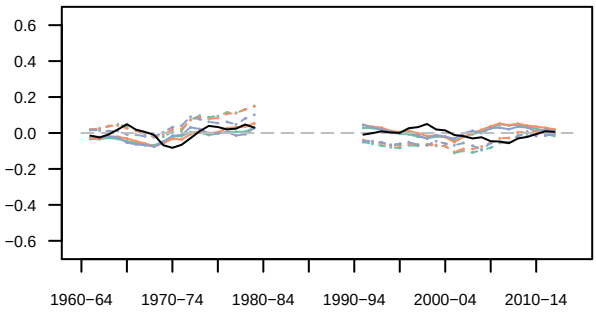
Le Vieux Jonc@Buellas [Corgenon] (U4216010)



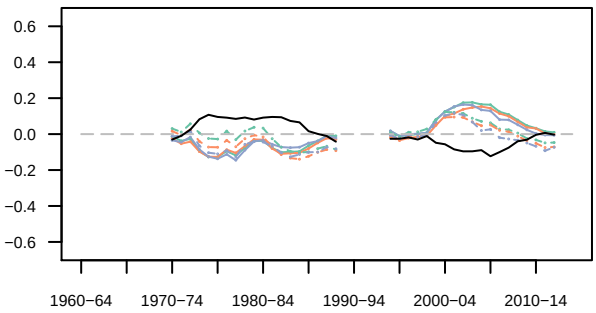
Le Renon@Neuville−les−Dames (U4235010)



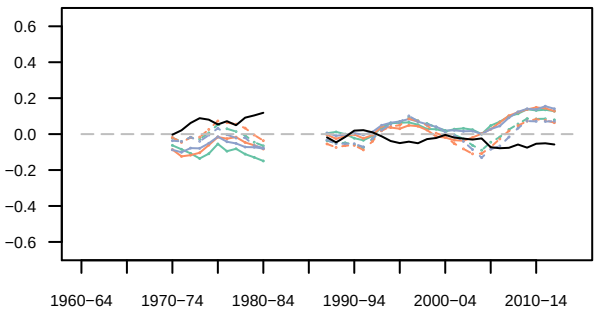
La Saône@Mâcon [amont] (U4300010)



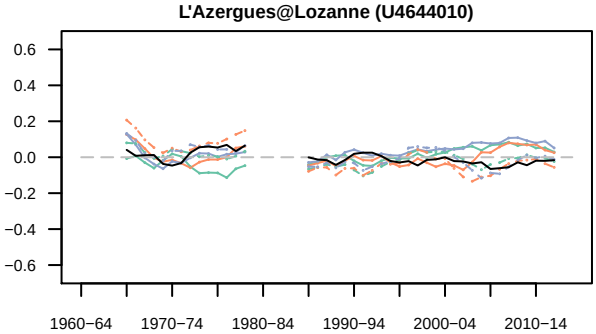
L'Ardieres@Beaujeu (U4505010)



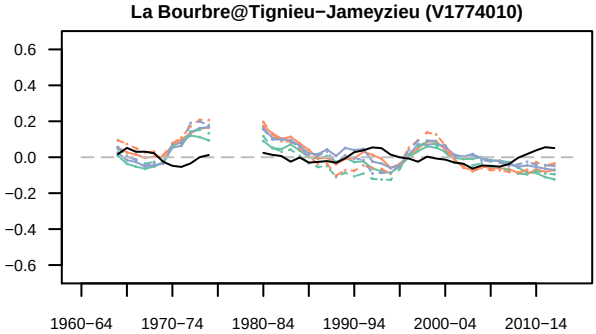
L'Azergues@Châtillon (U4624010)

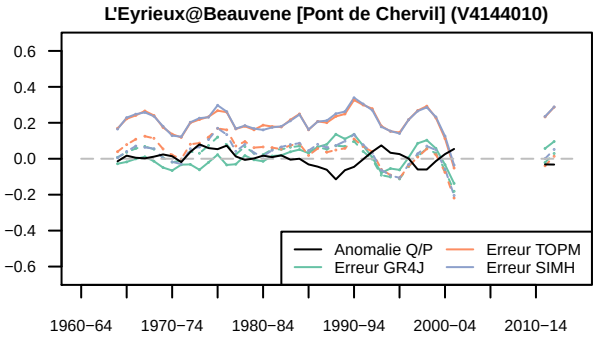


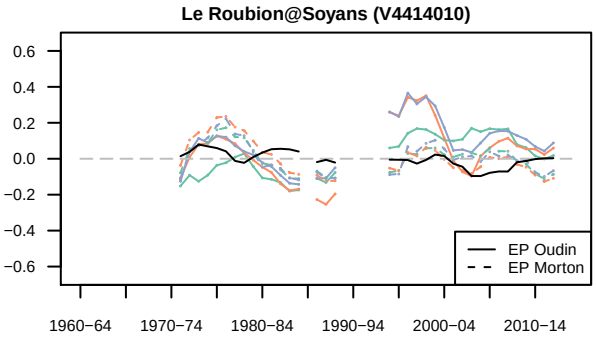
L'Azergues@Lozanne (U4644010)

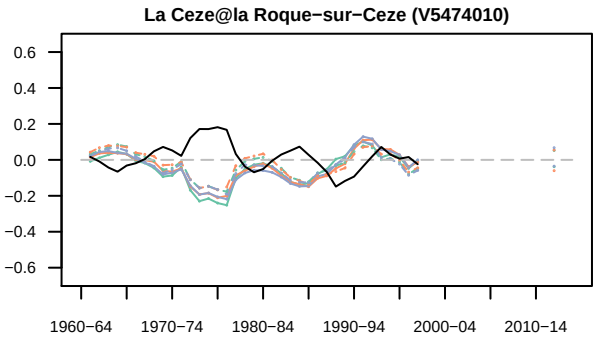


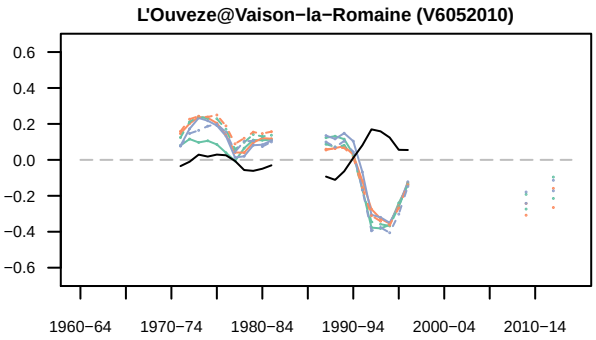
La Bourbre@Tignieu−Jameyzieu (V1774010)

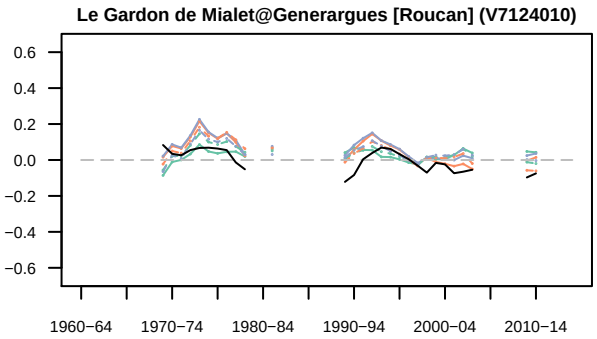


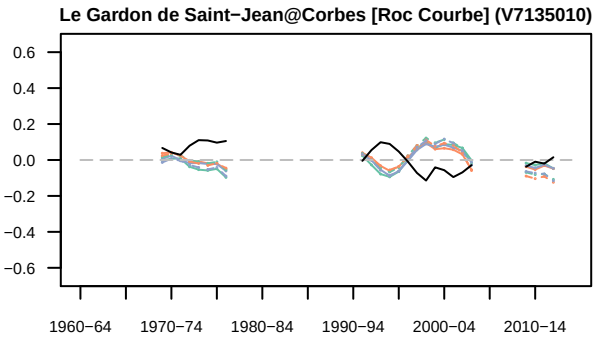

L'Eyrieux@Beauvene [Pont de Chervil] (V4144010)
Anomalie Q/P
Erreur GR4J
Erreur TOPM
Erreur SIMH
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

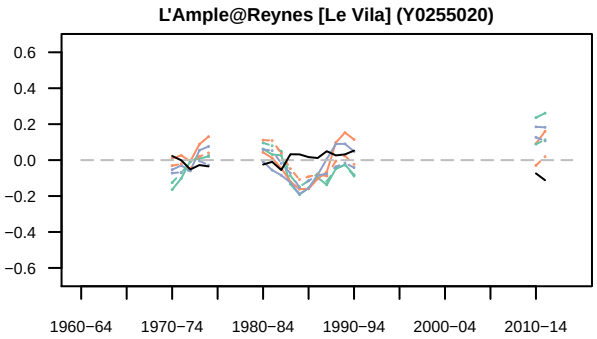

Le Roubion@Soyans (V4414010)
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

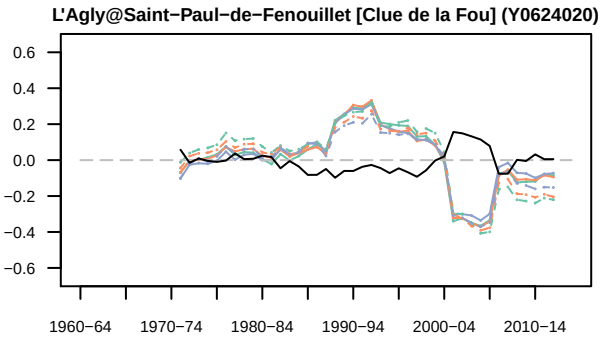

La Ceze@la Roque−sur−Ceze (V5474010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

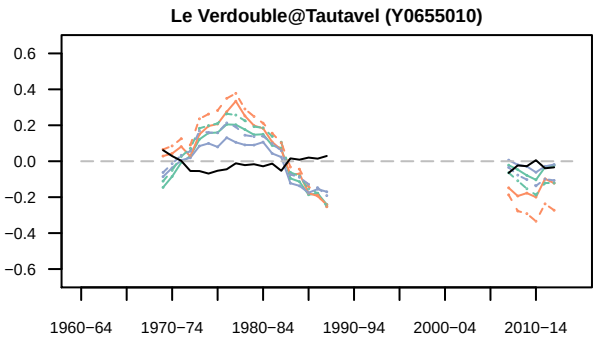

L'Ouveze@Vaison−la−Romaine (V6052010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

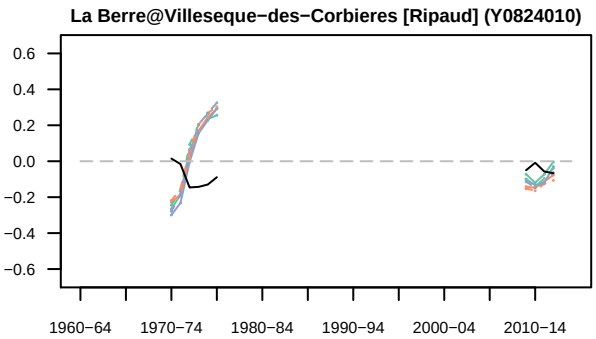

Le Gardon de Mialet@Generargues [Roucan] (V7124010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


Le Gardon de Saint−Jean@Corbes [Roc Courbe] (V7135010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

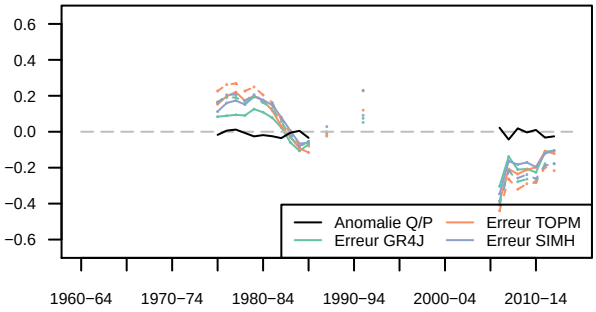

L'Ample@Reynes [Le Vila] (Y0255020)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


L'Agly@Saint−Paul−de−Fenouillet [Clue de la Fou] (Y0624020)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

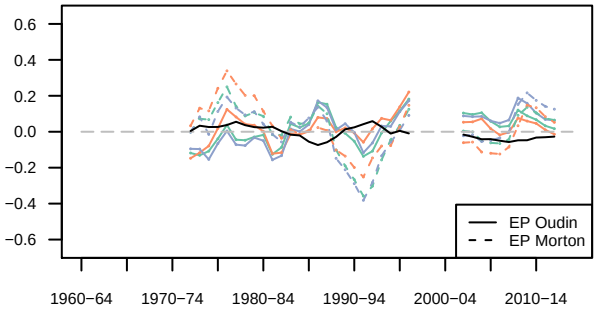

Le Verdouble@Tautavel (Y0655010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


La Berre@Villeseque−des−Corbieres [Ripaud] (Y0824010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

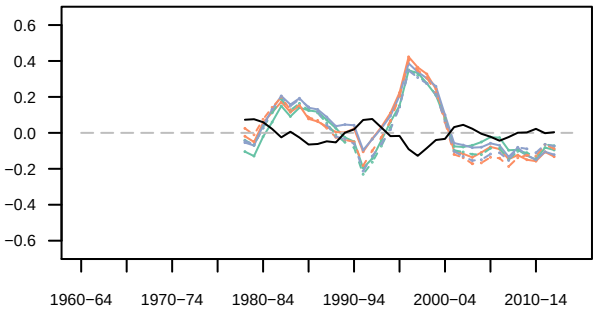
Le Lauquet@Greffeil (Y1225010)



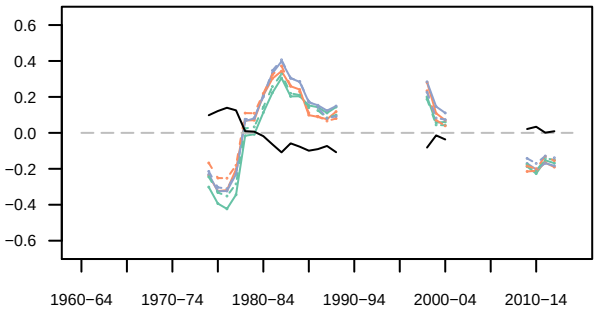
Le Treboul@Villepinte (Y1325010)



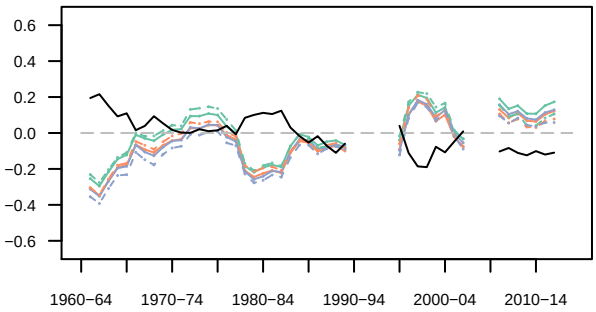
L'Orbiel@Bouilhonnac [Villedubert] (Y1415020)



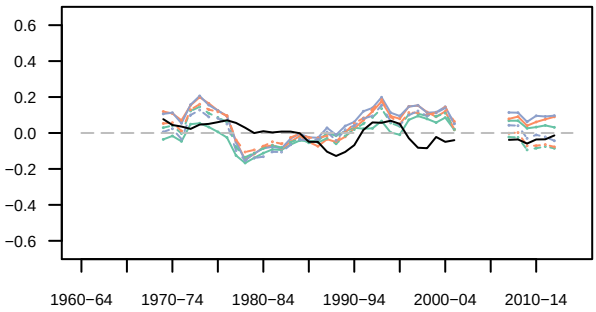
La Clamoux@Malves−en−Minervois (Y1416210)



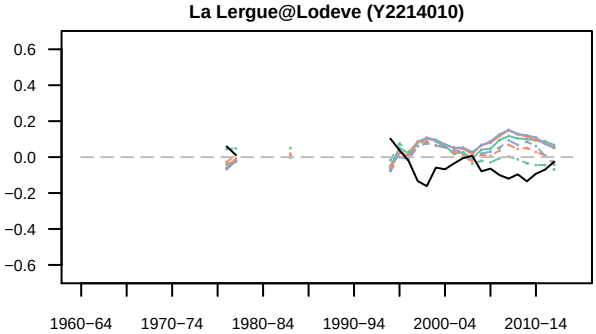
L'Arre au Vigan [La Terrisse] (Y2015010)



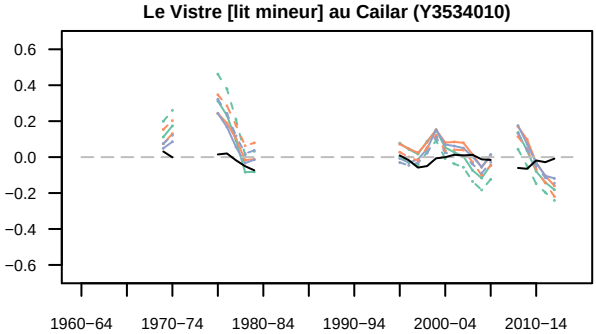
L'Herault@Laroque (Y2102010)



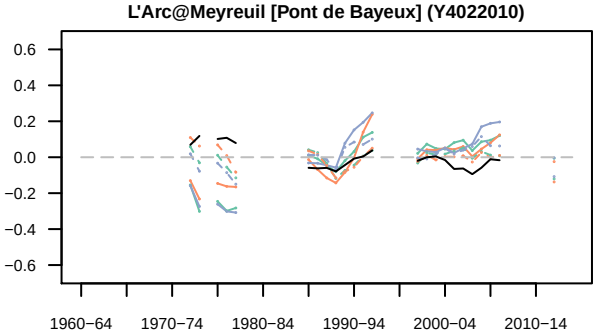
La Lergue@Lodeve (Y2214010)



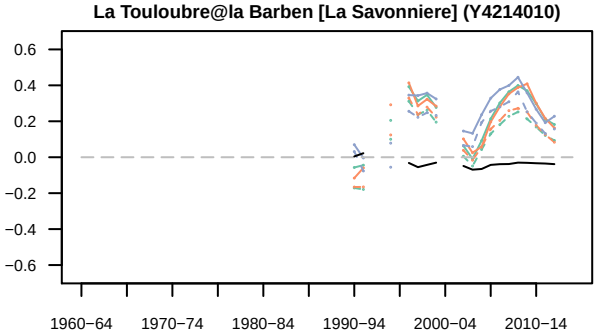
Le Vistre [lit mineur] au Cailar (Y3534010)

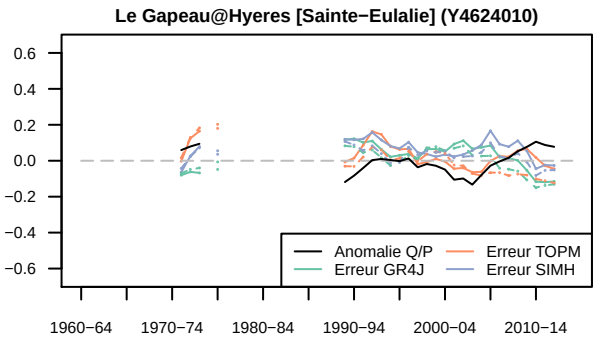


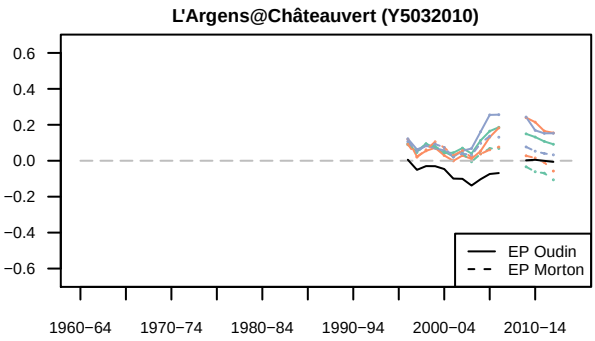
L'Arc@Meyreuil [Pont de Bayeux] (Y4022010)

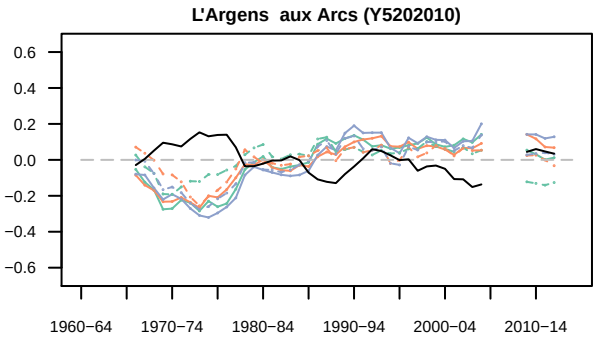


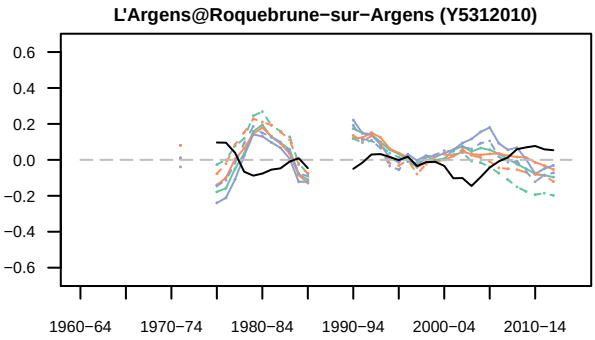
La Touloubre@la Barben [La Savonniere] (Y4214010)

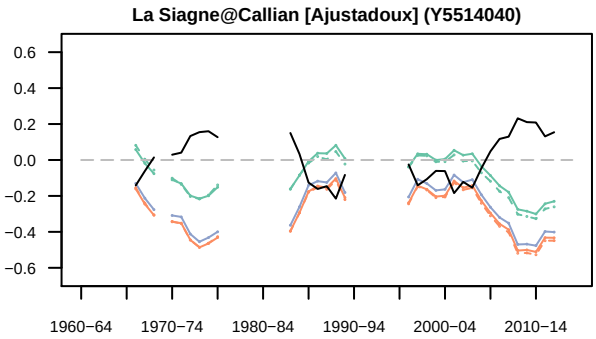


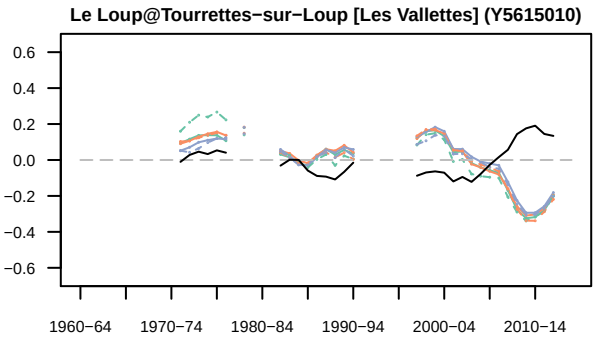

Le Gapeau@Hyeres [Sainte−Eulalie] (Y4624010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
Anomalie Q/P
Erreur GR4J
Erreur TOPM
Erreur SIMH
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

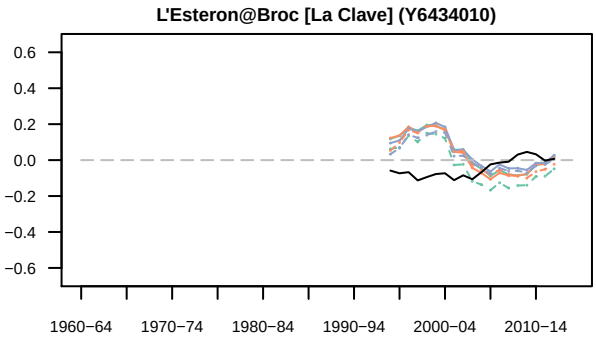

L'Argens@Châteauvert (Y5032010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
EP Oudin
EP Morton
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


L'Argens aux Arcs (Y5202010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


L'Argens@Roquebrune−sur−Argens (Y5312010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


La Siagne@Callian [Ajustadoux] (Y5514040)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


Le Loup@Tourrettes−sur−Loup [Les Vallettes] (Y5615010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14


L'Esteron@Broc [La Clave] (Y6434010)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1960−64
1970−74
1980−84
1990−94
2000−04
2010−14

Annexe F.

Courbes des biais glissants des modèles pour les bassins australiens de l'échantillon

F.1. Introduction

Cette annexe présente les courbes des biais glissants des modèles GR4J, TOPMO et SimHyd sur le sous-échantillon de 181 bassins versants asutraliens de la thèse. Les courbes ont été calculées pour ces modèles forcés avec les formules d'évaporation potentielle d'Oudin ou de Morton.

F.2. Mise en œuvre des courbes

La Figure F.1 illustre un exemple des courbes de biais glissants représentées.

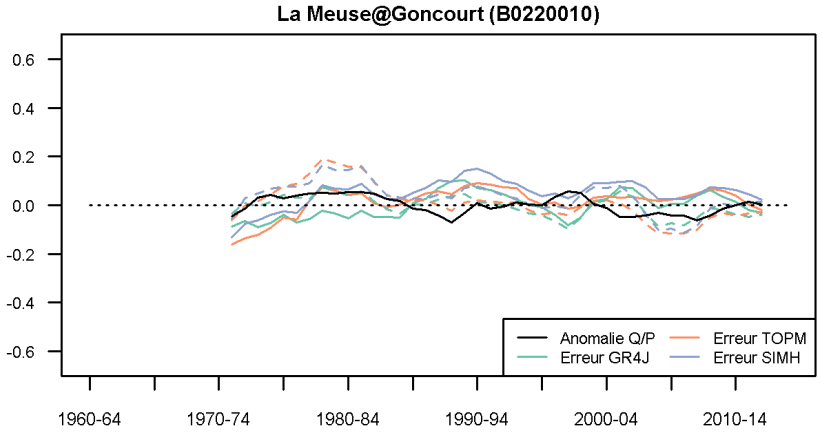
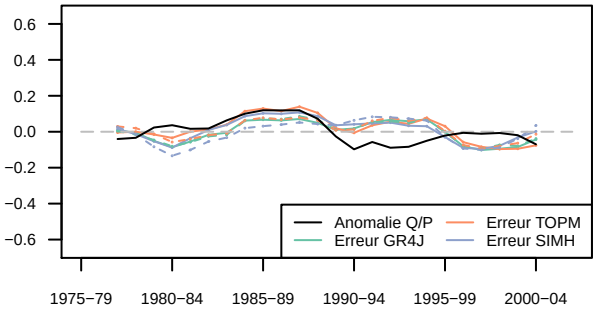


FIGURE F.1 : Courbes des biais glissants des modèles GR4J (vert), TOPMO (orange) et SimHyd (bleu), pour les formulations de l'évaporation potentielle d'Oudin (trait plein) et de Morton (trait pointillé). Les anomalies de rendement des précipitations sont représentées en noir. Les sous-périodes de 5 ans sont définies sur les années hydrologiques.

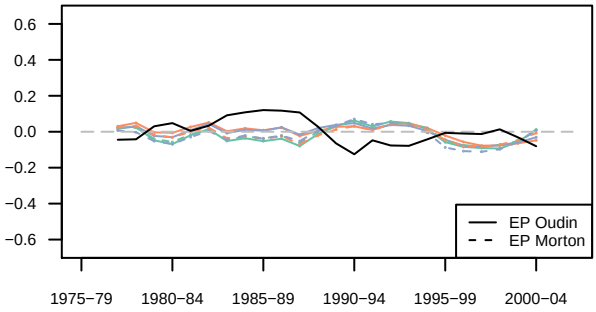
Le biais d'un modèle sur une sous-période de 5 ans a été calculé comme l'erreur moyenne du modèle sur la sous-période rapportée au débit moyen observé sur la période totale, c'est-à-dire de la même manière que dans le calcul du PMR (voir Chapitre 4). Ce n'est donc pas un biais à proprement parler, mais plutôt une erreur normalisée. Nous avons toutefois maintenu l'appellation « courbes des biais glissants ».
Les six courbes des biais glissants (trois modèles hydrologiques pour deux formulations de l'évaporation potentielle) ont été mises en perspective avec l'évolution de l'anomalie de rendement des précipitations au cours du temps, afin d'illustrer les résultats du Chapitre 5. Pour rappel, nous avons montré dans ce chapitre une anti-corrélation importante entre les biais des modèles hydrologiques et l'anomalie de rendement des précipitations. L'anomalie de rendement des précipitations sur une sous-période a été calculée comme l'écart entre le rendement moyen des précipitations sur la sous-période considérée et le rendement moyen des précipitations sur la période totale.

F.3. Courbes de biais glissants pour le sous-échantillon de bassins australien

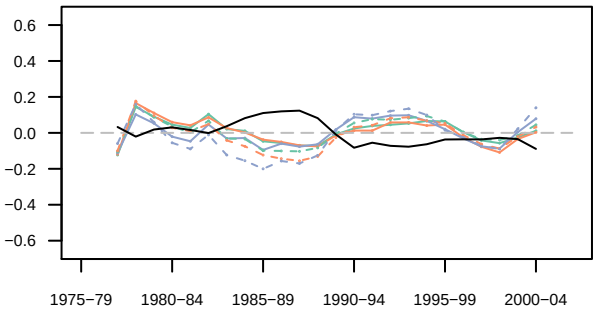
Oxley at Eungella (AU201001)



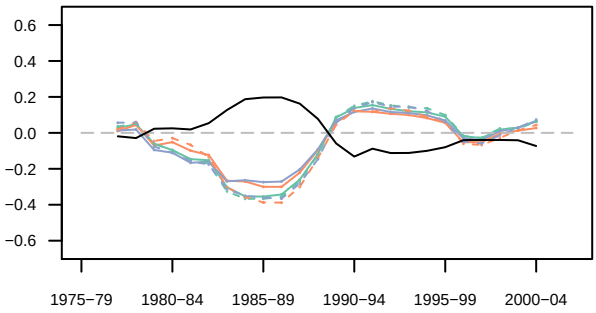
Coopers Ck at Repentance (AU203002)



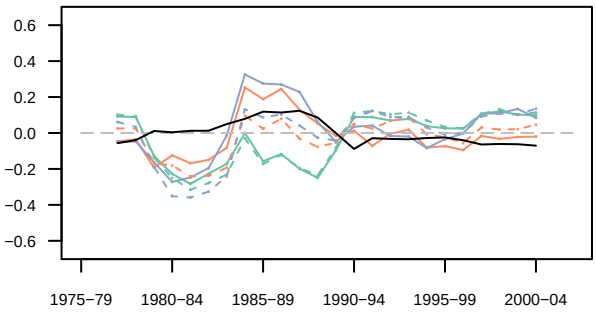
Richmond at Wiangaree (AU203005)



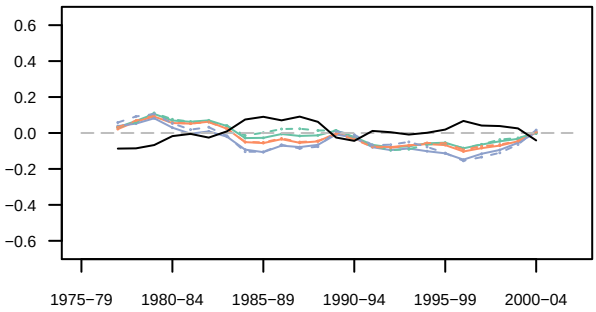
Leycester at Rock Valley (AU203010)



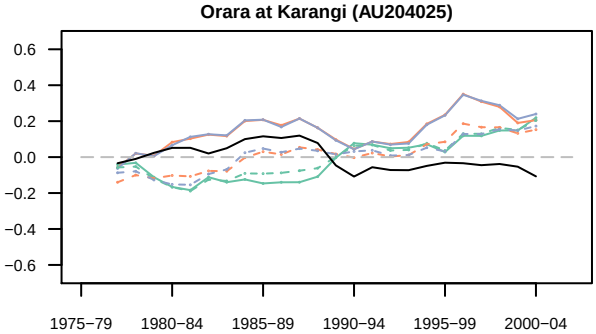
Myrtle Ck at Rappville (AU203030)



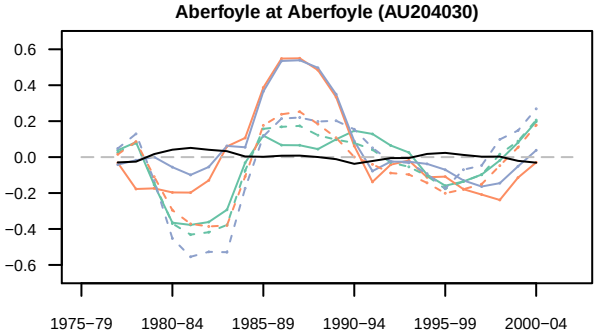
Bielsdown Ck at Dorrigo No.2 & No.3 (AU204017)



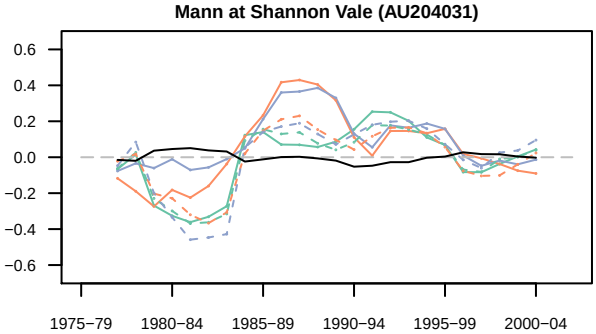
Orara at Karangi (AU204025)



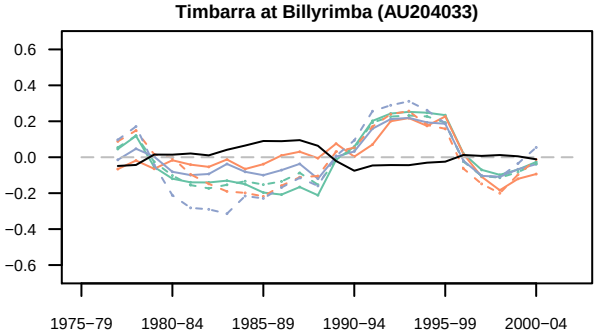
Aberfoyle at Aberfoyle (AU204030)

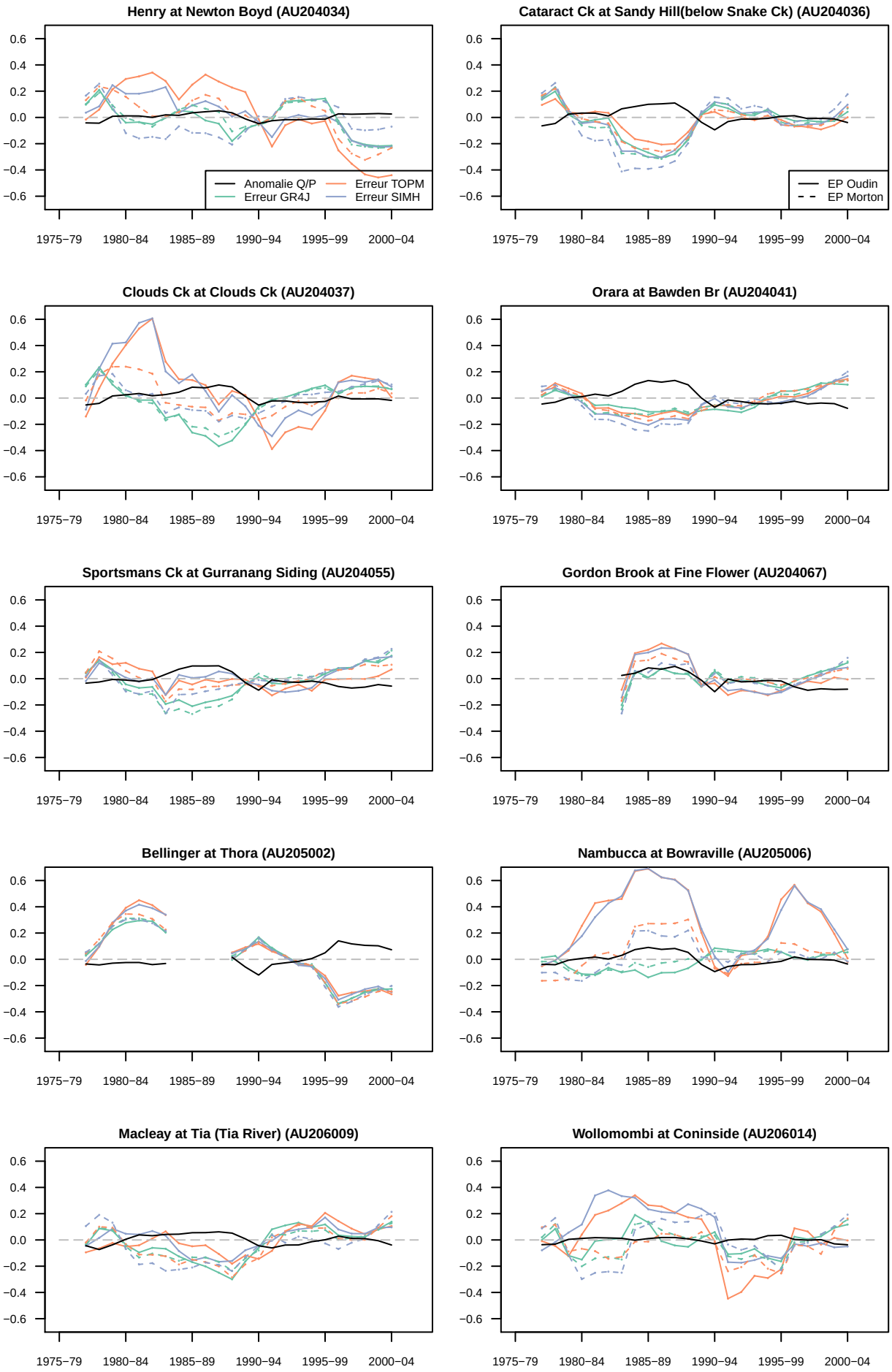


Mann at Shannon Vale (AU204031)

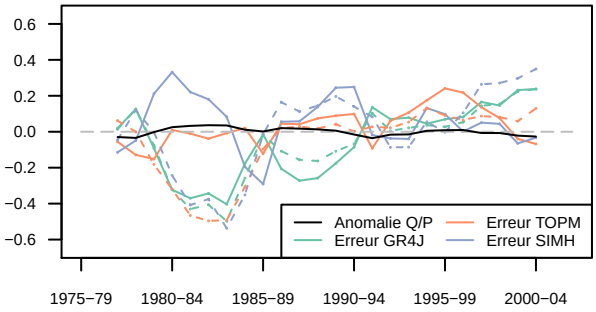


Timbarra at Billyrimba (AU204033)

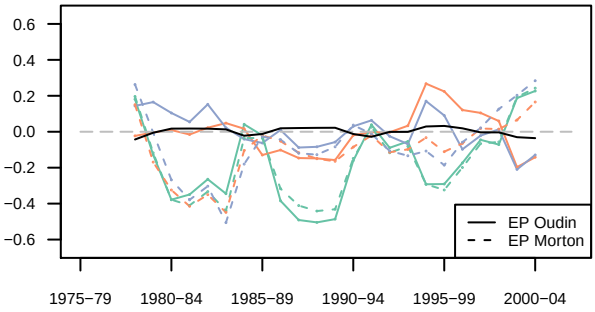



Henry at Newton Boyd (AU204034)
Cataract Ck at Sandy Hill(below Snake Ck) (AU204036)
Clouds Ck at Clouds Ck (AU204037)
Orara at Bawden Br (AU204041)
Sportsmans Ck at Gurranang Siding (AU204055)
Gordon Brook at Fine Flower (AU204067)
Bellinger at Thora (AU205002)
Nambucca at Bowraville (AU205006)
Macleay at Tia (Tia River) (AU206009)
Wollomombi at Coninside (AU206014)
Anomalie Q/P
Erreur GR4J
Erreur TOPM
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

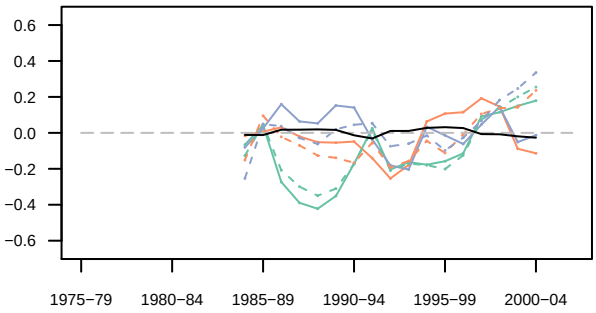
Apsley at Apsley Falls (AU206018)



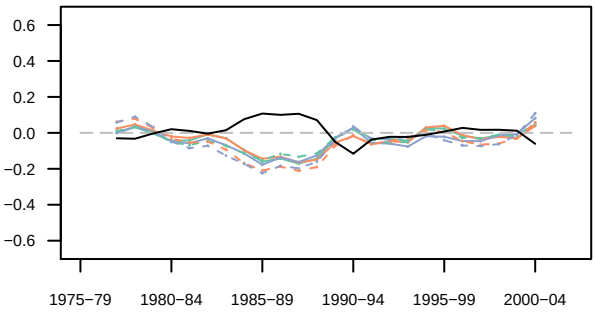
Salisbury Waters at near Dangar Falls (AU206025)



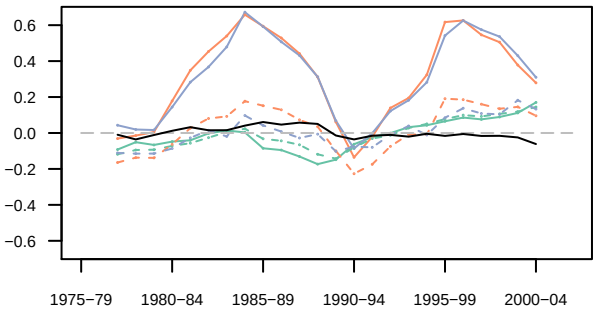
Mihi Ck at Abermala (AU206034)



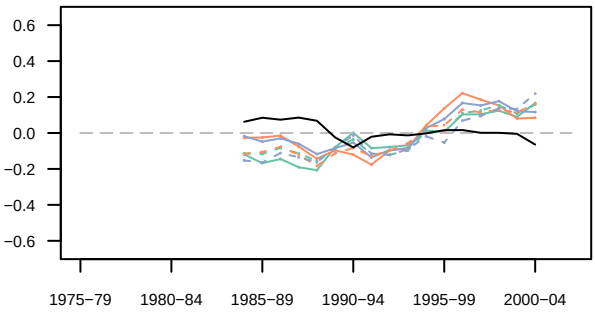
Forbes at Birdwood(Filly Flat) (AU207006)



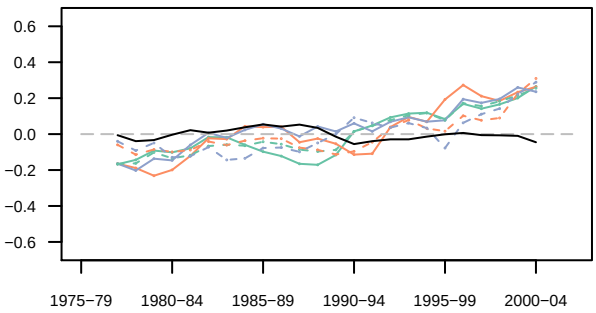
Ellenborough at D/S Bunnoo R Junction (AU207013)



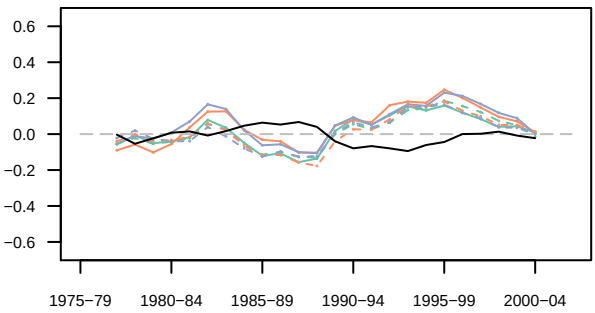
Hastings at Mount Seaview (AU207015)



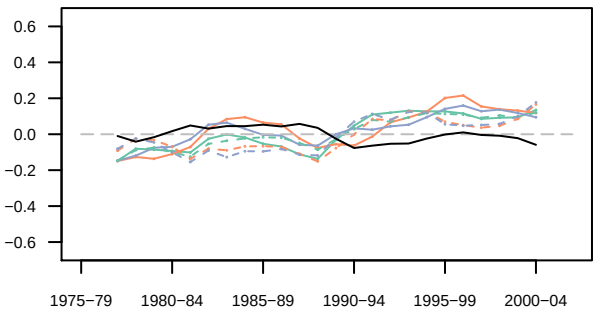
Manning at Rocks Crossing (Nowendoc River) (AU208005)



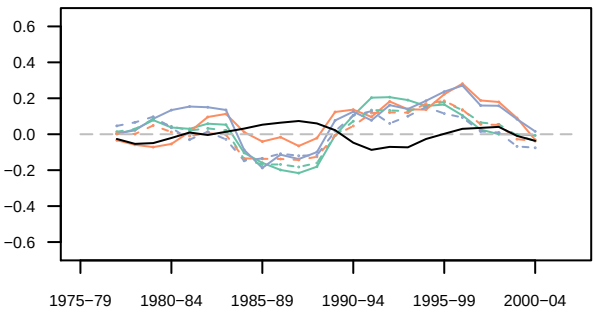
Barrington at Forbesdale (Causeway) (AU208006)

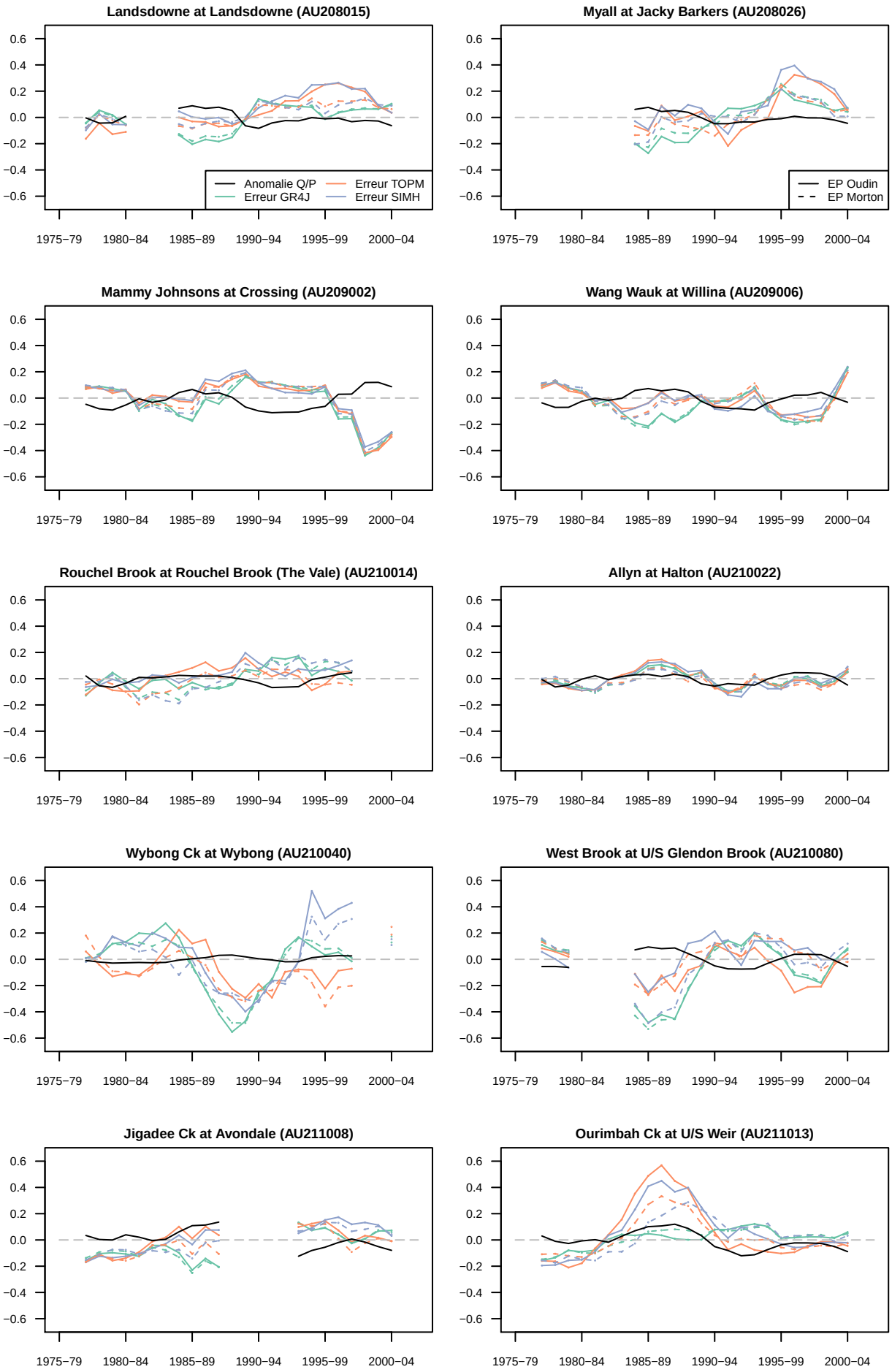


Nowendoc at Nowendoc (AU208007)

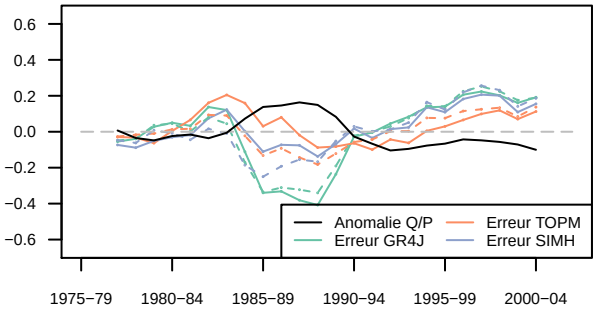


Barnard at Barry (AU208009)

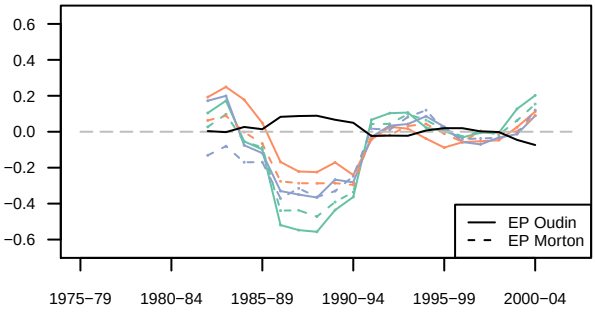



Landsdowne at Landsdowne (AU208015)
Myall at Jacky Barkers (AU208026)
Mammy Johnsons at Crossing (AU209002)
Wang Wauk at Willina (AU209006)
Rouchel Brook at Rouchel Brook (The Vale) (AU210014)
Allyn at Halton (AU210022)
Wybong Ck at Wybong (AU210040)
West Brook at U/S Glendon Brook (AU210080)
Jigadee Ck at Avondale (AU211008)
Ourimbah Ck at U/S Weir (AU211013)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

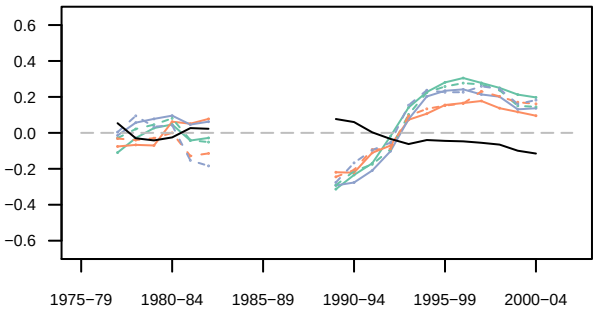
Wyong at Yarramalong (AU211014)



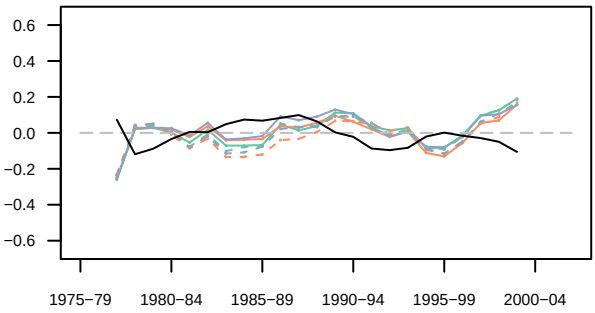
Kialla Ck at Pomeroy (AU212040)



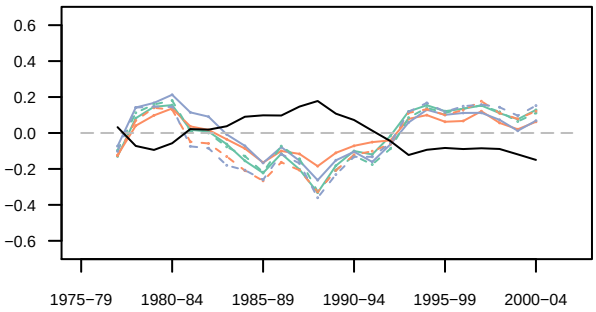
Shoalhaven at Warri (AU215002)



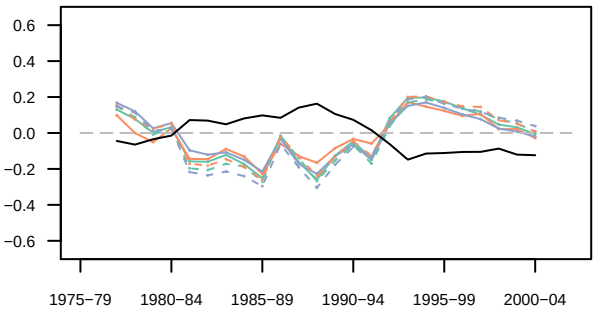
Corang at Hockeys (AU215004)



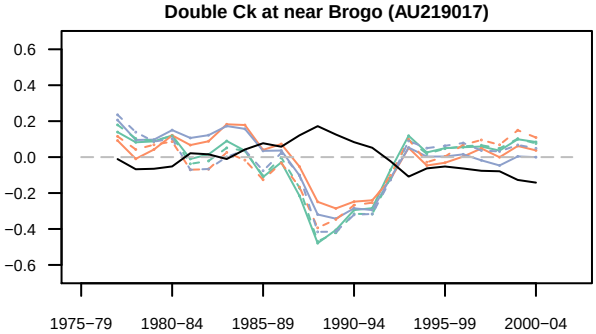
Tuross at Tuross Vale (AU218001)



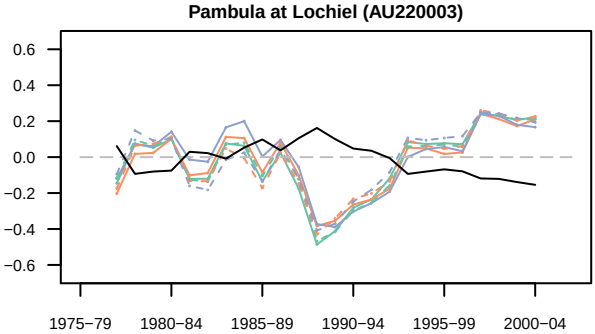
Wadbilliga at Wadbilliga (AU218007)



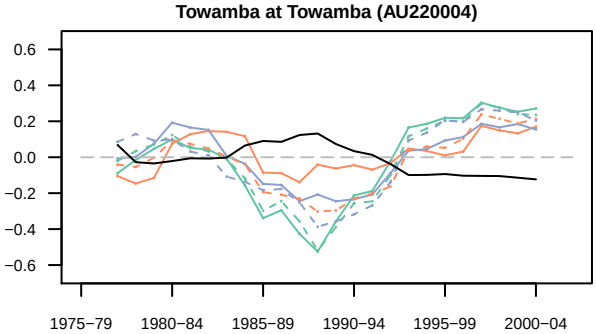
Double Ck at near Brogo (AU219017)



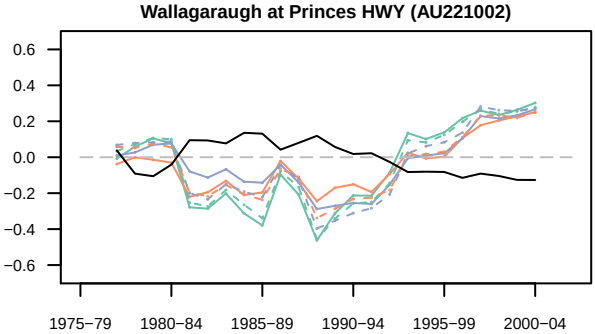
Pambula at Lochiel (AU220003)

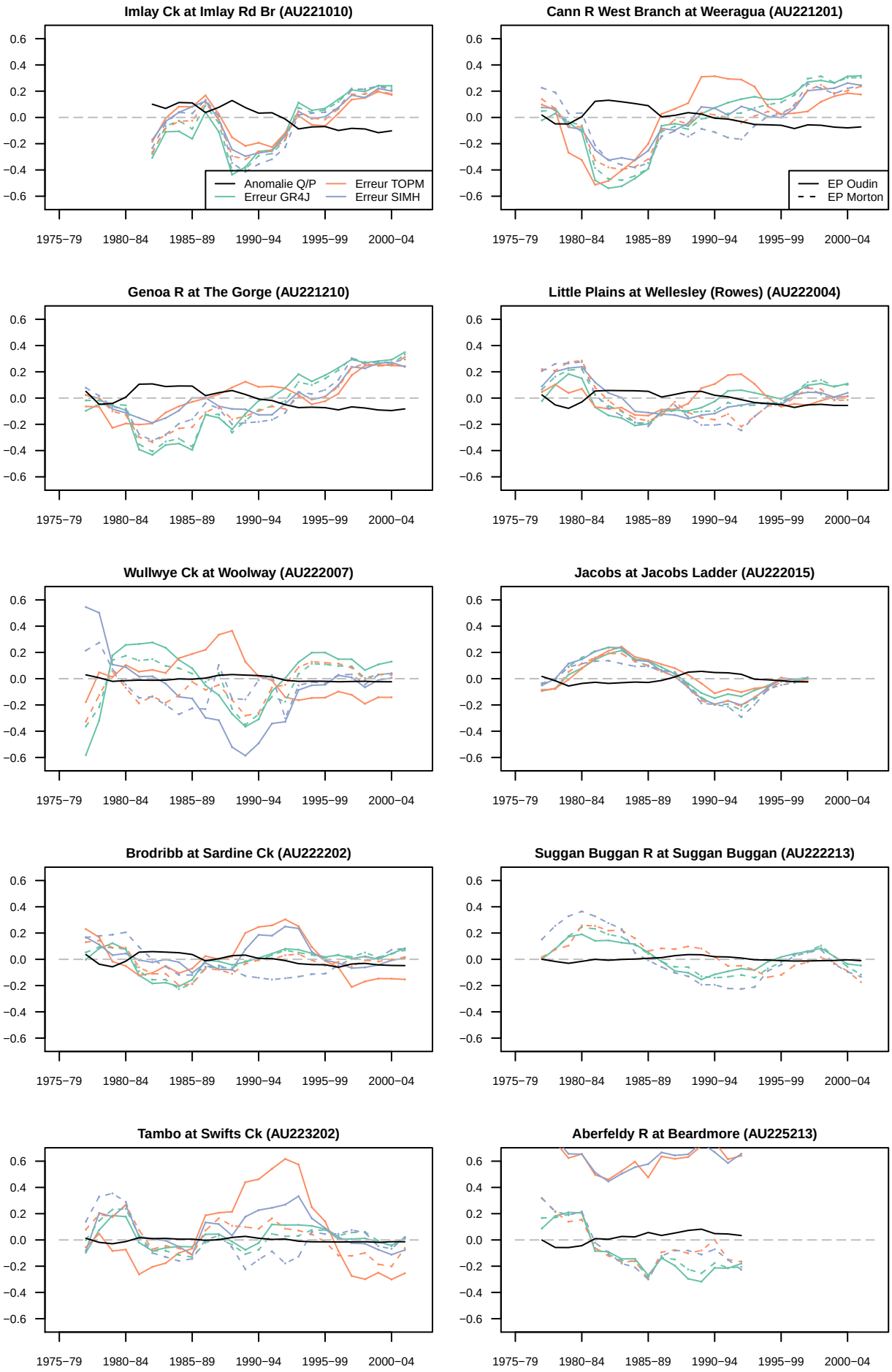


Towamba at Towamba (AU220004)

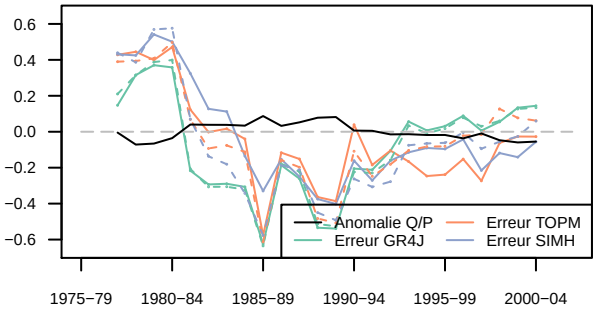


Wallagaraugh at Princes HWY (AU221002)

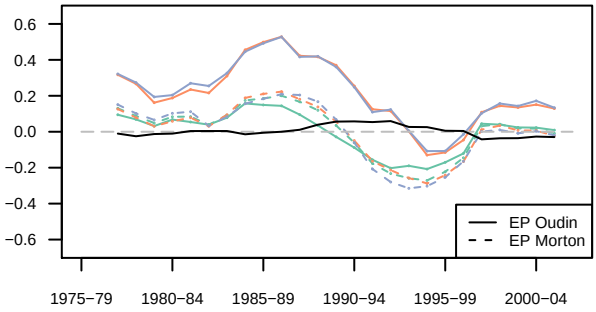



Imlay Ck at Imlay Rd Br (AU221010)
Cann R West Branch at Weeragua (AU221201)
Genoa R at The Gorge (AU221210)
Little Plains at Wellesley (Rowes) (AU222004)
Wullwye Ck at Woolway (AU222007)
Jacobs at Jacobs Ladder (AU222015)
Brodribb at Sardine Ck (AU222202)
Suggan Buggan R at Suggan Buggan (AU222213)
Tambo at Swifts Ck (AU223202)
Aberfeldy R at Beardmore (AU225213)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

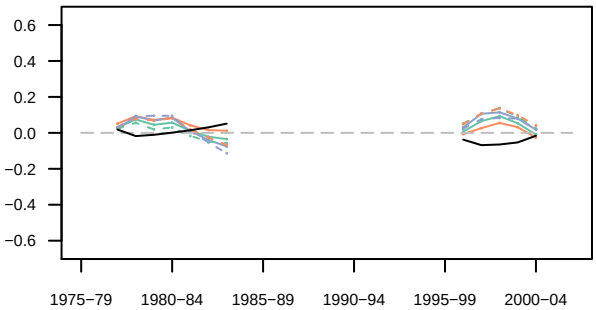
Freestone Ck at Briagolong (AU225218)



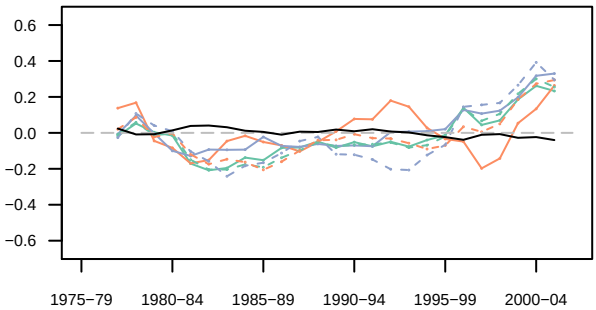
Macalister at Glencairn (AU225219)



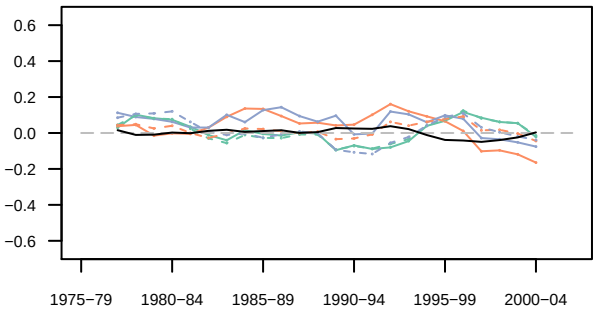
Narracan Ck at Thorpdale (AU226218)



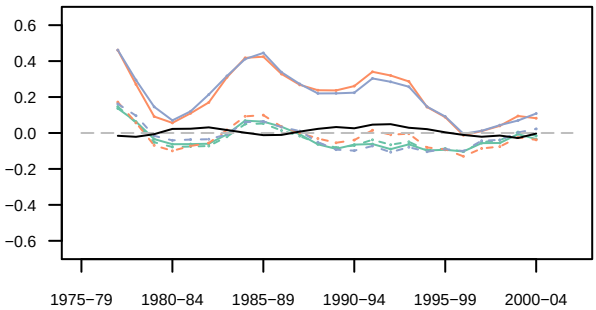
Tarra R at Yarram (AU227200)



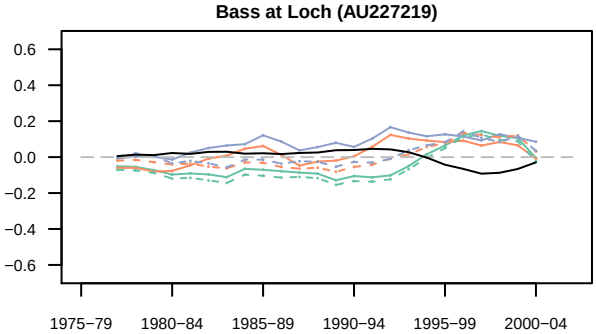
Tarwin at Meeniyan (AU227202)



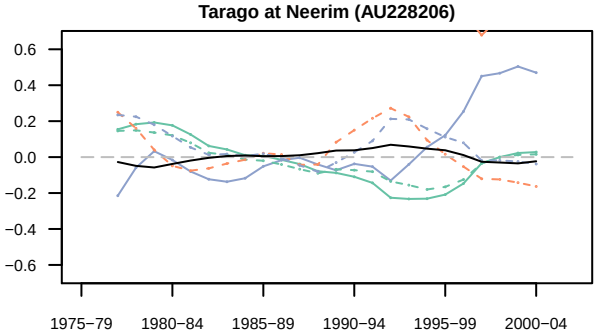
Agnes at Toora (AU227211)



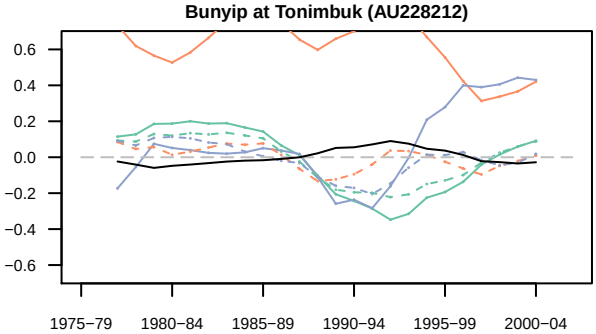
Bass at Loch (AU227219)



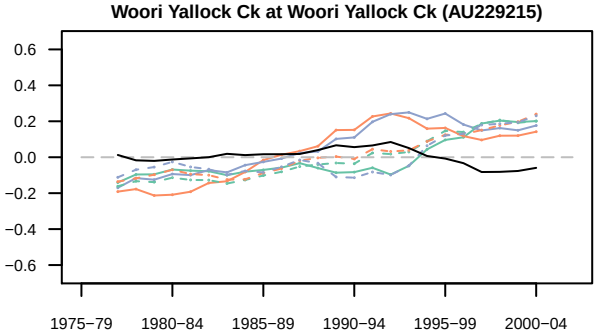
Tarago at Neerim (AU228206)

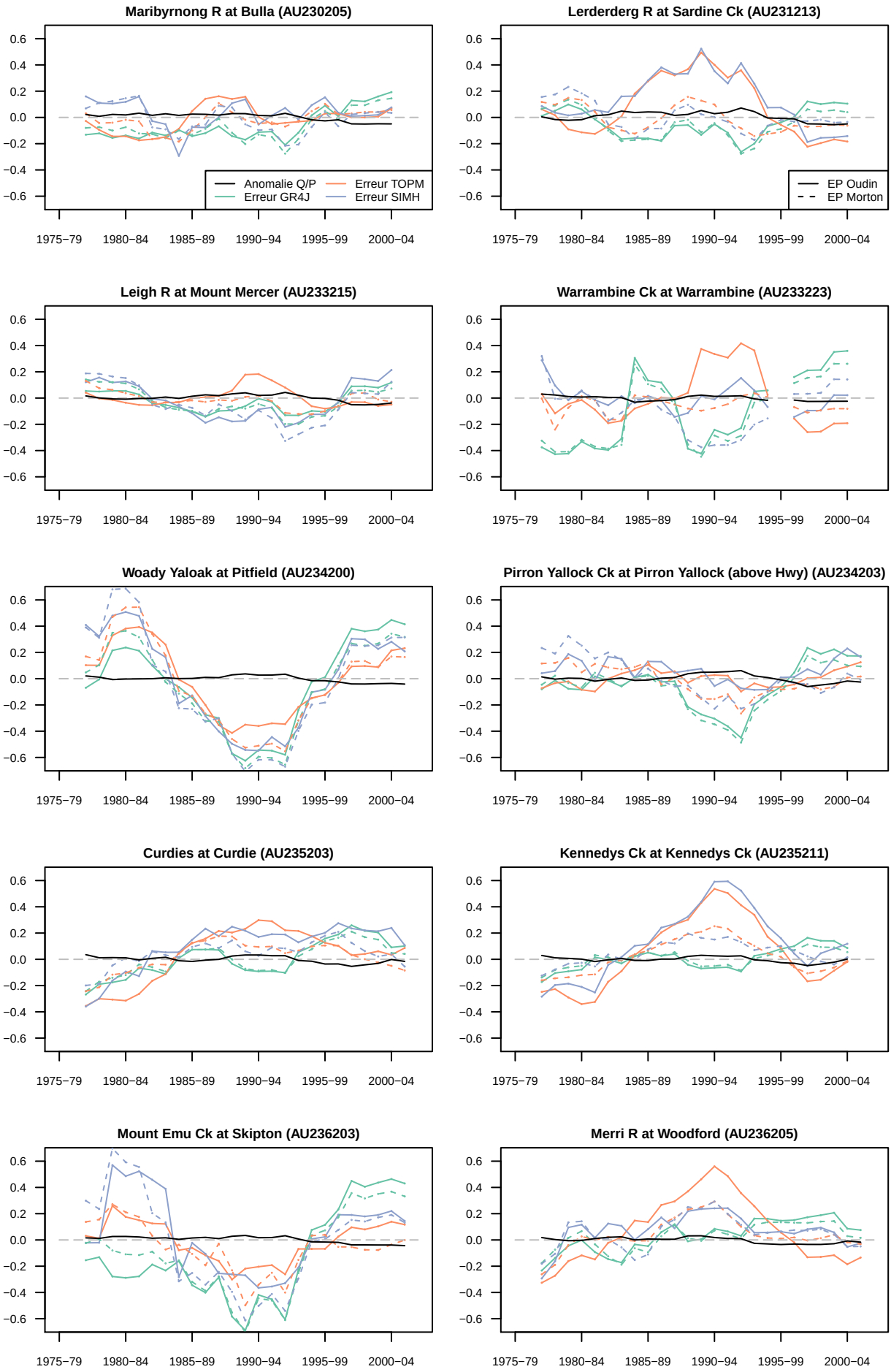


Bunyip at Tonimbuk (AU228212)

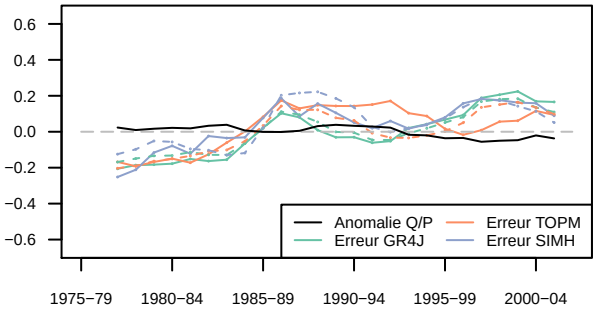


Woori Yallock Ck at Woori Yallock Ck (AU229215)

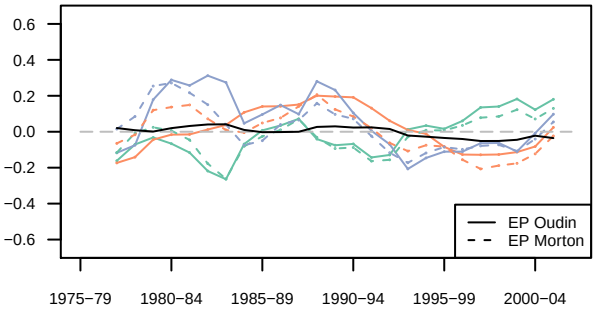



Maribyrnong R at Bulla (AU230205)
Lerderderg R at Sardine Ck (AU231213)
Leigh R at Mount Mercer (AU233215)
Warrambine Ck at Warrambine (AU233223)
Woady Yaloak at Pitfield (AU234200)
Pirron Yallock Ck at Pirron Yallock (above Hwy) (AU234203)
Curdies at Curdie (AU235203)
Kennedys Ck at Kennedys Ck (AU235211)
Mount Emu Ck at Skipton (AU236203)
Merri R at Woodford (AU236205)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

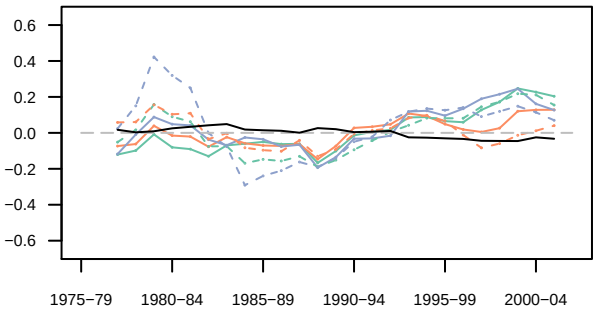
Brucknell Ck at Cudgee (AU236212)



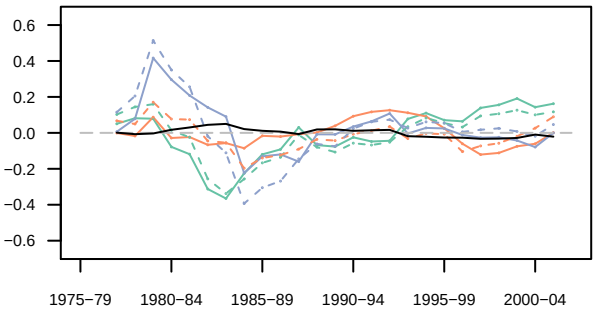
Moyne R at Toolong (AU237200)



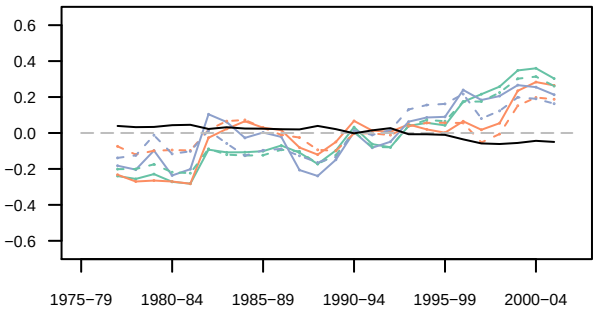
Darlot Ck at Homerton Br (AU237205)



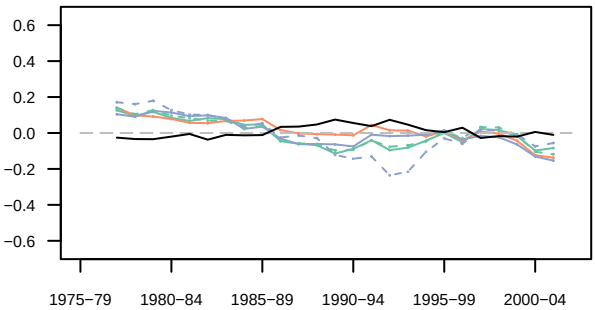
Eumeralla at Codrington (AU237206)



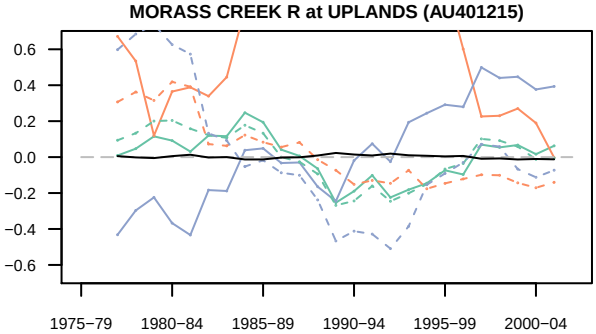
Wando at Wando Vale (AU238223)



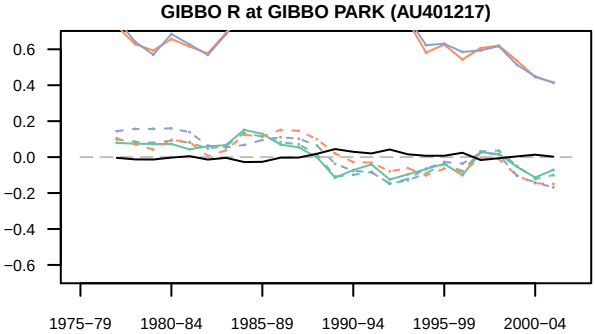
SNOWY CREEK R at BELOW GRANITE FLAT (AU401210)



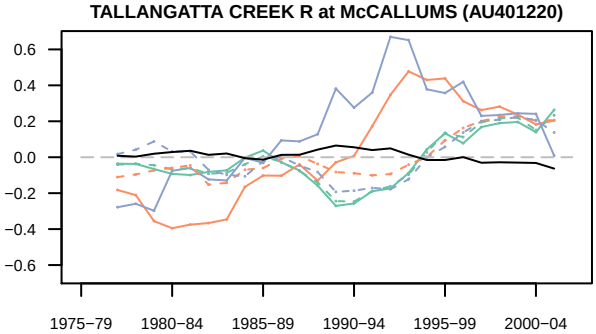
MORASS CREEK R at UPLANDS (AU401215)



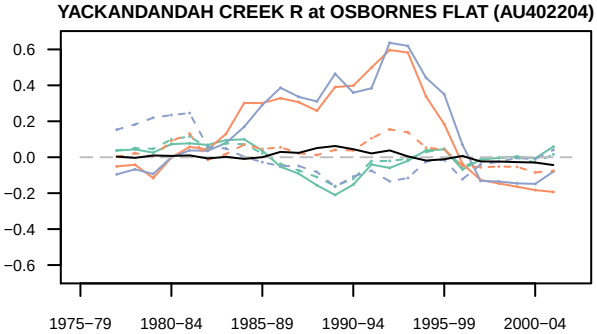
GIBBO R at GIBBO PARK (AU401217)

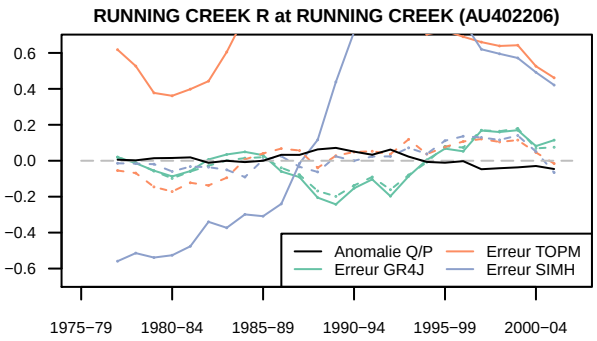


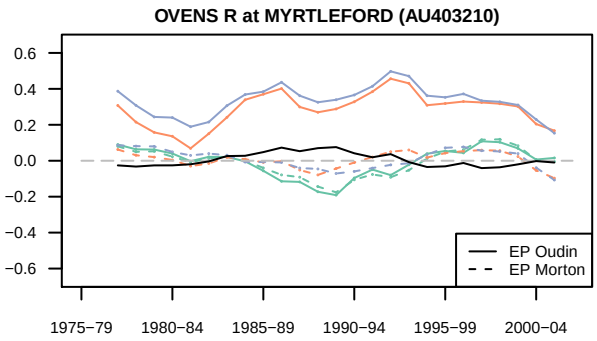
TALLANGATTA CREEK R at McCALLUMS (AU401220)

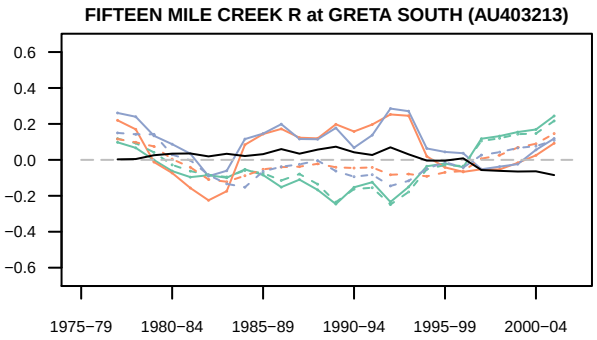


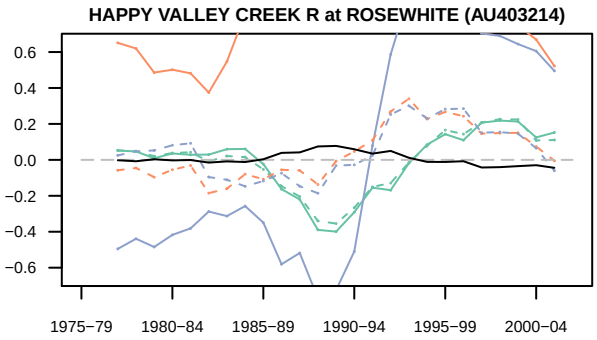
YACKANDANDAH CREEK R at OSBORNES FLAT (AU402204)

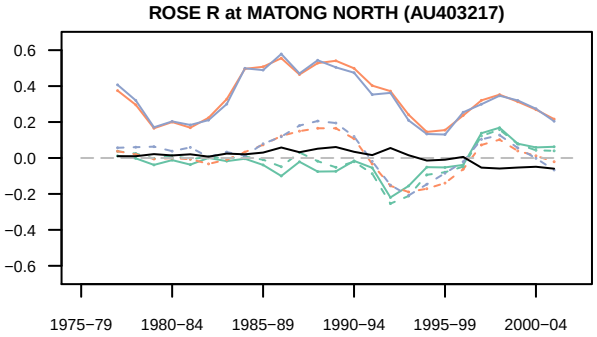


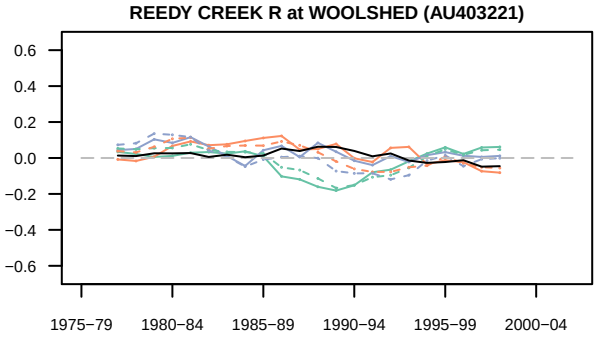

RUNNING CREEK R at RUNNING CREEK (AU402206)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

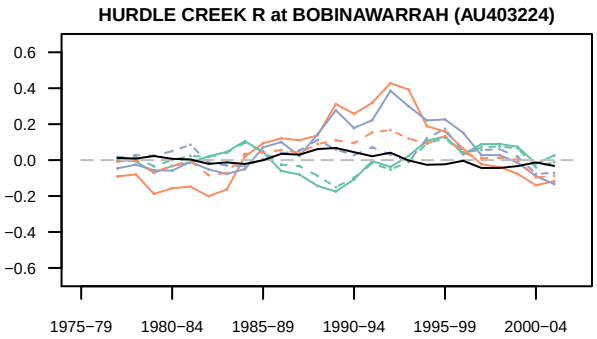

OVENS R at MYRTLEFORD (AU403210)
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

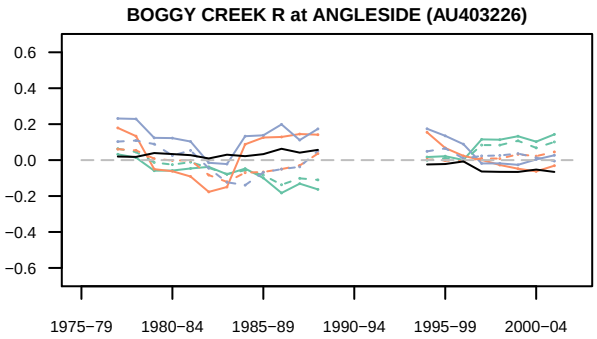

FIFTEEN MILE CREEK R at GRETA SOUTH (AU403213)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

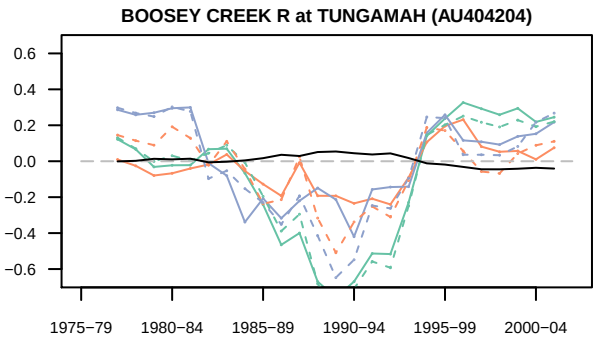

HAPPY VALLEY CREEK R at ROSEWHITE (AU403214)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

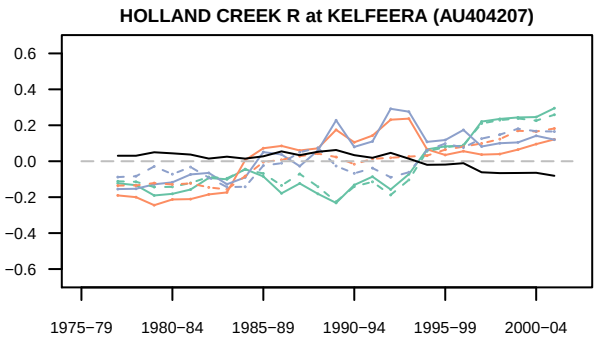

ROSE R at MATONG NORTH (AU403217)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04


REEDY CREEK R at WOOLSHED (AU403221)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

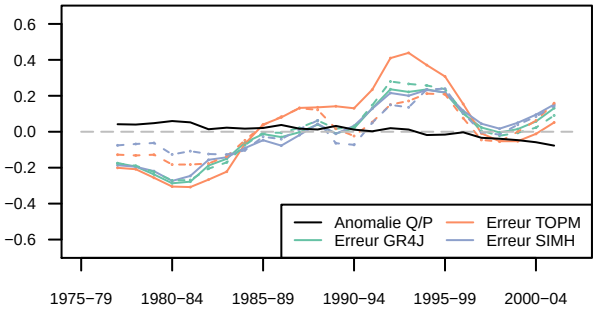

HURDLE CREEK R at BOBINAWARRAH (AU403224)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04


BOGGY CREEK R at ANGLESIDE (AU403226)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

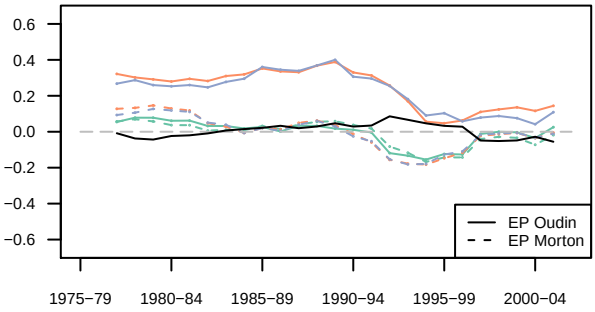

BOOSEY CREEK R at TUNGAMAH (AU404204)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04


HOLLAND CREEK R at KELFEERA (AU404207)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

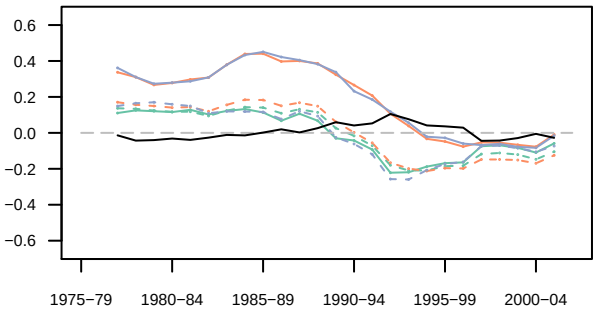
MOONEE CREEK R at LIMA (AU404208)



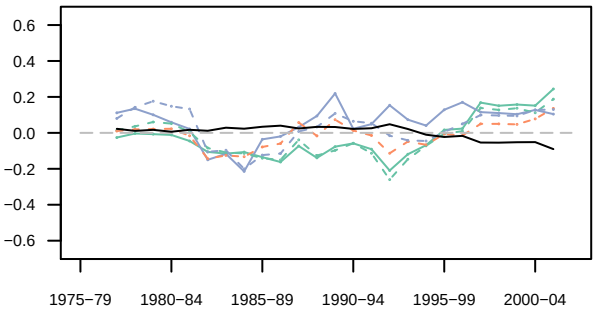
MURRINDINDI R at MURRINDINDI ABOVE COLWELLS (AU405205)



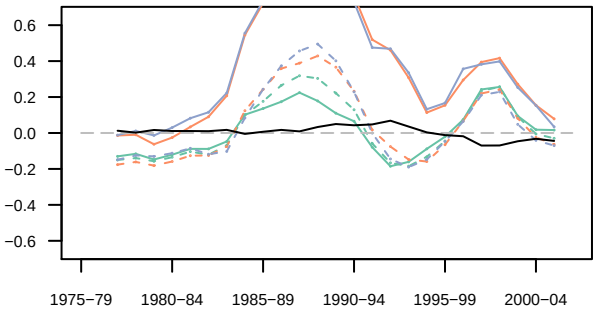
ACHERON R at TAGGERTY (AU405209)



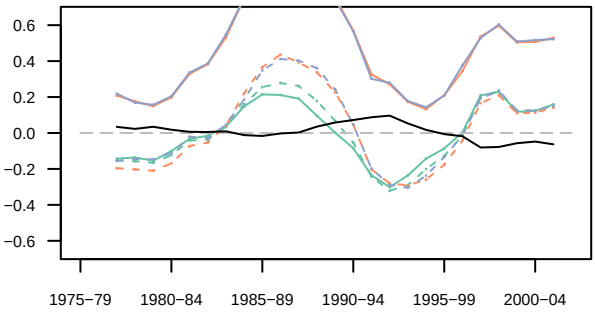
SUNDAY CREEK R at TALLAROOK (AU405212)



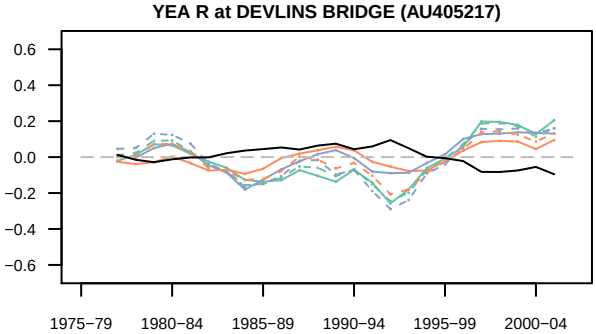
DELATITE R at TONGA BRIDGE (AU405214)



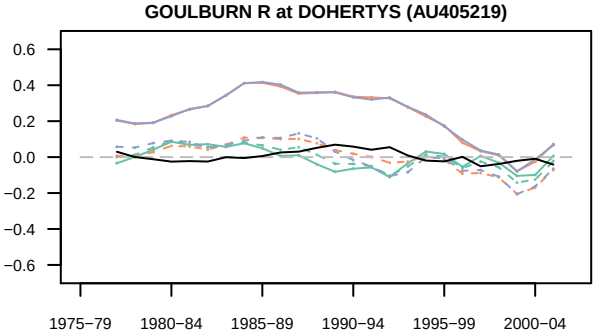
HOWQUA R at GLEN ESK (AU405215)



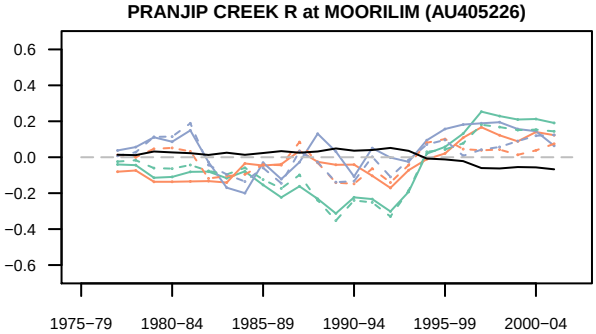
YEA R at DEVLINS BRIDGE (AU405217)



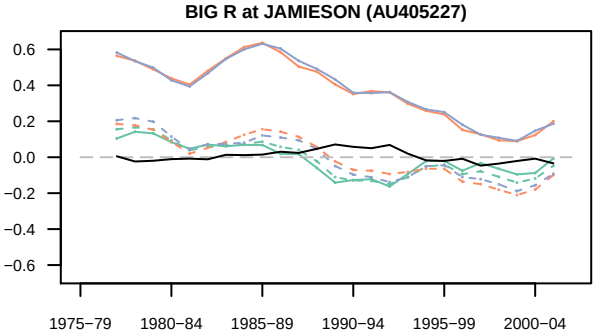
GOULBURN R at DOHERTYS (AU405219)

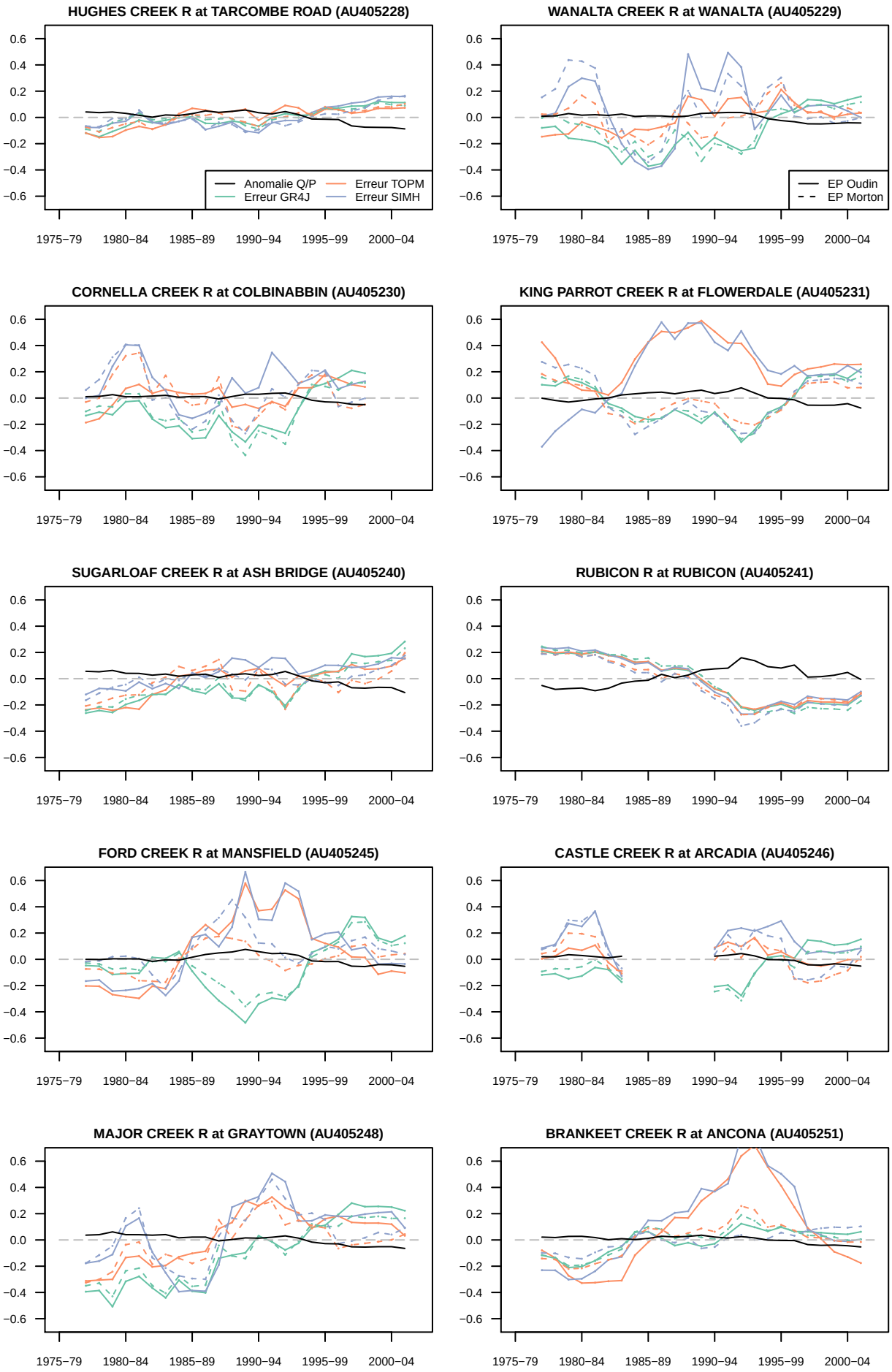


PRANJIP CREEK R at MOORILIM (AU405226)

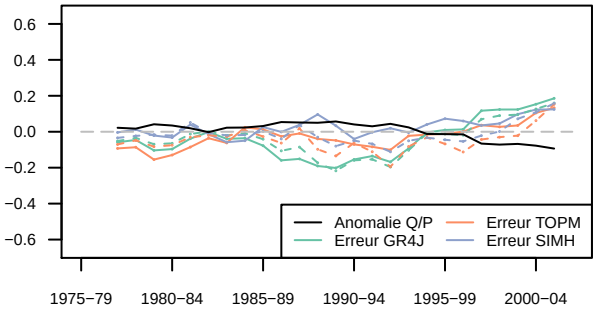


BIG R at JAMIESON (AU405227)

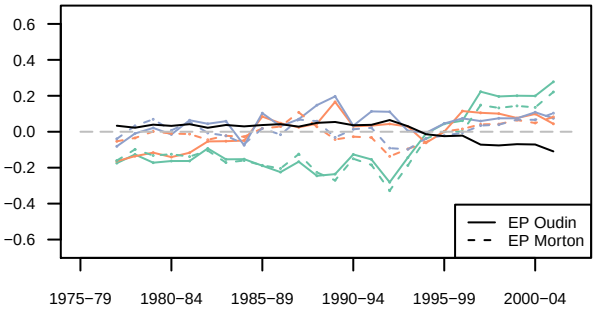



HUGHES CREEK R at TARCOMBE ROAD (AU405228)
WANALTA CREEK R at WANALTA (AU405229)
CORNELLA CREEK R at COLBINABBIN (AU405230)
KING PARROT CREEK R at FLOWERDALE (AU405231)
SUGARLOAF CREEK R at ASH BRIDGE (AU405240)
RUBICON R at RUBICON (AU405241)
FORD CREEK R at MANSFIELD (AU405245)
CASTLE CREEK R at ARCADIA (AU405246)
MAJOR CREEK R at GRAYTOWN (AU405248)
BRANKEET CREEK R at ANCONA (AU405251)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

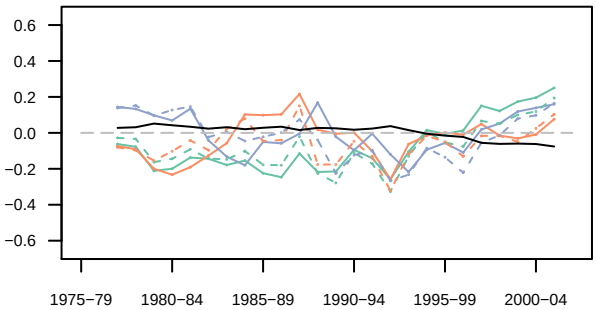
HOME CREEK R at YARCK (AU405274)



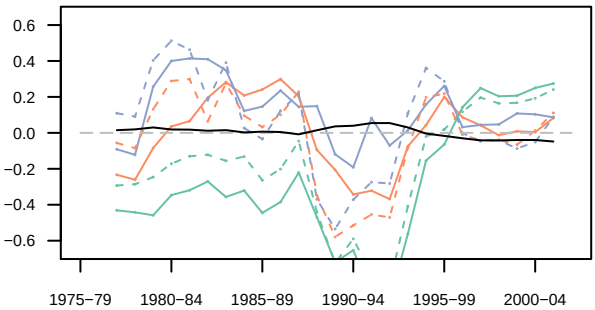
CAMPASPE R at REDESDALE (AU406213)



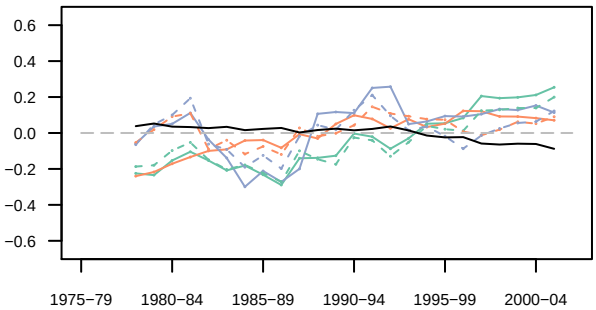
AXE CREEK R at LONGLEA (AU406214)



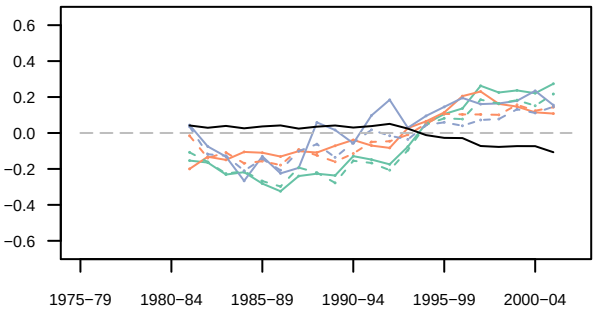
MOUNT PLEASANT CREEK R at RUNNYMEDE (AU406224)



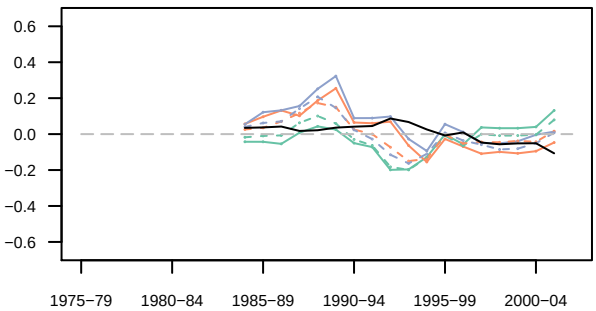
MOUNT IDA CREEK R at DERRINAL (AU406226)



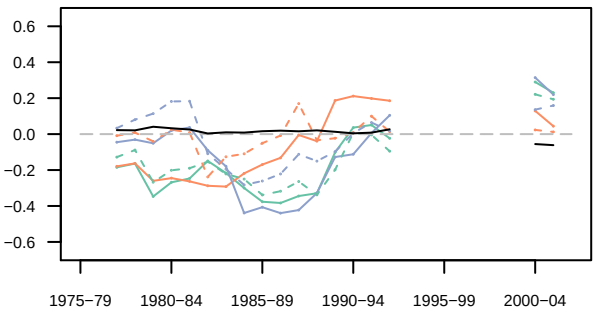
WILD DUCK CREEK R at US OF HEATHCOTE−MIA MIA RD (AU40623



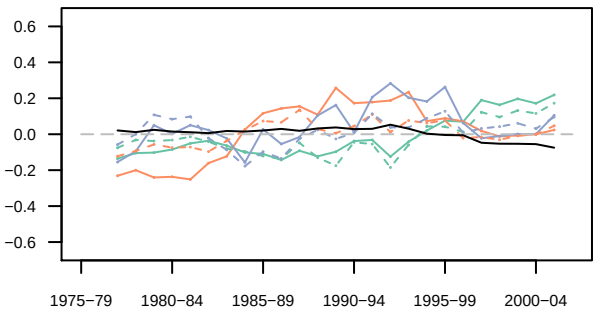
COLIBAN R at SPRINGHILL−TYLDEN ROAD (AU406250)



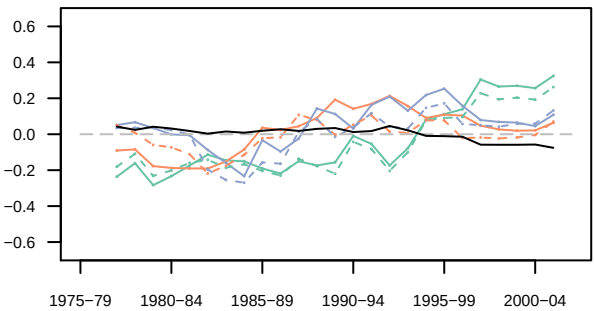
MCCALLUMS CREEK R at CARISBROOK (AU407213)

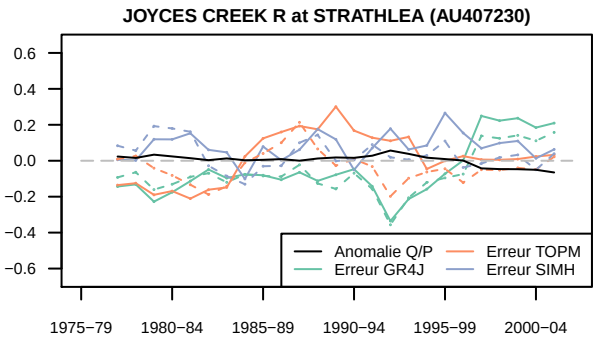


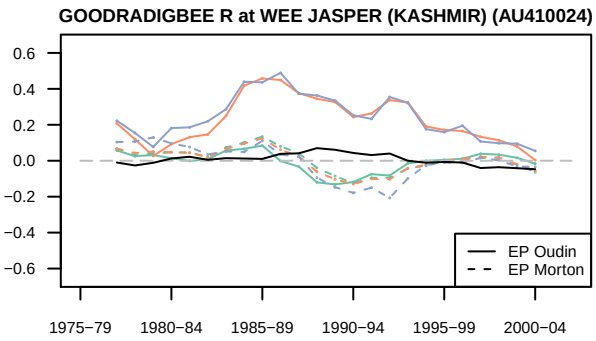
LODDON R at NEWSTEAD (AU407215)

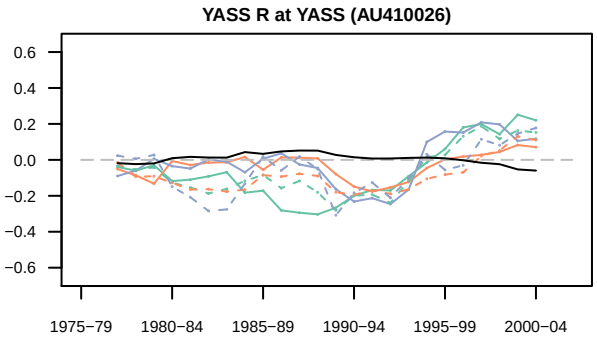


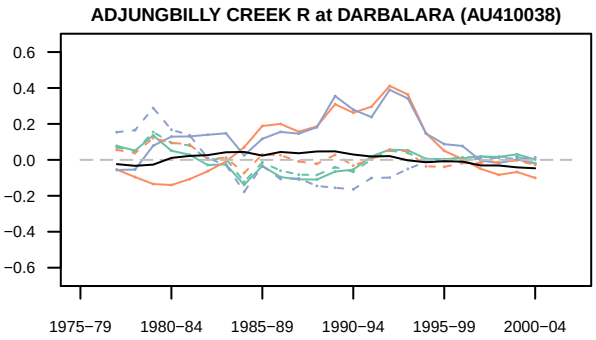
TULLAROOP CREEK R at CLUNES (AU407222)

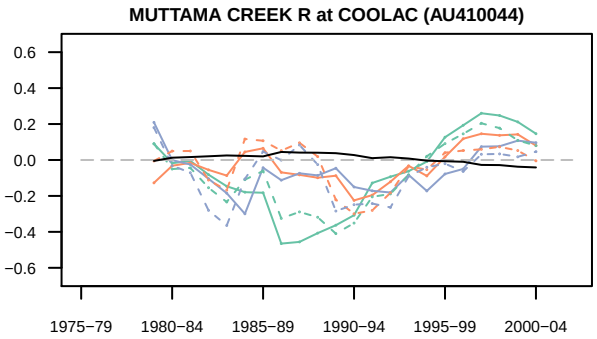


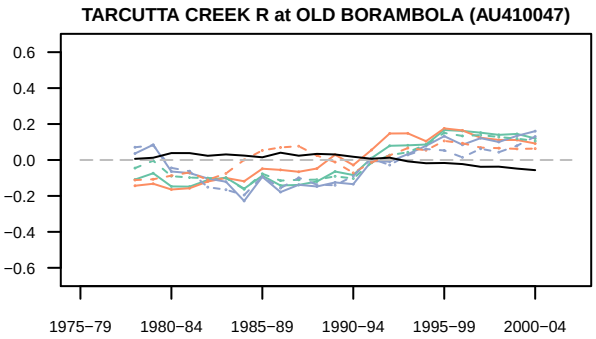

JOYCES CREEK R at STRATHLEA (AU407230)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

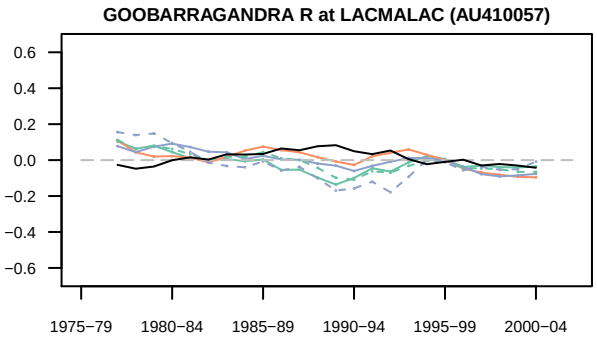

GOODRADIGBEE R at WEE JASPER (KASHMIR) (AU410024)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
EP Oudin
EP Morton
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

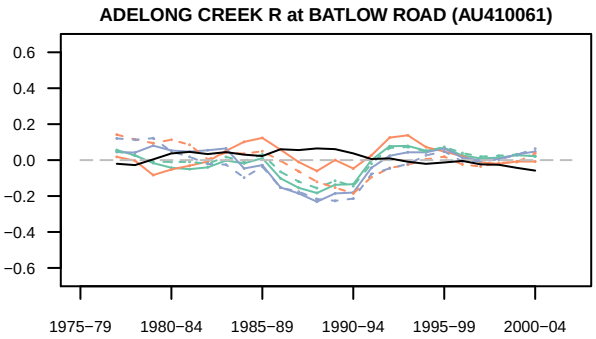

YASS R at YASS (AU410026)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

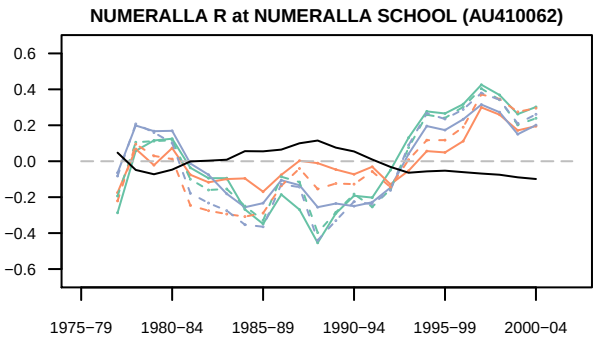

ADJUNGBILLY CREEK R at DARBALARA (AU410038)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

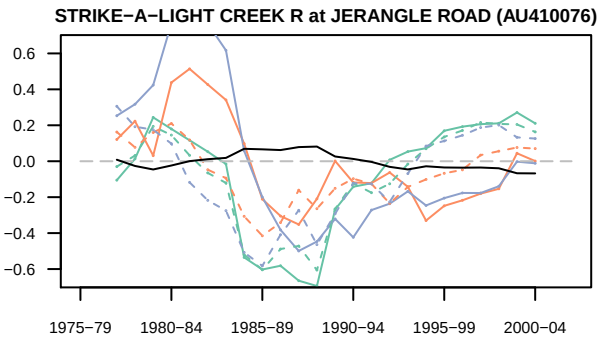

MUTTAMA CREEK R at COOLAC (AU410044)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04


TARCUTTA CREEK R at OLD BORAMBOLA (AU410047)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

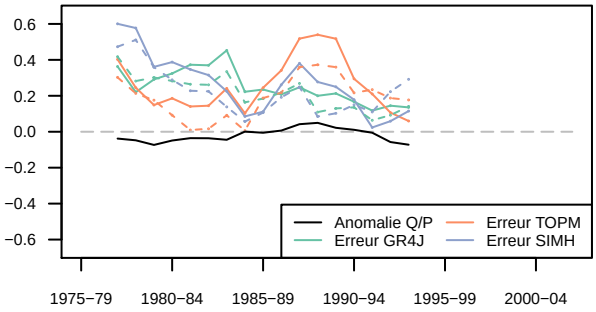

GOOBARRAGANDRA R at LACMALAC (AU410057)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04


ADELONG CREEK R at BATLOW ROAD (AU410061)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

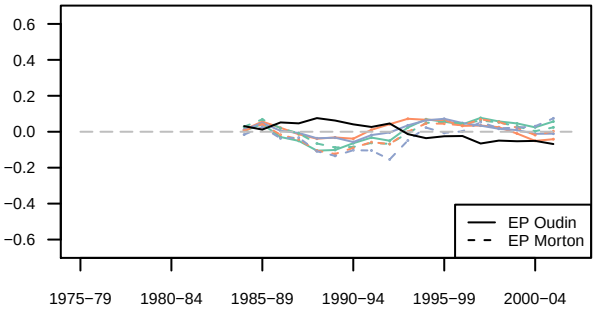

NUMERALLA R at NUMERALLA SCHOOL (AU410062)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04


STRIKE−A−LIGHT CREEK R at JERANGLE ROAD (AU410076)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

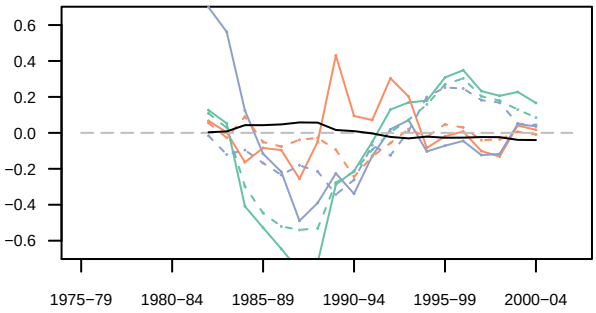
COOMA CREEK R at COOMA NO.2 (THE GRANGE) (AU410081)



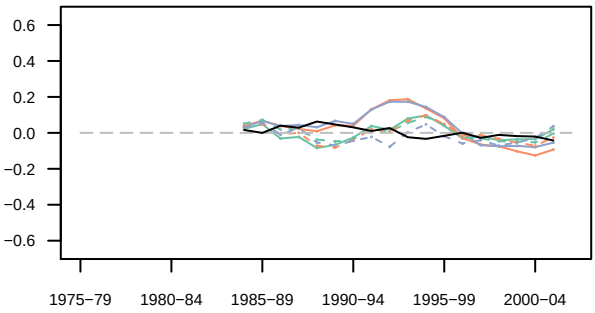
JOUNAMA CREEK R at ABOVE JOUNAMA PONDAGE (AU410094)



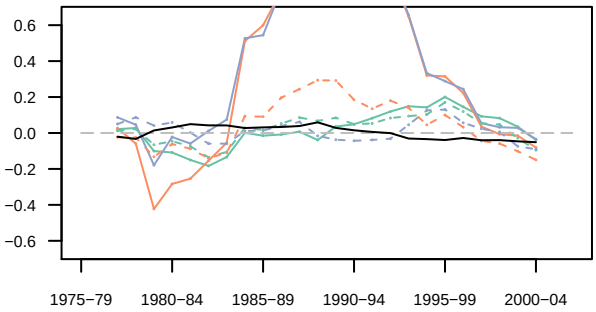
MICALIGO CREEK R at MICHELAGO (AU410141)



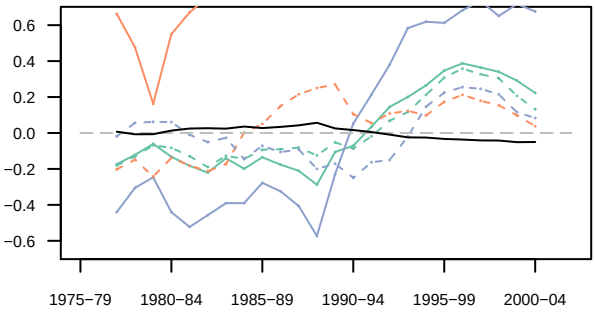
YARRANGOBILLY R at RAVINE (AU410574)



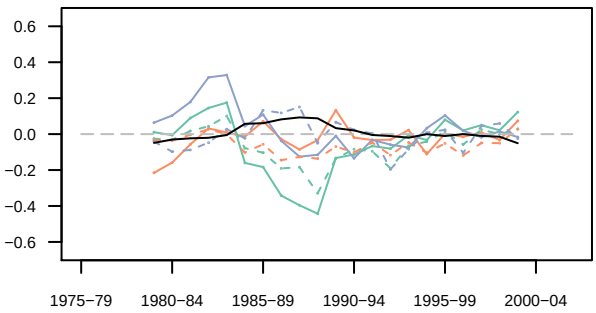
PADDY'S R at RIVERLEA (AU410713)



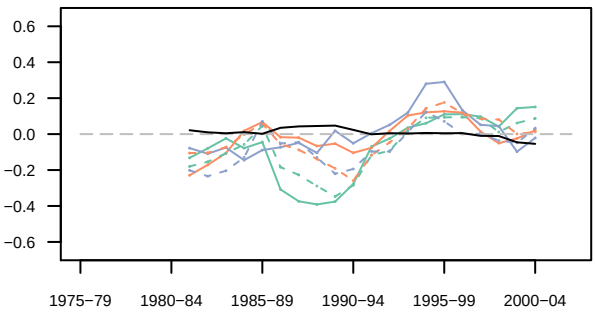
GUDGENBY R at TENNENT (AU410731)



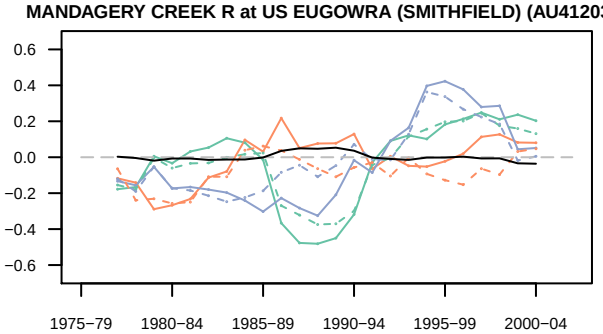
BUTMAROO CREEK R at BUTMAROO (AU411003)



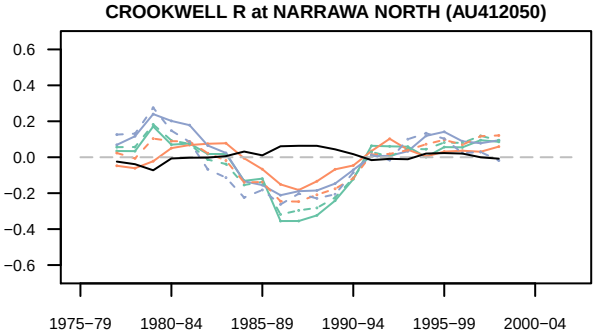
BOOROWA R at PROSSERS CROSSING (AU412029)

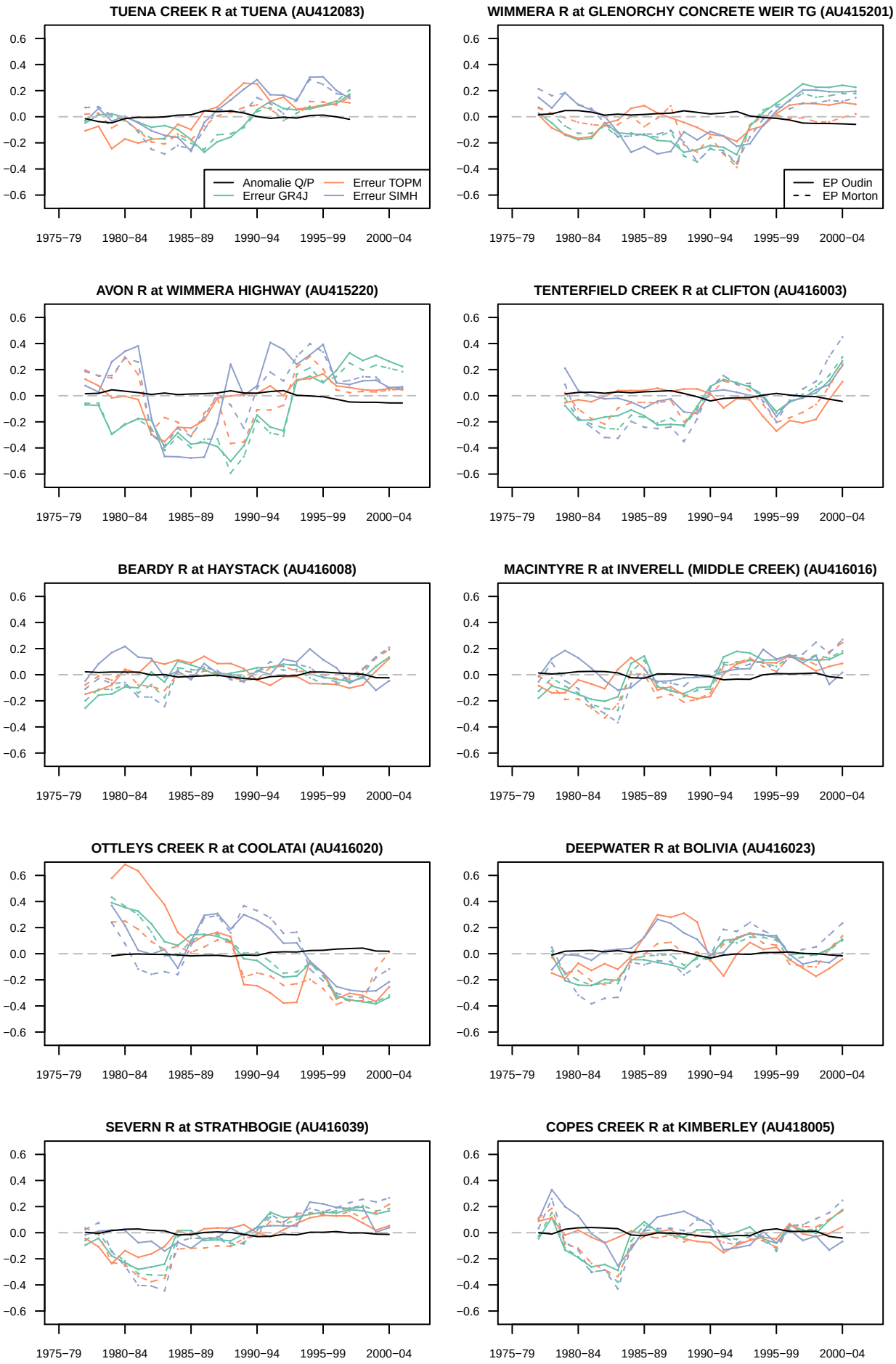


MANDAGERY CREEK R at US EUGOWRA (SMITHFIELD) (AU412030

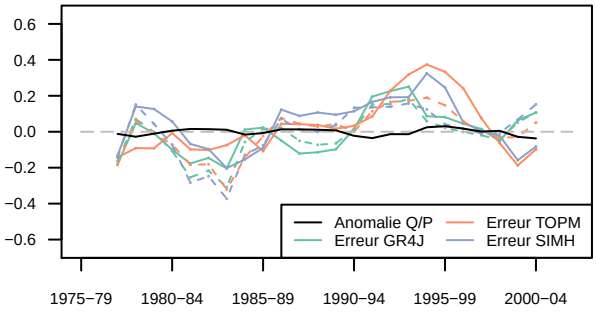


CROOKWELL R at NARRAWA NORTH (AU412050)

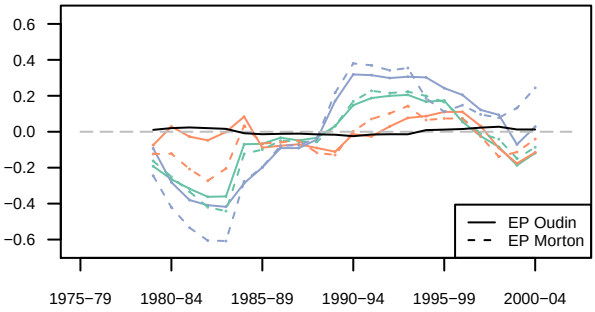



TUENA CREEK R at TUENA (AU412083)
WIMMERA R at GLENORCHY CONCRETE WEIR TG (AU415201)
AVON R at WIMMERA HIGHWAY (AU415220)
TENTERFIELD CREEK R at CLIFTON (AU416003)
BEARDY R at HAYSTACK (AU416008)
MACINTYRE R at INVERELL (MIDDLE CREEK) (AU416016)
OTTLEYS CREEK R at COOLATAI (AU416020)
DEEPWATER R at BOLIVIA (AU416023)
SEVERN R at STRATHBOGIE (AU416039)
COPES CREEK R at KIMBERLEY (AU418005)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04

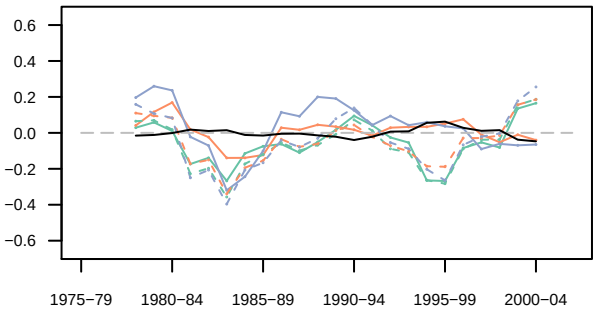
GWYDIR R at YARROWYCK (AU418014)



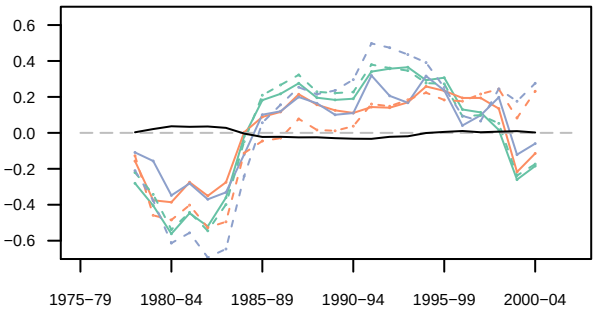
MYALL CREEK R at MOLROY (AU418017)



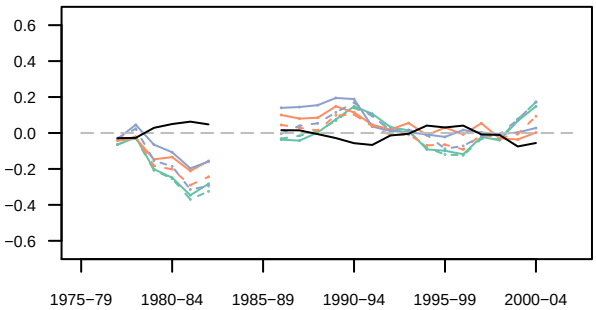
LAURA CREEK R at LAURA (AU418021)



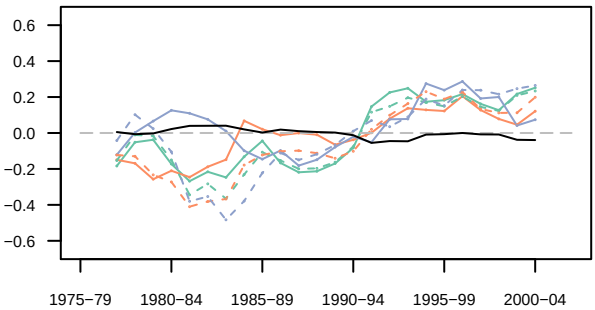
HALLS CREEK R at BINGARA (AU418025)



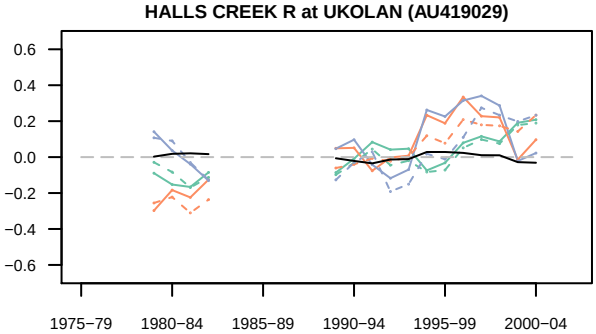
HORTON R at HORTON DAM SITE (AU418027)



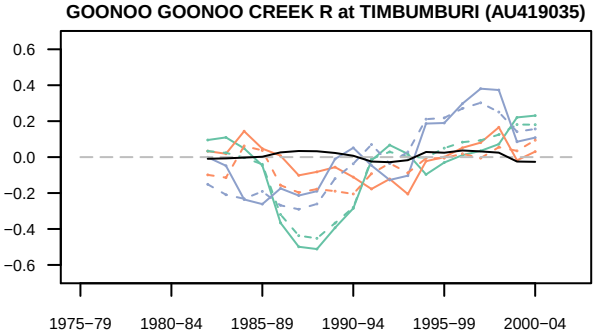
COCKBURN R at MULLA CROSSING (AU419016)



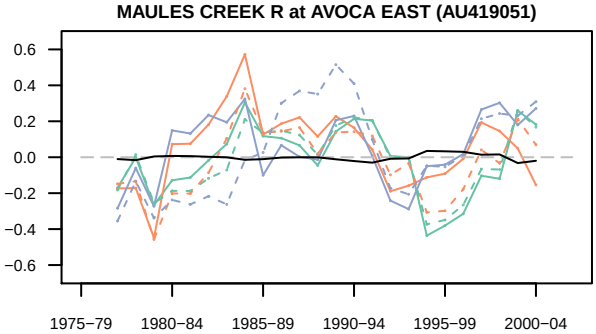
HALLS CREEK R at UKOLAN (AU419029)



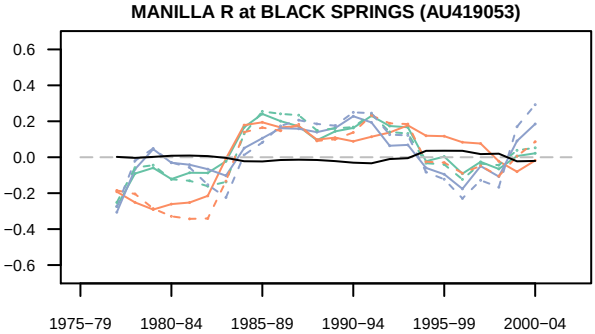
GOONOO GOONOO CREEK R at TIMBUMBURI (AU419035)

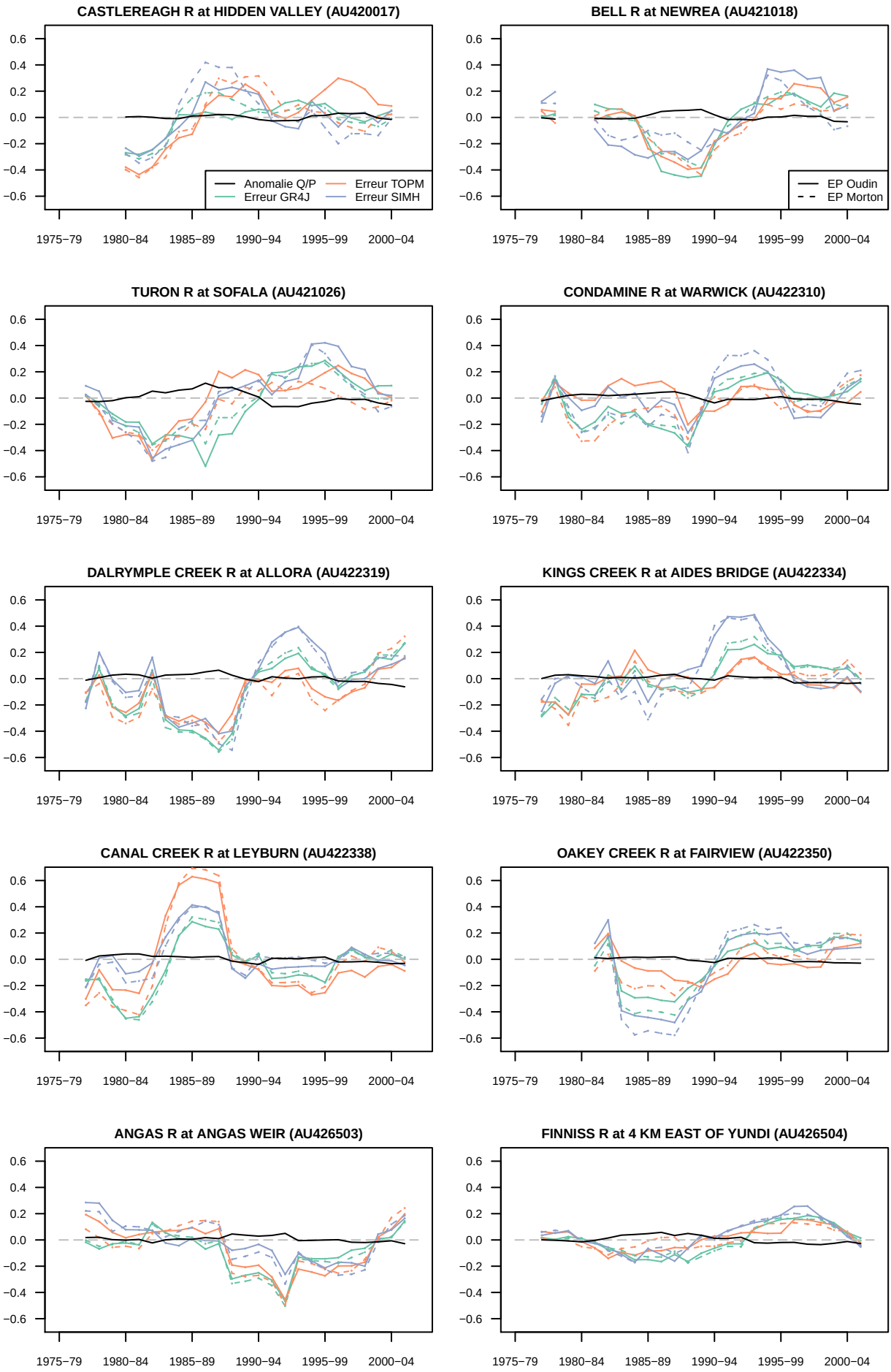


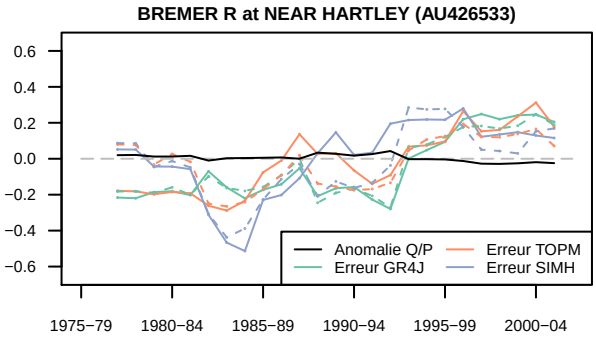
MAULES CREEK R at AVOCA EAST (AU419051)



MANILLA R at BLACK SPRINGS (AU419053)




CASTLEREAGH R at HIDDEN VALLEY (AU420017)
BELL R at NEWREA (AU421018)
TURON R at SOFALA (AU421026)
CONDAMINE R at WARWICK (AU422310)
DALRYMPLE CREEK R at ALLORA (AU422319)
KINGS CREEK R at AIDES BRIDGE (AU422334)
CANAL CREEK R at LEYBURN (AU422338)
OAKEY CREEK R at FAIRVIEW (AU422350)
ANGAS R at ANGAS WEIR (AU426503)
FINNISS R at 4 KM EAST OF YUNDI (AU426504)
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
EP Oudin
EP Morton
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04


BREMER R at NEAR HARTLEY (AU426533)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
Anomalie Q/P
Erreur TOPM
Erreur GR4J
Erreur SIMH
1975−79
1980−84
1985−89
1990−94
1995−99
2000−04