



HAL
open science

Les tests statistiques de significativité appliqués à l'hydrologie

Imane Haddadi

► **To cite this version:**

Imane Haddadi. Les tests statistiques de significativité appliqués à l'hydrologie. Hydrologie. 2016. hal-02604535

HAL Id: hal-02604535

<https://hal.inrae.fr/hal-02604535>

Submitted on 16 May 2020

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.



1, rue Pierre Gilles de Gennes
CS 10030
92761 Antony France



24, avenue des Landais
63170 Aubière France

Master 1 Statistiques et Traitements des données – 2015/2016

Imane HADDADI

LES TESTS STATISTIQUES DE SIGNIFICATIVITÉ APPLIQUÉS À L'HYDROLOGIE



Le bassin d'orage de la Vallée du Bédât à Clermont-Ferrand, Photo I. Haddadi.

Maîtres de stage

M. Olivier DELAIGUE
M. Guillaume THIREL

Enseignante de suivi

Mme. Anne-Françoise YAO LAFOURCADE

Formation master en Laboratoire de Mathématiques à l'UBP Clermont-Ferrand
Stage de 3 mois du 01 avril au 30 juin 2016 effectué à l'IRSTEA d'Antony
Unité de recherche Hydrosystèmes et Bioprocédés, Équipe hydrologie des bassins versants

Remerciements

À travers ce mémoire, je tiens à exprimer mes sincères remerciements à tous ceux qui ont contribué à l'aboutissement de ce travail :

Mes hommages pour le centre d'IRSTEA d'Antony, qui m'a inculqué des connaissances pertinentes en hydrologie, en programmation et en analyse de données, en l'occurrence l'équipe hydrologie des bassins versants qui m'a bien accueillie. Mes chaleureux remerciements pour M. Guillaume Thirel et M. Olivier Delaigue, mes maîtres de stage qui m'ont offert une atmosphère propice pour une réussite souhaitée de mon stage imbibée de patience, de sympathie, de disponibilité, et de conseils judicieux. Je tiens à remercier aussi mes colocataires de bureau Morgane et Arnaud avec qui j'avais eu de nombreuses discussions sympathiques, de bonne courtoisie et qui m'ont supportée pendant les trois mois de mon stage. Un remerciement tout spécial pour Carine et Andrea que j'ai trop dérangés (j'avoue !) avec toutes mes questions et qui n'ont jamais hésité de répondre. Mes respects aussi pour tous les autres membres du personnel de l'équipe : Charles, Pierre, Vazken, Carina, Julie, Maria-Helena, Cédric, Alban, Laure, Léonard, Sylvia, Philippe, Julien, Manon et Jeffrey pour leur accueil chaleureux, leur soutien, leur serviabilité, et leur bonne humeur.

Ma gratitude pour mon université, mes professeurs qui m'ont appris un savoir-faire précieux qui pourrait être une valeur ajoutée pour le développement et le progrès humain, en particulier Mme. Anne-Françoise Yao-Lafourcade mon enseignante de suivi.

Enfin, un grand merci pour ma famille, qui m'a toujours accompagnée avec amour et sérénité pour dépasser les obstacles moraux et matériels.

Résumé

L'objectif de ce stage était de proposer des tests statistiques dont l'application rigoureuse permettrait d'appuyer l'évaluation des performances des modèles hydrologiques pluie-débit, que j'ai ensuite appliqués à des modèles hydrologiques de mon équipe d'accueil, à savoir GR4J, GR5J et GR6J.

Pour cela, j'ai utilisé un échantillon de bassins versants qui devait représenter diverses conditions climatiques et hydrologiques. L'assemblage de différentes données : pluie, débit, et évapotranspiration collectées a permis d'élaborer un grand échantillon de 605 bassins versants.

Tout d'abord, nous avons étudié les données pour déterminer le respect ou non des différentes conditions d'application des tests statistiques, et pour cela trois étapes ont été suivies. La première a concerné l'étude de l'autocorrélation des débits observés des bassins versants ; les résultats ont confirmé que les bassins versants de grandes surfaces sont plus autocorrélés que ceux de petites surfaces. Ensuite, nous avons étudié la corrélation et la corrélation croisée des bassins versants en utilisant la variable débit observé, et le découpage administratif des cours d'eau en France. Les résultats ont montré que plus la distance augmente entre deux bassins versants, plus la corrélation entre eux est faible. La dernière étape a consisté à étudier la normalité des scores d'erreurs en utilisant les critères d'évaluation typiquement utilisés en hydrologie. Nous avons porté en particulier notre attention sur la droite de Henry et le test statistique Kolmogorov Smirnov, et nous avons trouvé que les performances des modèles ne suivent pas la loi normale. Ainsi, l'utilisation des tests statistiques non paramétriques pour la comparaison de la performance de modèles hydrologiques, n'est pas valide.

Ensuite, nous avons appliqué des tests statistiques pour étudier la différence entre les modèles sur des échantillons de bassins respectant leurs hypothèses, à savoir le test de permutation, le test de Mann-Whitney, le test de Friedman et le test de Cochran. Le test de permutation donne des résultats totalement différents des résultats des autres tests ; le test de Mann-Whitney qui est valable pour comparer des modèles deux à deux, donne une différence entre les modèles GR4J et GR5J, GR5J et GR6J et pas de différence entre GR4J et GR6J ; le test de Friedman qui normalement s'applique

seulement pour les bassins non corrélés, a montré des différences entre les modèles. Finalement, le test de Cochran a mis en évidence des différences entre les modèles pour certains critères mais pas tous.

Si ce stage ne permet pas de répondre à tous les cas de figures et toutes les questions que se posent les hydrologues, il permet de poser un premier cadre de travail qui servira aux futurs travaux d'Irstea.

Table des matières

Table des illustrations.....	1
Introduction	3
1. Le cadre de stage	5
1.1. Présentation de l'organisme d'accueil : IRSTEA – HBAN.....	5
1.1.1. L'IRSTEA	5
1.1.2. L'unité de recherche Hydrosystèmes et Bioprocédés.....	5
1.2. Contexte de stage.....	7
1.3. Objectifs de stage	7
1.4. Méthodologie.....	7
2. L'état de l'art	8
2.1. L'hydrologie des bassins versants	8
2.1.1. Le cycle de l'eau et le bilan hydrologique.....	8
2.1.2. Les précipitations.....	8
2.1.3. L'évapotranspiration potentielle.....	9
2.1.4. Qu'est-ce qu'un bassin versant ?.....	9
2.1.5. La mesure des débits	10
2.2. La modélisation hydrologique	11
2.2.1. Les modèles hydrologiques	11
2.2.2. Les modèles conceptuels	12
2.2.3. Les modèles GR	12
2.2.3.1. Le modèle GR4J	13
2.2.3.2. Les modèle GR5J et GR6J	15
2.2.4. Le calcul des simulations des débits.....	16
2.2.4.1. L'utilisation de modèle.....	16
2.2.4.2. Le calage du modèle.....	17
2.2.5. Les critères d'évaluation et de calage.....	17
2.3. Les tests statistiques	18
2.3.1. Le type de test, la statistique et la p-value.....	19
2.3.2. Les tests paramétriques.....	19
2.3.3. Les tests non paramétriques.....	21
2.3.4. Les tests statistiques dans l'équipe hydrologie des bassins versants	23
3. Les données et le travail réalisé	25
3.1. Les données	25
3.2. La chaîne de traitement réalisée	26
4. L'étude des variables	27

4.1. L'autocorrélation des bassins versants	28
4.2. La corrélation des bassins versants.....	32
4.3. La normalité des critères d'évaluation	37
5. Les résultats d'application des tests de significativité.....	40
Conclusion.....	43
Références bibliographiques	45

Table des illustrations

1. FIGURES

Figure 1 : Organigramme de l'Unité de Recherche Hydrosystèmes et Bioprocédés en avril 2016.....	6
Figure 2 : Cycle hydrologique.....	8
Figure 3 : Exemple de hétérogramme à la station de l'Ill à Altkirch (cumuls journaliers pour l'année 2000).	9
Figure 4 : Schéma d'un bassin versant.....	10
Figure 5 : Exemple d'hydrogramme à la station de l'Ill à Altkirch (moyennes journalières pour l'année 2000).	10
Figure 6 : Schéma synoptique du modèle GR4J (Perrin et al. 2003).....	13
Figure 7 : Schéma synoptique des modèles GR5J (Le Moine 2008) et GR6J (Pushpalatha 2013).	16
Figure 8 : Localisation des 605 bassins versants étudiés au sein du découpage en zones hydrographiques d'ordre 4.	26
Figure 9 : Schéma des étapes de travail durant le stage.....	27
Figure 10 : Autocorrélation du bassin de la Neste de Cap de Long à Aragnouet - 5,29km ²	29
Figure 11 : Autocorrélation du bassin de l'Abert-Benoit à Plabennec - 26,96km ²	29
Figure 12 : Autocorrélation du bassin du Strengbach à Ribeauvillé - 31,24km ²	30
Figure 13 : Autocorrélation du bassin de la Moselle à Toul - 3345,73 km ²	31
Figure 14 : Autocorrélation du bassin de l'Oise à Pont-Sainte-Maxence – 13483,53 km ²	31
Figure 15 : Autocorrélation du bassin La Saône à Mâcon – 26941,12 km ²	32
Figure 16 : Corrélation croisée entre deux bassins versants des zones hydrologiques A et B.	35
Figure 17 : Corrélation croisée entre deux bassins versants des zones hydrologiques A et X.	35
Figure 18 : Corrélation croisée entre deux bassins versants des zones hydrologiques H et D.	36
Figure 19 : Corrélation croisée entre deux bassins versants des zones hydrologiques H et Y.	36
Figure 20 : Q-Q plot des scores de KGE'.....	37
Figure 21 : Q-Q plot des scores de NSE.....	38
Figure 22 : Q-Q plot des scores de RMSE.....	38
Figure 23 : Q-Q plot des scores normalisés de KGE'.....	39
Figure 24 : Q-Q plot des scores normalisés de NSE.....	39
Figure 25 : Q-Q plot des scores normalisés de RMSE.....	40
Figure 26 : Distributions de critères de performance en fonction des modèles pour 605 bassins versants.	41
Figure 27 : Distributions de critères de performance en fonction des modèles pour 22 bassins versants.	42
Figure 28 : Démarche d'application d'un test.....	44

2. TABLEAUX

Tableau 1 : Deux tests statistiques paramétriques pour deux échantillons appariés.	20
Tableau 2 : Deux tests statistiques non paramétriques pour deux échantillons appariés.....	21
Tableau 3 : Trois tests statistiques non paramétriques pour $k > 2$ échantillons appariés.....	22
Tableau 4 : Différents cas d'utilisation des tests statistiques dans l'équipe HYDRO.....	24
Tableau 5 : la proportion de la corrélation égale à 0,7 des bassins versants.....	33
Tableau 6 : Interprétation des niveaux de la corrélation.....	34
Tableau 7 : Les p-values résultant des tests de normalité Kolmogorov-Smirnov et lillietest sur les critères d'évaluation des modèles hydrologiques.	38
Tableau 8 : Les p-values des tests statistiques de significativité pour 605 bassins versant (en gras, les résultats significatifs).....	41
Tableau 9 : Les p-values du test de Friedman pour 605 bassins versant.....	41
Tableau 10 : Les p-values des tests statistiques de significativité pour les 22 bassins non corrélés.....	42
Tableau 11 : Les p-values du test de Friedman pour les 22 bassins non corrélés.....	42

Introduction

Dans le cadre de ma première année en master Statistiques et Traitements des données à l'université Blaise Pascal à Clermont-Ferrand, j'ai effectué un stage de trois mois au sein de l'équipe « hydrologie des bassins versants » à IRSTEA (Institut national de Recherche en Sciences et Technologies pour l'Environnement et l'Agriculture) d'Antony, du 1^{er} avril au 30 juin 2016. Ce stage, effectué sous la direction de Guillaume Thirel et Olivier Delaigue, m'a permis de découvrir le domaine des statistiques appliquées à l'hydrologie.

L'amélioration des outils de modélisation hydrologique est une thématique de recherche très active. Cette amélioration s'évalue grâce au calcul de la performance des modèles pluie-débit, donc la capacité d'un modèle à reproduire des débits observés, qui s'estime à l'aide des critères de performances. Cependant, une amélioration de simulation des débits observés sur un bassin donné peut ne pas être reproductible sur un autre bassin, que ce soit à cause des caractéristiques hydrologiques et des conditions climatiques variable dans le temps, ou de mauvais choix de modélisation par exemple.

IRSTEA développe et améliore des modèles qui permettent, à partir des connaissances des pluies et températures passées, de simuler des débits en un point d'une rivière. Ces outils sont appliqués sur des bassins versants différents et sur des périodes différentes à des fins de généralisation des résultats. Ces modèles hydrologiques comportent un certain nombre de paramètres, qui doivent être déterminé par calage, donc on fait appel à une fonction dite fonction objectif qui est un critère numérique qu'on essaye d'optimiser, mesurant l'écart entre les observations et les simulations.

L'objectif de ce stage est de trouver des tests statistiques, qui sont des outils intéressants afin de renforcer la confiance que l'on peut avoir en des expérimentations, pour étudier la différence entre plusieurs ensembles des performances. Ces ensembles de performances représentant typiquement des critères de performances pour deux modèles différents, sur un ensemble de bassins versants.

Le présent mémoire présente donc l'environnement de stage, à savoir IRSTEA d'Antony, l'unité de recherche HBAN et l'équipe HYDRO. Ensuite, j'énonce le contexte, la problématique et les objectifs

du stage. Je détaille aussi l'état de l'art de la modélisation hydrologique et des tests statistiques. Puis, j'évoque les données et la chaîne de modélisation mise en place et l'étude des variables. Enfin, je présente et discute les résultats obtenus.

1. Le cadre de stage

1.1. Présentation de l'organisme d'accueil : IRSTEA – HBAN

1.1.1. L'IRSTEA

Créé en 1981, CEMAGREF (Centre national du machinisme agricole, du génie rural, des eaux et des forêts) est un établissement public à caractère scientifique et technologique français de recherche dans le domaine de l'environnement et de l'agriculture sous l'autorité des ministères chargés de la Recherche et de l'Agriculture. En 2012, il est renommé IRSTEA (Institut national de recherche en sciences et technologies pour l'environnement et l'agriculture).

L'objectif de l'IRSTEA est de produire des solutions concrètes au bénéfice de la décision publique et débouchant sur l'action. À ce titre, ce centre de recherche finalisée mène principalement une recherche partenariale à forte composante technologique et méthodologique et une recherche en appui aux politiques publiques (ministères, services déconcentrés de l'État, agences publiques)

Répartie au sein des neuf centres régionaux, la recherche de l'IRSTEA est organisée en trois grands départements : Eaux (ressource en eau, pollutions, écosystèmes aquatiques, risques naturels, irrigation, etc.), Écotecnologies (agriculture et agroalimentaire, gestion et stockage des déchets, épuration, etc.) et Territoires (gestion des relations entre milieu naturel et activités humaines).

1.1.2. L'unité de recherche Hydrosystèmes et Bioprocédés

L'unité de recherche Hydrosystèmes et Bioprocédés (HBAN) est basée au centre IRSTEA d'Antony. Ses travaux de recherche ont pour objectif la maîtrise qualitative et quantitative des eaux continentales de surface.

Les modèles et l'expertise de l'unité s'appliquent à :

- l'ingénierie des stations d'épuration et des installations de stockage traitement des déchets (stockage, méthanisation),
- la prévision des ressources en eau (fleuves, retenues) et des événements exceptionnels (crues, étiages),
- la gestion des populations de poissons et de l'occupation des bassins versants fortement anthropisés, permettant une approche intégrée de la gestion des hydrosystèmes.

L'unité HBAN travaille en collaboration avec plusieurs établissements publics (ANR, EDF Météo-France, ONEMA, IRD, CNRS, etc.) et privés (Suez, Veolia, etc.), et des associations scientifiques. Elle est composée de 8 équipes de recherche dont l'équipe Hydrologie des bassins versants (HYDRO), à laquelle j'étais intégrée.

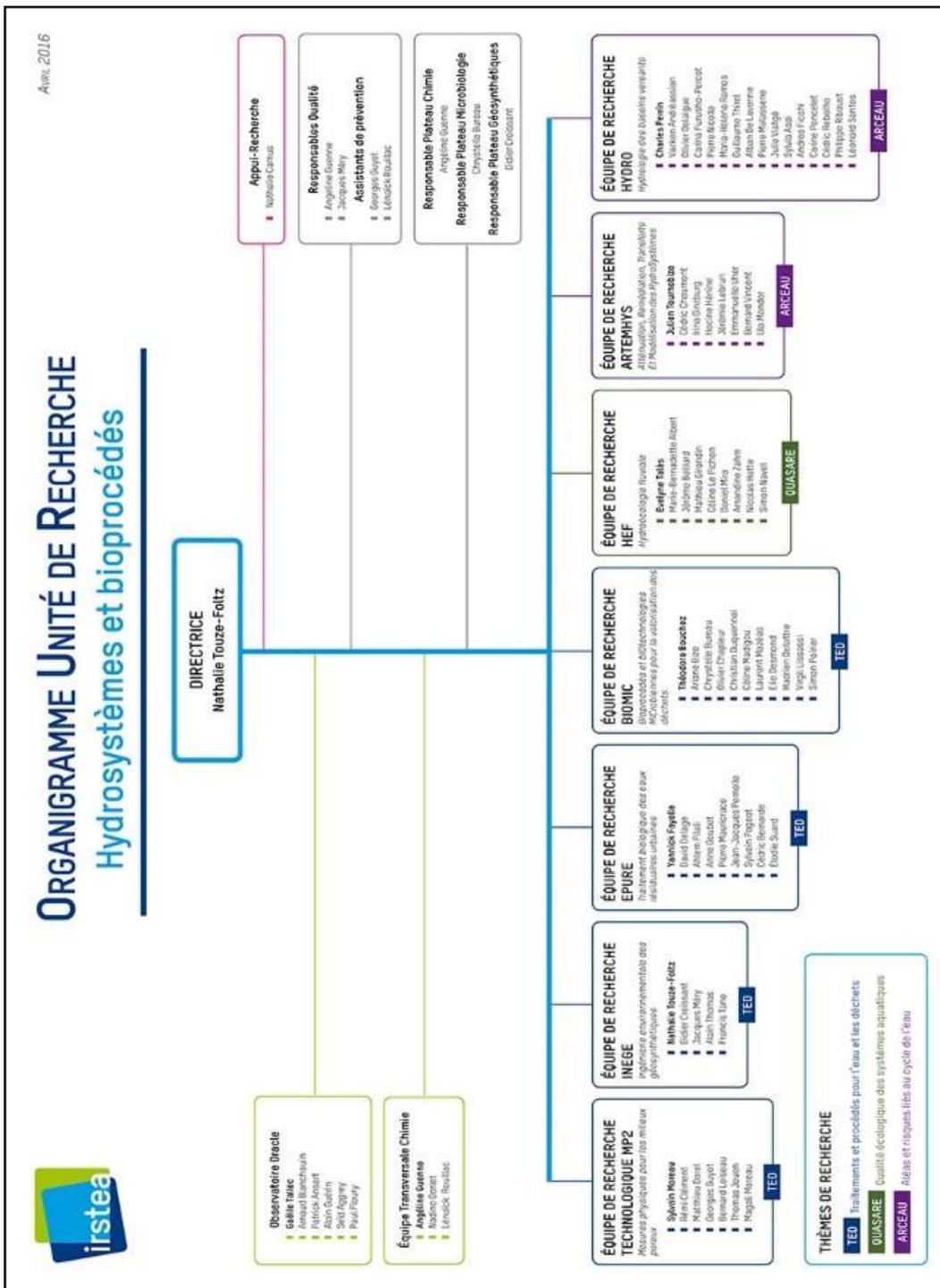


Figure 1 : Organigramme de l'Unité de Recherche Hydrosystèmes et Bioprocédés en avril 2016.

L'équipe Hydrologie des bassins versants (HYDRO) développe des travaux de recherche visant à améliorer les connaissances des comportements hydrologiques des bassins versants et se focalise sur la production d'outils de modélisation applicables pour des objectifs opérationnels (prévision, gestion, etc.). Les modèles développés dans ce cadre sont utiles dans plusieurs contextes, par exemple les prévisions des crues ou des étiages, les études d'impact de changements climatiques sur l'hydrologie et la simulation de débits.

Le site web¹ de l'équipe présente l'ensemble de recherches réalisées par l'équipe HYDRO.

1.2. Contexte de stage

Dans un objectif d'amélioration des outils de modélisation hydrologique, il convient d'évaluer la capacité d'un modèle à simuler au mieux les débits, c'est-à-dire à reproduire les débits observés, avant et après modification de sa structure, et ce, à l'aide de critères de performances afin de pouvoir juger de l'amélioration. Cependant, il est important de s'assurer de la généralité de l'amélioration des performances d'un modèle. En effet, il est tout à fait possible qu'une amélioration de simulations de débits observés sur un bassin versant donné ne soit pas reproductible sur un autre bassin versant, des caractéristiques hydrologiques pouvant différer entre ceux-ci. De même, une amélioration notée sur une période donnée peut ne pas être reproduite sur une autre période, les conditions climatiques étant variables dans le temps.

Pour s'assurer de la bonification des simulations d'un outil, le modélisateur doit multiplier les tests sur des bassins versants différents et sur des périodes différentes. Au lieu de comparer deux performances (un seul et unique bassin avec deux structures de modèles différentes), il est alors nécessaire de comparer deux ensembles de performances. Comme il est rare que l'évolution de tous les éléments de ces ensembles soit dans le sens d'une amélioration, ou dans le sens d'une détérioration (i.e. certains bassins verront leur performance améliorée, d'autres non), il est nécessaire de disposer d'outils pour déterminer une amélioration « moyenne » significative.

Dans ce contexte, les tests statistiques représentent des outils intéressants pour une étude objective des différences entre plusieurs ensembles de performances.

1.3. Objectifs de stage

Le stage vise, d'une part, à trouver des tests statistiques adaptés aux questions de recherche de l'équipe hydrologie des bassins versants au centre IRSTEA d'Antony, et d'autre part à identifier les tests statistiques déjà utilisés dans l'équipe.

1.4. Méthodologie

Dans le but de mieux gérer le stage, et d'atteindre ses objectifs, nous avons adopté une méthodologie de travail dont la démarche est la suivante :

- Réaliser une synthèse bibliographique sur l'hydrologie, les modèles hydrologiques, les critères d'évaluation des modèles et les tests statistiques,
- Discuter avec les doctorants et chercheurs de l'équipe concernés par le sujet,
- Identifier des tests statistiques déjà utilisés,
- Critiquer l'adéquation de ces tests et proposer d'autres plus adaptées,
- Décrire la mise en œuvre des tests pour les différents aspects des travaux de l'équipe,
- Étudier la base de données des bassins versants à utiliser,

¹ <http://webgr.irstea.fr/>

- Implémenter dans un langage de programmation libre R les tests statistiques sur un certain nombre de bassins.

2. L'état de l'art

2.1. L'hydrologie des bassins versants

2.1.1. Le cycle de l'eau et le bilan hydrologique

L'hydrologie est la science qui étudie la distribution et la circulation de l'eau à la surface de la terre, dans les sols, dans les souterrains et dans l'atmosphère. L'eau est indispensable à la vie des êtres vivants. Sur la planète, elle se retrouve sous trois états, appelés « les phases de l'eau » : liquide, solide ou gazeux. Le changement de phase de l'eau dépend de la température et de la pression.

L'ensemble des processus de transformation et de transfert de l'eau forme le cycle hydrologique (Figure 2), qui est un ensemble des phénomènes naturels, composé par : les précipitations, l'évapotranspiration, l'interception, le ruissellement de surface, l'infiltration, la percolation et le stockage dans les dépressions.

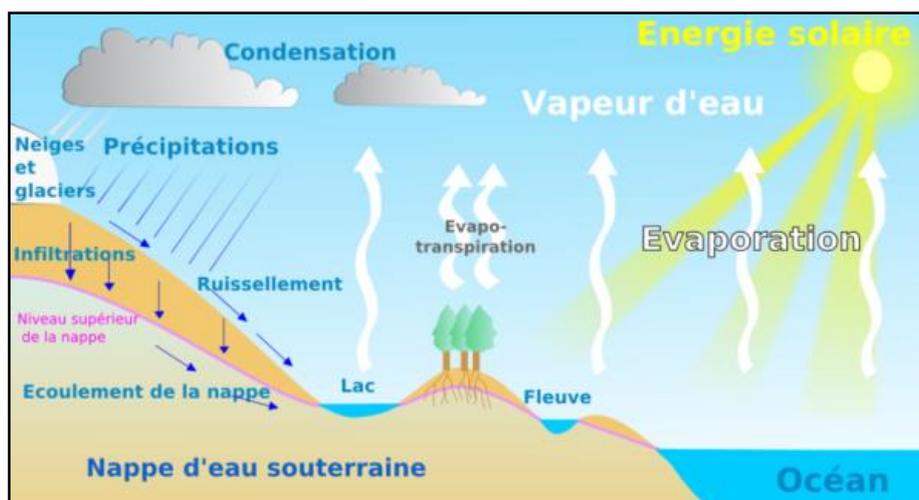


Figure 2 : Cycle hydrologique².

Les données nécessaires à l'utilisation des modèles hydrologiques utilisés dans le cadre de ce stage sont les données de précipitations et d'évapotranspiration.

2.1.2. Les précipitations

Les précipitations sont une des composantes principales du cycle hydrologique. Elles représentent toutes les eaux météoriques qui tombent sur la surface de la terre, sous forme liquide ou solide en provenance directe ou indirecte de la condensation de la vapeur d'eau atmosphérique. Elles sont mesurées en millimètres le plus souvent grâce à des pluviomètres. Il existe plusieurs formes différentes de précipitations : la pluie, la pluie verglaçante, le grésil et la neige, qui dépendent de

² <http://www.futura-sciences.com/magazines/environnement/>

l'épaisseur des nuages, de leur taux d'humidité, de la température de l'air ambiant et de celle du sol, mais aussi de la pression atmosphérique.

En hydrologie, on représente classiquement l'intensité des précipitations en fonction du temps sous la forme d'un histogramme, appelé « hyétogramme » (Figure 3).

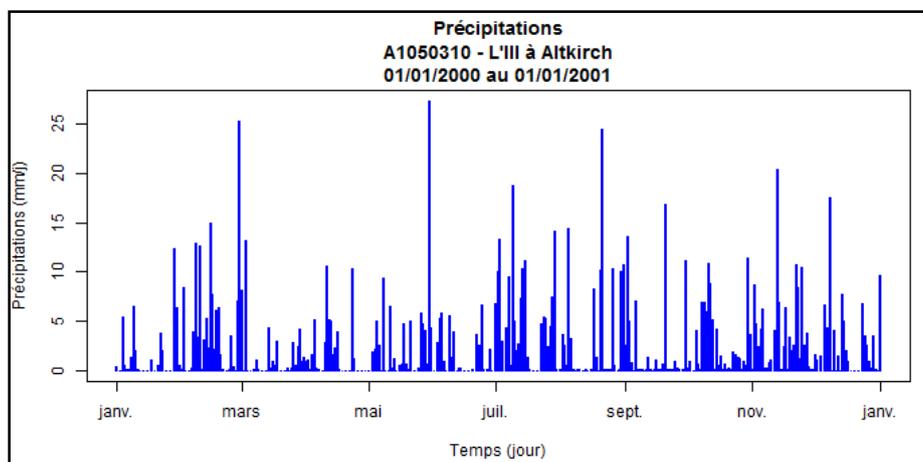


Figure 3 : Exemple de hyétogramme à la station de l'III à Altkirch (cumuls journaliers pour l'année 2000).

2.1.3. L'évapotranspiration potentielle

L'évapotranspiration est la combinaison de l'évaporation directe à partir des surfaces d'eau libre, et de la transpiration végétale qui est un processus par lequel l'eau des végétaux est transférée dans l'atmosphère sous forme de vapeur. Elle dépend de cinq facteurs : la température, la pression atmosphérique, l'humidité, le vent, et la surface d'eau à l'air libre.

L'évapotranspiration réelle étant très difficilement mesurable, on utilise classiquement une approximation de cette dernière appelée évapotranspiration potentielle (ETP), qui correspond à la quantité maximale d'eau susceptible d'être perdue. Elle est mesurée en millimètre.

2.1.4. Qu'est-ce qu'un bassin versant ?

Le bassin versant est l'échelle géographique adoptée pour gérer l'eau (Figure 4). Il est délimité par les crêtes (ligne de points hauts d'un relief séparant deux versants opposés) au-delà desquelles les pentes dirigent les précipitations tombées vers un autre cours d'eau. Ces frontières sont appelées « lignes de partage des eaux ». Autrement dit, le bassin versant représente la zone irriguée par un réseau hydrographique dont toutes les précipitations reçues suivent une pente naturelle et se concentrent vers ce que l'on appelle un exutoire.

Il est possible d'estimer les volumes et les distributions temporelles de ces précipitations à l'exutoire. Bien que le débit puisse techniquement être mesuré, les moyens à mettre en œuvre sont lourds, et son estimation par des moyens mathématiques est donc parfois rendue nécessaire. Par ailleurs, ces outils mathématiques sont indispensables à des fins d'anticipation des débits. Ce sont des objectifs de la modélisation hydrologique.

Le bassin versant d'un grand cours d'eau se subdivise en un certain nombre de sous-bassins versants de ses affluents.

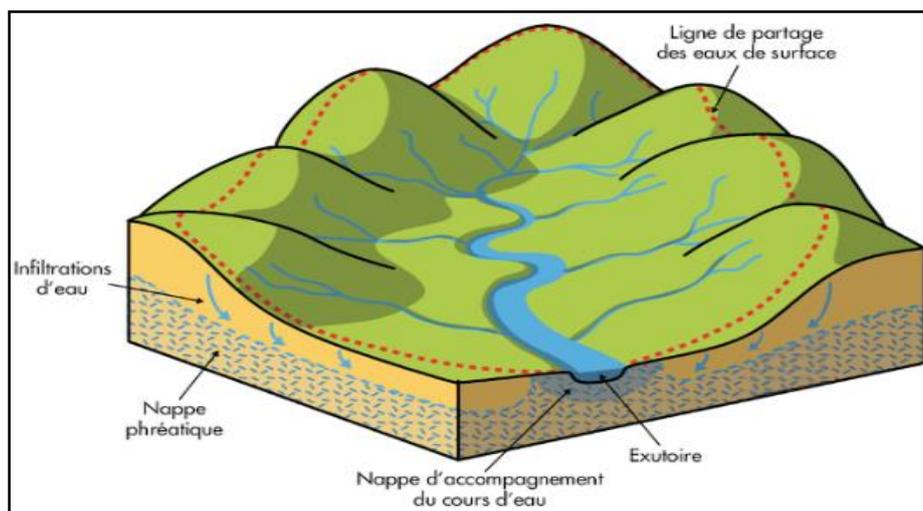


Figure 4 : Schéma d'un bassin versant.

2.1.5. La mesure des débits

Le débit Q d'un cours d'eau, exprimé en m^3/s ou l/s , correspond au volume d'eau liquide traversant une section transversale de l'écoulement par unité de temps. Le débit reflète physiquement le comportement du bassin versant, et peut être interprété dans le temps et dans l'espace. À la station hydrométrique, on ne dispose pas d'une mesure directe et continue des débits, mais seulement d'un enregistrement des hauteurs d'eau en une section donnée. Le passage de la courbe des hauteurs d'eau en fonction du temps $H = f(t)$ (appelée « limnigramme ») à celle des débits $Q = f(t)$ appelée « hydrogramme » (Figure 5) se fait par l'établissement d'une courbe de tarage $Q = f(H)$ à l'aide d'un jaugeage.

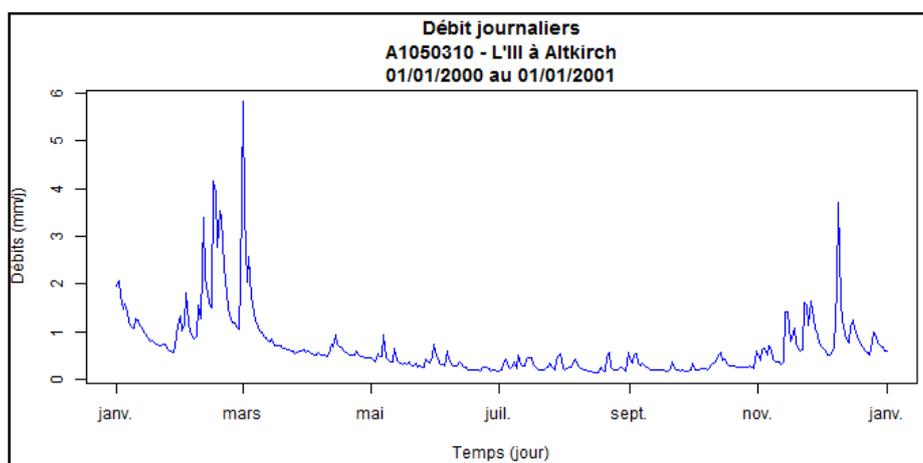


Figure 5 : Exemple d'hydrogramme à la station de l'III à Altkirch (moyennes journalières pour l'année 2000).

La mesure des hauteurs d'eau :

Anciennement, la mesure des hauteurs d'eau ou la limnimétrie s'effectuait de manière discontinue par la lecture d'une règle graduée (échelle limnimétrique) fixée sur un support. Pour connaître en continu les variations d'un plan d'eau, il est désormais courant d'utiliser des limnigraphes, qui fournissent sur un support un enregistrement continu des variations du niveau d'eau dans la rivière en fonction du temps.

La mesure des débits :

Pour mesurer le débit d'un écoulement naturel, il existe quatre méthodes :

- Les méthodes volumétriques (appelés « jaugeages capacitifs ») permettent de déterminer le débit directement à partir du temps nécessaire pour remplir d'eau un récipient d'une contenance déterminée. Compte tenu des aspects pratiques inhérents à la méthode de mesure (taille du récipient nécessaire, incertitude sur la mesure du temps, aménagement spécifique éventuel), cette méthode n'est généralement pratiquée que pour des débits très faibles, de quelques l/s.
- Les méthodes d'exploration du champ de vitesse consistent à déterminer la vitesse de l'écoulement en différents points de la section, tout en mesurant la surface de la section mouillée. Ces techniques nécessitent un matériel spécifique et un personnel formé à son utilisation.
- Les méthodes hydrauliques tiennent compte des forces qui régissent l'écoulement (pesanteur, inertie, viscosité...). Ces méthodes obéissent aux lois de l'hydraulique.
- Les méthodes physico-chimiques prennent en compte les variations, lors de l'écoulement, de certaines propriétés physiques du liquide (concentration en certains éléments dissous). Ces méthodes consistent généralement à injecter dans le cours d'eau un corps en solution, et à suivre l'évolution de sa concentration au cours du temps. Ce sont les méthodes dites « par dilution » ou encore « chimique ».

2.2. La modélisation hydrologique

2.2.1. Les modèles hydrologiques

Le modèle pluie-débit est un outil mathématique qui intègre les processus hydrologiques à l'échelle du bassin versant pour déterminer la réponse de celui-ci. De nombreuses structures, classifications et définitions des modèles pluie-débit ont été proposées. En ce qui concerne la représentation des processus, deux principales approches se distinguent (Sivapalan et al. 2003). La première est basée sur des fondements physiques, c'est-à-dire qu'elle utilise par exemple des équations issues de la mécanique des fluides. La seconde est appelée conceptuelle, dans le sens où elle cherche à représenter les principaux processus de la relation pluie-débit sans utiliser les lois physiques. Par ailleurs, des modèles dits « boîtes noires » ou statistiques peuvent aussi parfois être utilisés.

Les modèles utilisés dans ce stage étant conceptuels, nous ne nous attarderons par la suite que sur ceux-ci. Toutefois, ce travail est parfaitement adaptable à des modèles physiques ou statistiques.

2.2.2. Les modèles conceptuels

Dans ces modèles, les processus hydrologiques sont conceptualisés à partir d'une interprétation des observations à l'échelle du bassin versant. On essaie ici de représenter uniquement les principaux processus menant à la formation des débits. Ces modèles sont souvent dits « globaux », car ils utilisent des structures qui ne tiennent pas compte explicitement de la variabilité spatiale des données d'entrée météorologiques et des caractéristiques physiques des bassins versants. D'après Kampf et al. (2007), le modèle conceptuel global peut être défini comme un modèle à une dimension qui intègre les processus hydrologiques sur les trois dimensions spatiales du bassin versant pour calculer un débit à son exutoire. La famille des modèles GR³, développés au centre IRSTEA d'Antony depuis plus de 30 ans, et utilisés dans ce stage, appartiennent à cette catégorie de modèles.

2.2.3. Les modèles GR

Les modèles hydrologiques pluie-débit de type conceptuel, appliqués à l'échelle des bassins versants, permettent de simuler les débits à partir des précipitations. Ils sont relativement aisés d'utilisation, car ils ne requièrent que très peu de données d'entrées, à savoir : les précipitations, l'évapotranspiration potentielle et éventuellement la température pour les bassins fortement enneigés. Ces modèles sont le plus souvent parcimonieux (Anctil et al. 2003, Perrin et al. 2003) et permettent de simuler de manière performante le fonctionnement hydrologique des bassins versants avec une complexité limitée (évapotranspiration, ruissellement, pertes souterraines), sans nécessiter une connaissance fine du fonctionnement physique du bassin.

Depuis une trentaine d'années, l'équipe Hydrologie des bassins versants d'IRSTEA⁴ a développé toute une famille de modèles hydrologiques connus sous le nom de « modèles GR » (pour « génie rural »), permettant de prédire au mieux les débits des cours d'eau, et ce sur une très grande variété de bassins versants, tout en ne requérant que très peu de données d'entrée, données qui, par ailleurs, sont aisément disponibles. Ces modèles permettent de travailler à différents pas de temps (de l'heure à l'année), mais aussi dans des contextes d'utilisations variés comme par exemple : l'évaluation de la ressource en eau, la prévision des crues et les projections hydrologiques en climat futur. L'utilisation de ces modèles est déjà relativement répandue tant dans le monde de la recherche académique (ENPC, universités Paris 6, Nice-Sophia Antipolis, Grenoble Alpes) que dans les services opérationnels (Services de prévision des crues [S.P.C.] en France, Bureau of Meteorology [B.O.M.] en Australie) ou dans le monde privé (bureaux d'études, EDF). Ces modèles ont pour avantage d'avoir un faible nombre de paramètres à optimiser, ce qui évite les problèmes liés à la sur-paramétrisation et réduit considérablement les temps de calcul.

Dans le cadre de ce stage, trois modèles ont été utilisés, à savoir le modèle hydrologique GR4J, et ses versions dérivées que sont GR5J, et GR6J. Tous fonctionnent à des pas de temps journaliers. Ces

³ <http://webgr.irstea.fr/modeles/bref-historique/>

⁴ <http://webgr.irstea.fr/>

modèles sont peu exigeants en données puisqu'ils n'utilisent que des données de pluie et de l'ETP en entrée.

2.2.3.1. Le modèle GR4J

Le modèle GR4J est un modèle journalier à 4 paramètres (Perrin et al. 2003). Les paramètres à caler sont les suivants :

- X_1 : capacité maximum du réservoir de production,
- X_2 : coefficient d'échanges souterrains,
- X_3 : capacité maximum du réservoir de routage,
- X_4 : temps de base de l'hydrogramme unitaire HU1.

Son développement a été initié à l'IRSTEA au début des années 1980, avec pour objectif de mettre au point un modèle de simulation pluie-débit fiable en vue d'utilisations pour des applications de gestion de la ressource en eau. Sa structure associe un réservoir de production et un réservoir de routage, des hydrogrammes unitaires ainsi qu'une fonction d'ouverture sur l'extérieur non atmosphérique permettant de simuler le gros du comportement hydrologique du bassin. La Figure 6 schématise le fonctionnement du modèle GR4J.

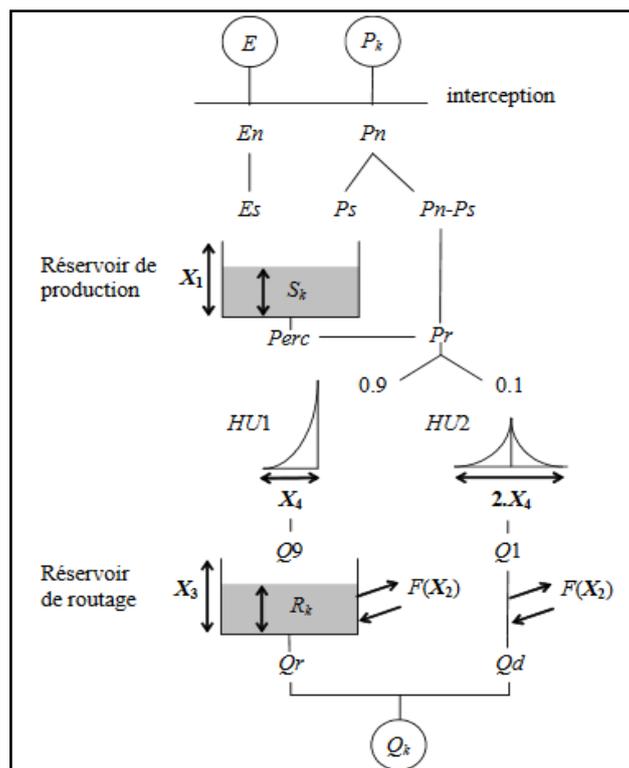


Figure 6 : Schéma synoptique du modèle GR4J (Perrin et al. 2003).

La production évalue la partie des précipitations qui alimentent les eaux de ruissellement à l'exutoire du bassin versant, ainsi que la partie qui s'évapore. Le routage a un rôle de reconstitution de la réponse hydrologique du bassin versant à une impulsion de pluie ; il modélise conceptuellement l'écoulement jusqu'à la rivière.

La production :

P_k : La pluie journalière du jour k,

E : L'évapotranspiration moyenne pour le jour k,

P_n : La pluie nette,

E_n : L'évapotranspiration nette.

L'interception représente le processus par lequel une partie des précipitations est captée et retenue par la végétation, puis évaporée sans avoir atteint la surface du sol.

Lors de la phase de neutralisation, une quantité de l'ETP est produite de la pluie brute. P_n représente la pluie nette, et E_n l'ETP nette.

Si $P_k \geq E$ alors, $P_n = P_k - E$ et $E_n = 0$, et

Si $P_k \leq E$ alors, $P_n = 0$ et $E_n = E - P_k$

Dans le cas où P_n est différente de 0, une partie P_s de P_n alimente le réservoir de production et est calculée par :

$$P_s = \frac{X_1 * \left(1 - \left(\frac{S_k}{X_1}\right)^2\right) * \tanh \frac{P_n}{X_1}}{1 + \frac{S_k}{X_1} * \tanh \frac{P_n}{X_1}} \quad (1)$$

Où S_k est le niveau du réservoir de la production.

Et si E_n est différent de 0, une quantité d'évaporation E_s est retirée du réservoir de production :

$$E_s = \frac{S_k * \left(2 - \frac{S_k}{X_1}\right) * \tanh \frac{E_n}{X_1}}{1 + \left(1 - \frac{S_k}{X_1}\right) * \tanh \frac{E_n}{X_1}} \quad (2)$$

Alors, le S_k est mis à jour, et la capacité maximale est déterminée par un premier paramètre libre X_1 exprimé en mm. S_k est calculé par :

$$S_k = S_k - E_s + P_s \quad (3)$$

La percolation issue du réservoir de production S_k est alors calculée :

$$\text{perc} = S_k * \left(1 - \left(1 + \left(\frac{4}{9} * \frac{S_k}{X_1}\right)\right)^{\frac{-1}{4}}\right) \quad (4)$$

De nouveau, S_k devient :

$$S_k = S_k - \text{perc} \quad (5)$$

La vidange peut se faire par deux chemins distincts l'évapotranspiration et la percolation.

Et on obtient une quantité d'eau P_r entrante dans la partie routage du modèle.

$$P_r = \text{perc} + (P_n - P_s) \quad (6)$$

Le routage :

L'hydrogramme unitaire, dont l'allure dépendra du paramètre X_4 , assimilable au délai de restitution du bassin versant, est la première étape qui distribue temporellement les volumes d'eau venant du réservoir de production. Plus le paramètre est grand, plus l'hydrogramme s'étendra dans le temps et donc plus la réponse pluvieuse, restituée sous forme de débit, sera lissée dans le temps.

Le réservoir de routage cherche à modéliser la capacité d'infiltration de l'eau dans le sol. Le volume d'eau restitué à la sortie de l'hydrogramme UH1 viendra alimenter le réservoir de routage, dont la capacité maximum est déterminée par le paramètre X_3 en mm. UH2 gère le transfert de 10 % d'eau restant vers l'exutoire.

L'eau transférée, soit vers le réservoir de routage, soit directement à l'exutoire, est également soumise à une fonction d'échange avec le sous-sol, grâce à un coefficient d'échange exprimé en mm.

La fonction d'échange joue plusieurs rôles à la fois. Elle est à la base censée modéliser les échanges souterrain, mais a également un rôle de correction des volumes d'un pas de temps à l'autre ou bien prend en compte l'évapotranspiration au niveau des sols. La fonction d'échange est en fonction du paramètre X_2 , dit paramètre d'échange mais également du paramètre X_3 , capacité maximal du réservoir de routage.

Pour le flux passant direct à l'exutoire Q_d est donné par :

$$Q_d = \max(0, Q_1 + F) \quad (7)$$

Où F est la fonction d'échange. Sa formule est :

$$F(x_2) = x_2 \left(\frac{R_k}{X_3} \right)^{\frac{7}{2}} \quad (8)$$

Pour le flux $Q_r < R_k$ en sortie du réservoir,

$$Q_r = R_k \left(1 - \left(1 + \left(\frac{R_k}{X_3} \right)^4 \right)^{\frac{-1}{4}} \right) \quad (9)$$

Nous obtenons après un nouveau niveau dans le réservoir :

$$R_k = R_k - Q_r \quad (10)$$

Enfin, le débit en sortie de la simulation sera la somme des débits obtenus Q_d et Q_r .

$$Q = Q_d + Q_r \quad (11)$$

2.2.3.2. Les modèles GR5J et GR6J

Les modèles GR5J (Le Moine 2008) et GR6J (Pushpalatha 2013) sont des évolutions du modèle GR4J. Ils sont donc basés sur la même structure, à savoir une association des deux réservoirs de production et de routage, de deux hydrogrammes unitaires et une fonction d'échange souterrain. Il existe des différences au niveau de la fonction d'échange. Pour le modèle GR5J, un paramètre libre adimensionnel X_5 est ajouté à cette fonction d'échange ; ce paramètre indique le sens de l'échange avec

la nappe et dépend du niveau R du réservoir de routage. Le paramètre ajouté au modèle GR6J est lié à l'ajout d'un réservoir exponentiel R_2 ; il s'agit de sa capacité en mm.

X_5 : fonction d'échange en eau améliorée.

X_6 : coefficient de vidange exponentielle du second réservoir de routage.

Le modèle GR5J, grâce à son paramètre supplémentaire, prend mieux en compte les effets d'échange d'eau avec la nappe par rapport au modèle GR4J, tandis que le modèle GR6J réalise de meilleures simulations en étiage grâce à son réservoir exponentiel, que GR4J.

La Figure 7 présente les deux modèles, à gauche GR5J, et à droite GR6J.

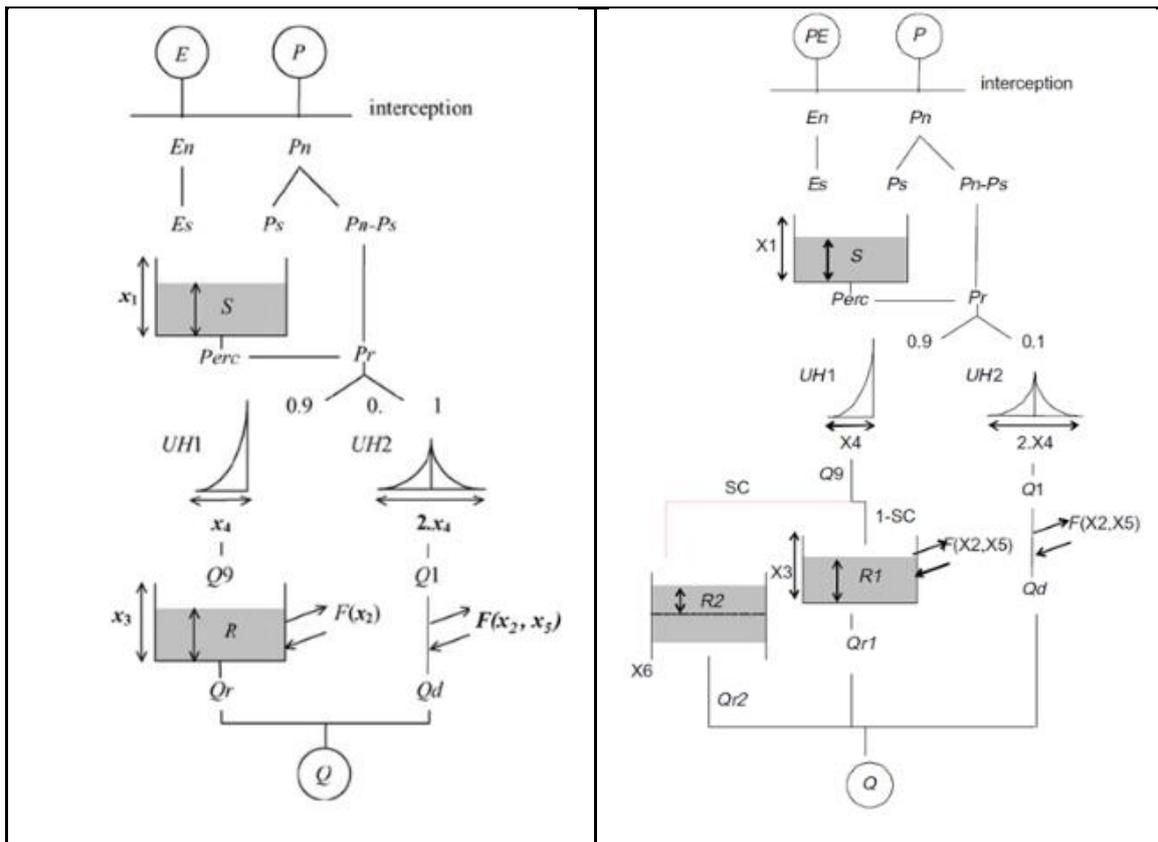


Figure 7 : Schéma synoptique des modèles GR5J (Le Moine 2008) et GR6J (Pushpalatha 2013).

2.2.4. Le calcul des simulations des débits

2.2.4.1. L'utilisation de modèle

Le modèle hydrologique adopté doit être calé afin de s'adapter aux propriétés du bassin versant étudié. En effet, la simplification de la réalité implique que certains, voire la totalité des paramètres utilisés pour décrire le modèle, ne peuvent être reliés directement à des mesures de terrain, d'où la nécessité d'un calage mathématique qui s'effectue sur une période où la pluie, le débit et l'évapotranspiration potentielle sont connus. Ainsi, nous pouvons comparer les débits simulés avec les débits observés grâce à des critères appelés fonctions-objectif.

2.2.4.2. Le calage du modèle

Dans cette partie, il y a deux étapes : la première appelée calage et la seconde appelée contrôle. Lors du calage, on fait varier les paramètres du modèle et l'on utilise un score de performance pour quantifier l'écart entre les débits simulés et observés et ainsi apprécier la qualité de simulation du modèle. Lors du contrôle, on simule les débits sur une période différente de celle utilisée pour le calage en utilisant le jeu de paramètres donnant le critère optimal trouvé lors de l'étape de calage. Les débits simulés lors du contrôle sont alors comparés à l'observation à l'aide d'un critère d'évaluation. Les erreurs calculées donnent une estimation des performances du modèle.

2.2.5. Les critères d'évaluation et de calage

Plusieurs critères sont couramment utilisés pour estimer la qualité des simulations des débits Q_{sim} par rapport aux débits observés Q_{obs} . Ces critères sont nommés « fonction objectif ».

Les critères quadratiques basés sur les moindres carrés sont ceux qui sont le plus traditionnellement employés dans le domaine de l'hydrologie. L'erreur quadratique s'écrit :

$$Eq = \sum_{t=1}^n (Q_{obs}(t) - Q_{sim}(t))^2 \quad (12)$$

Avec :

- $Q_{sim}(t)$ est le débit simulé à l'instant t ,
- $Q_{obs}(t)$ est le débit observé à l'instant t ,
- n le nombre totale de pas de temps.

Un problème de ce critère vient du fait que l'erreur commise est du même ordre de grandeur que le débit moyen du bassin, ce qui rend la comparaison avec d'autres bassins versants difficile. Nash et al. (1970) ont donc proposé de normaliser ce critère par la variance des débits observés. Le NSE est considéré le critère historiquement le plus utilisé en hydrologie pour étudier la performance des modèles pluie-débit. Il prend des valeurs entre $-\infty$ et 1.

$$NSE = 1 - \frac{\sum_{t=1}^n (Q_{obs}(t) - Q_{sim}(t))^2}{\sum_{t=1}^n (Q_{obs}(t) - \overline{Q_{obs}(t)})^2} \quad (13)$$

Tel que :

- Si le $NSE \leq 0$: le modèle n'est pas meilleur que la moyenne des débits observés,
- Si le $NSE > 0$: le modèle est meilleur que la moyenne des débits observés,
- Si le $NSE = 1$: le modèle correspond parfaitement aux débits observés.

Si le NSE est positif, le modèle donne de meilleurs résultats qu'un modèle naïf donnant à chaque pas de temps un débit constant égal à la moyenne des débits observés, et s'il est négatif, c'est l'inverse.

En 2009, Gupta et al. (2009) ont proposé une modification du critère NSE qui améliore la performance des modèles pluie-débit en simulation, ce nouveau critère est nommé KGE. Il est basé sur le biais, le rapport des variances, et le coefficient de corrélation entre débits observés et cumulés.

$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (14)$$

Avec :

- $\alpha = \frac{\sqrt{(Q_{sim}(t) - \overline{Q_{sim}(t)})^2}}{\sqrt{(Q_{obs}(t) - \overline{Q_{obs}(t)})^2}}$ présente le rapport des écarts-types,
- $\beta = \frac{\overline{Q_{sim}}}{\overline{Q_{obs}}}$ est le biais,
- r est la corrélation entre Q_{obs} et Q_{sim} , donnée par :

$$r = \frac{1}{n} \sum \frac{(Q_{obs}(t) - \overline{Q_{obs}(t)})(Q_{sim}(t) - \overline{Q_{sim}(t)})}{\sigma_{obs}\sigma_{sim}} \quad (15)$$

Où :

- σ_{obs} est la variance des débits observés,
- σ_{sim} est la variance des débits simulés.

Nous citons aussi la RMSE, *Root Mean Square Error* ou la racine carrée de l'erreur quadratique moyenne. Il est rarement utilisé. Il permet d'évaluer la qualité moyenne des prévisions et ainsi d'avoir une vue d'ensemble de la qualité des modèles.

Il peut être calculé sur les débits. Il est défini par :

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Q_{obs}(t) - Q_{sim}(t))^2}{n}} \quad (16)$$

2.3. Les tests statistiques

Dans le but d'étudier la différence significative entre des modèles hydrologiques, l'utilisation des tests statistiques semble intéressante. En effet, comme détaillé en introduction, l'évaluation de modèles hydrologiques passe par le calcul de critères de performances sur un nombre parfois élevé de bassins versants. On se retrouve alors avec un nombre élevé de critères de performances, pour deux ou plus modèles, qu'il convient de comparer. Si la comparaison simple de la moyenne des critères pour chacun des modèles, ou bien la comparaison de certains quantiles, peuvent donner une indication de la hiérarchie entre les modèles, il est parfois difficile de conclure. En effet, ces valeurs peuvent être influencées par des valeurs extrêmes ou bien ne pas représenter de manière adéquate ce que l'on veut évaluer. Il est alors possiblement intéressant de mettre en place un cadre statistique pour appuyer ou non les conclusions des études hydrologiques, avec des tests statistiques.

Chaque test est une démarche permettant d'évaluer la validité d'une hypothèse statistique, rejeter ou ne pas rejeter l'hypothèse, en fonction d'un échantillon d'observations que nous avons à tester.

Lorsque nous définissons la nature de la population et d'observations, nous établissons un modèle statistique, c'est-à-dire une formulation mathématique des hypothèses faites sur les observations. À chaque test statistique est associé un modèle et des contraintes de mesure. Ce test n'est alors valide que si les conditions imposées par le modèle et les contraintes de mesure sont respectées.

Dans notre cas, la modélisation hydrologique, les individus statistiques sont des valeurs de critères pour chacun des bassins versants de l'échantillon, les populations correspondent à des valeurs de critère pour des modèles hydrologiques différents, et les échantillons dans chaque population sont un tirage identique parmi les populations, donc il s'agit d'échantillons appariés (c'est le cas où nous disposons de deux séries de mesures ou plus portant sur le même échantillon des individus qui sont dépendants entre eux, c'est-à-dire une intensité de liaison existe entre les individus).

Par la suite, le cadre d'étude concernera l'évaluation de la différence significative entre les performances des trois modèles hydrologiques GR4J, GR5J et GR6J. Ces évaluations consistent à évaluer les erreurs commises obtenues par les critères KGE', NSE et RMSE. Nous appliquerons quelques tests statistiques de significativité pour évaluer, pour chaque critère, la différence de performance des trois modèles hydrologiques. La vérification de l'applicabilité des différents tests selon les cas considérés sera au cœur de ce travail.

Alors quels sont les tests statistiques qui peuvent être appliqués ?

Il existe deux types de tests : des tests paramétriques, qui nécessitent de vérifier la normalité et l'homogénéité des variances, et des tests non paramétriques qui utilisent les rangs et qui ne nécessitent pas l'estimation de la moyenne et de la variance et qui sont plus puissants que les tests paramétriques dans le cas où la distribution n'est pas normale.

2.3.1. Le type de test, la statistique et la p-value

Les tests statistiques peuvent être soit bilatéraux « two-sided », soit unilatéraux « one-sided », en fonction de type de conclusions que nous cherchons à faire et de l'hypothèse que nous posons au départ. Les tests unilatéraux servent à montrer qu'un traitement est supérieur ou inférieur à un autre traitement de référence et qui correspond à faire l'hypothèse que le traitement doit devenir meilleur, et dans ce cas mieux utiliser un test unilatéral qu'un test bilatéral. Les tests bilatéraux sont utilisés dans le cas où nous ne pouvons pas spécifier de direction particulière pour l'hypothèse alternative. Il importe peu que le paramètre soit plus grand ou plus petit, ce qui compte, c'est qu'il diffère de la valeur supposée en hypothèse. En hydrologie, par exemple si nous faisons une modification sur un modèle pour l'améliorer, nous pouvons faire l'hypothèse qu'il doit devenir meilleur et donc nous utilisons un test unilatéral.

La statistique de test est la valeur numérique que nous obtenons et que nous pouvons comparer au seuil choisi (5 %, 0.1 %...) dans la table du test en question. Sous tous les logiciels de statistique, et particulièrement R, la p-value évite de comparer la statistique de test à la valeur théorique de la table (par rapport au seuil) et donne directement la conclusion de rejeter ou non.

2.3.2. Les tests paramétriques

Dans cette partie, nous citons deux tests statistiques (Tableau 1) appliqués sur deux échantillons d'observations appariées.

Tableau 1 : Deux tests statistiques paramétriques pour deux échantillons appariés.

Tests	Nature des individus	Hypothèses sur les observations	Pourquoi ?	Commande R
Test F	Dépendants	Normalité	Comparaison des variances	<code>var.test(paired=TRUE)</code>
Test t de student	Indépendants	Normalité et homoscedasticité	Comparaison des moyennes	<code>t.test(paired=TRUE, alternative=" ")</code>

Le test F :

Le test F est un test paramétrique. Il sert à étudier la différence entre deux échantillons quand il existe une corrélation entre les paires d'observations $(x_1, y_1), (x_2, y_2), \dots (x_n, y_n)$. Soit n la taille des observations et A et B les deux échantillons associés. Le rapport de la variance est calculé par :

$$F = \frac{S_1^2}{S_2^2} \quad (17)$$

Tel que :

$$s_1^2 = \frac{\sum(x_i - \bar{x})^2}{n} \text{ et } s_2^2 = \frac{\sum(y_i - \bar{y})^2}{n} \quad (18)$$

La corrélation r est donnée par :

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{(\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2)^{\frac{1}{2}}} \quad (19)$$

Enfin la statistique du test est :

$$Y_F = \frac{F - 1}{((F + 1)^2 - 4r^2F)^{\frac{1}{2}}} \quad (20)$$

Avec $n - 2$ degrés de liberté.

Le test t de student :

Ce test paramétrique repose sur la comparaison des moyennes. Il permet de comparer deux échantillons appariés des données quantitatives. Lorsque l'on souhaite comparer plus de deux échantillons, il convient d'utiliser l'ANOVA.

L'hypothèse nulle du test est : « $H_0 : \mu_a - \mu_b = 0$ ».

Le principe du test est le suivant. Nous avons deux échantillons A et B à évaluer. Nous calculons la différence d_i entre A et B pour chaque individu, et nous sommes d'un côté toutes les différences (S_d), et les carrés des différences (S_d^2). Ensuite, nous calculons la moyenne des différences :

$$m_d = \frac{S_d}{n} \quad (21)$$

Avec n le nombre des individus.

La somme des carrés des différences vaut :

$$SCE_d = S_d^2 - \frac{S_d^2}{n} \quad (22)$$

Enfin, la valeur critique du test est donnée par :

$$t = \frac{\frac{S_d}{n}}{\sqrt{\frac{SCE_d}{n(n-1)}}} \quad (23)$$

Avec $n - 1$ degrés de liberté.

Pour la conclusion, nous fixons un seuil de significativité, et nous comparons la valeur t critique du test avec ce seuil. Si $t < \text{seuil}$, alors nous rejetons H_0 .

2.3.3. Les tests non paramétriques

Nous pouvons distinguer deux types de tests non paramétriques : des tests non paramétriques pour deux échantillons d'observations appariées (Tableau 2), et des tests non paramétriques pour des échantillons d'observations appariées supérieurs à deux (Tableau 3).

Tableau 2 : Deux tests statistiques non paramétriques pour deux échantillons appariés.

Tests	Nature des individus	Hypothèses sur les observations	Pourquoi ?	Commande R
Test de Mann-Whitney	Dépendants	-	Comparaison des rangs	<code>wilcox.test (paired=TRUE, alternative=" ")</code> Ou library (coin) <code>wilcoxsign_test (x, y, distribution = "exact")</code>
Permutation	Dépendants	-	Comparaison des distributions	library (perm) <code>permTS (alternative=" ", method="pclt")</code>

Le test de Mann-Whitney :

Le test de Mann-Whitney⁵ est un test qui permet de comparer deux échantillons d'observations A et B appariés, appliqués aux individus dépendants, dont l'hypothèse nulle est « H_0 : pas de différence entre les moyennes ». Ce test est basé sur les rangs, il donne plus de poids à une paire qui montre une large différence entre les deux échantillons d'observations qu'à une paire ayant une faible différence. Cela implique que nous pouvons dire quel membre d'une paire est plus grand que l'autre en donnant le signe de la différence, mais aussi que nous pouvons ranger les différences en ordre croissant.

Soit d_i la différence entre chaque paire, représentant la différence entre les données appariées obtenues lors des deux échantillons. Chaque paire a un d_i . Nous trions les valeurs absolues de ces différences, nous attribuons à chaque rang le signe de la différence et nous sommes les rangs négatifs et les rangs positifs.

Si les traitements A et B sont équivalents, donc H_0 est vraie, la somme des rangs ayant un signe positif et celle des rangs ayant un signe négatif devraient être à peu près égale. Mais si la somme des

⁵ <http://www.cons-dev.org/elearning/stat/St4.html>

rangs de signes positifs est très différente de celle des rangs de signes négatifs, nous en concluons que l'échantillon A diffère de l'échantillon B, donc nous rejetons l'hypothèse nulle.

Le test de permutation :

Les tests de permutation, appelés aussi randomisation, sont particulièrement utiles lorsque nous avons à comparer des échantillons d'observations appariées ne vérifiant pas la normalité de leurs distributions et/ou qu'ils sont petits, deux échantillons ou plus. Le principe de ces tests est de trouver une statistique pour évaluer dans quelle mesure les scores présentent la structure en question (Rudolph 1995). Nous appliquons cette statistique aux valeurs des données d'origine, et nous la comparons avec la distribution obtenue en permutant les données. Nous répétons cette procédure un nombre important de fois (1000 ou autre) pour pouvoir conclure.

Si l'hypothèse nulle est vraie, alors toutes les combinaisons possibles des observations sont équiprobables.

Tableau 3 : Trois tests statistiques non paramétriques pour $k > 2$ échantillons appariés.

Tests	Nature des individus	Hypothèses sur les observations	Pourquoi ?	Commande R
Friedman	Indépendants	-	Comparaison des variances	<code>friedman.test(valeurs_numeriques~echantillon individu)</code> Ou library(agricolae) <code>friedman(individu, echantillon, valeurs_numeriques)</code>
Permutation	Dépendants	-	Comparaison des moyennes	library(perm) <code>permKS(alternative=" ", method="pclt")</code>
Cochran	Dépendants	-	Comparaison des moyennes	library(RVAideMemoire) <code>cochran.qtest(valeurs_binaires~echantillon individu)</code>

Le test de Friedman :

Le test de Friedman est appliqué sur des individus indépendants et échantillons d'observations appariées. Il est basé sur les rangs, nous trions les données pour chaque individu entre les échantillons, en prenant en compte les rangs identiques tel que pour des données identiques nous leur associons le rang moyen de l'ensemble. Ensuite, nous sommions les rangs par échantillon d'observations, et nous obtenons un tableau des rangs au lieu des valeurs de base (Friedman 1937). L'hypothèse nulle est « H_0 : pas de différence entre les variances des observations ».

La formule de la statistique pour le test de Friedman est :

$$F = \left(\frac{12}{nk(k+1)} \sum_{i=1}^k R_i^2 \right) - 3n(k+1) \quad (24)$$

Où :

- R_i^2 : la somme des rangs pour l'échantillon d'observation i ,

- k : le nombre d'échantillon
- n : la taille des échantillons

Pour conclure, nous utilisons la table de χ^2 et nous comparons la valeur de la statistique F avec la valeur critique de la table, dans le cas où la statistique F est supérieure à la valeur critique, nous rejetons l'hypothèse H_0 , sinon nous ne la rejetons pas.

Test Q de Cochran :

Le Q test de Cochran est un test de comparaison des distributions d'une variable qualitative binaire (0 : Echec / 1 : Succès) au sein de trois échantillons d'observations appariées ou plus. D'après certains auteurs, ce test est un des cas de test de Friedman ; il est basé sur la médiane (Cochran 1950).

Dans le cas où les données ne sont pas binaires, nous devons les transformer pour avoir un nouveau tableau qui ne contient que les valeurs 0 et 1. Pour cela, nous regroupons les données dans un seul vecteur, nous calculons sa moyenne, et nous la comparons avec les données de base. Si la valeur est supérieure à la médiane, nous remplaçons la valeur avec 1, sinon nous mettons 0. Au final, nous obtenons un nouveau tableau avec des valeurs 0 et 1 qui représente en ligne les individus, et en colonne les échantillons d'observations.

La statistique de ce test est donnée par :

$$Q = n(n - 1) \frac{\sum_{i=1}^T (C_i - \frac{S}{T})^2}{\sum_{k=1}^n L_k (K - L_k)} \quad (25)$$

Avec :

- T : le nombre des échantillons,
- C_i : le nombre de "1" dans la i^e colonne,
- L_k : le nombre de "1" dans la k^e ligne,
- S : la somme totale des "1" du tableau,
- N : le nombre des individus

Pour la conclusion, nous utilisons la table de χ^2 : si Q est supérieure à la χ^2 , nous rejetons H_0 et sinon, nous ne la rejetons pas.

2.3.4. Les tests statistiques dans l'équipe hydrologie des bassins versants

Certains tests statistiques ont été déjà mis en place dans des travaux de recherche de l'équipe. Le Tableau 4 suivant montre différents cas d'utilisation de ces tests.

Les individus statistiques correspondent, dans tous les cas, aux bassins versants. Les échantillons sont différents d'un cas à l'autre selon le sujet de recherche.

Tableau 4 : Différents cas d'utilisation des tests statistiques dans l'équipe HYDRO.

Test statistique	Pourquoi ?	Conditions d'utilisation
Friedman	Comparaison des différences significatives dans la performance du modèle GR4J à différents pas de temps (6min à 1 jour) (Ficchi et al. 2016).	Non vérifiées : problème d'indépendance des bassins.
Friedman	Évaluation des méthodes de prévisions des débits (Crochemore 2016).	Faiblement vérifiées : possible problème d'indépendance des bassins.
Permutation	Comparaison des deux systèmes de prévision des débits pour 569 jours et pour une échéance de prévision limitée à 2 jours (Thirel et al. 2008).	Vérifiées.
Permutation	Détection des changements progressifs dans le comportement hydrologique des bassins versants (Andréassian et al. 2003).	Vérifiées.

Ficchi et al. (2016) ont essayé d'examiner dans quelle mesure la performance de la modélisation hydrologique est améliorée, à l'aide d'un grand ensemble de 240 bassins versants français, pour 2400 événements d'inondation sélectionnés. Le modèle utilisé est le modèle pluie-débit GR4 adapté pour prendre en entrée des précipitations à pas de temps variables (dans la publication, ont été testés huit pas de temps 6, 12, 30 min ; 1, 3, 6, 12 h ; 1 jour). Pour tester la significativité de différences entre les résultats des modèles à différents pas de temps, le test de Friedman est appliqué. Cependant, ce test est appliqué sans tenir compte de la corrélation des débits des 240 bassins versants : les pluies qui tombent sur un bassin versant peuvent se retrouver quelques heures plus tard sur un autre bassin, d'où l'existence d'une corrélation entre les bassins.

Crochemore (2016) a essayé, dans sa thèse, de faire progresser les connaissances sur la prévision saisonnière pour la gestion de réservoirs multi-usages. Les travaux ont évalué la qualité des prévisions saisonnières de pluie et de débit dans seize bassins versants français. Elle a utilisé le test de Friedman pour évaluer de nouvelles méthodes de prévision des débits dans ces bassins, notamment pour la prévision des étiages. Ce test était utilisé sans s'assurer de la condition de non corrélation des bassins. Cependant, les bassins étaient suffisamment éloignés les uns des autres pour laisser penser que l'indépendance était respectée.

L'article de Thirel et al. (2008) présente l'évaluation des performances de deux systèmes de prévisions afin de mieux anticiper les événements hydrologiques et plus généralement pour estimer la qualité des deux systèmes pour une période allant du 10 mars 2005 au 30 septembre 2006. Les deux ensembles de prévisions ont été utilisés comme entrée pour le même modèle hydrométéorologique, et ont été évalués sur l'ensemble de la France, 881 stations, pour 569 jours et pour une période limitée à 2

jours. Cette évaluation des systèmes s'est effectuée avec le test de permutation, qui ne nécessite aucune condition d'utilisation particulière.

Finalement, Andréassian et al. (2003) ont visé à détecter des changements progressifs dans le comportement hydrologique des bassins versants en utilisant le test statistique de permutation. Ils ont utilisé le modèle GR4J, pour trois cas de bassins versants. Tous les trois bassins ont une longue période de pluie et de ruissellement record (60, 35 et 40 ans). De même que l'article précédent, le test de permutation est appliqué en respectant les conditions d'utilisation.

3. Les données et le travail réalisé

3.1. Les données

Dans le cadre de ce stage, nous avons travaillé sur un échantillon de 605 bassins versants situés en France. Étant donné que nous avons utilisé les modèles GR4J, GR5J et GR6J et que certains bassins versants pouvaient être fortement enneigés, nous avons eu besoin des données utilisées en modélisation pluie-débit : précipitations, débits, évapotranspiration potentielle et températures au pas de temps journalier.

Les données hydrométriques sont issues de la banque HYDRO⁶, du ministère de l'écologie et de développement durable. Les données climatiques, quant à elles proviennent des réanalyses SAFRAN (Vidal et al. 2010) de Météo-France.

Les 605 bassins versants de notre échantillon sont globalement répartis sur l'ensemble du territoire français métropolitain, exception faite de la Corse ; ils sont présents dans l'ensemble des zones hydrographiques d'ordre 4 (SANDRE 2002), (Figure 8). Ils couvrent la plupart des situations hydrologiques que l'on peut rencontrer en France (bassins versants de montagnes, méditerranéens, soumis à de fortes crues ou de fortes sécheresses, etc.). Néanmoins, ils ne représentent qu'une partie des données disponibles dans la Banque Hydro, qui contient les données de plusieurs milliers de stations hydrométriques). Les données des réanalyses SAFRAN se présentent sous la forme de chroniques climatiques couvrant une résolution spatiale de mailles de 8 km de côté. Quel que soit le bassin versant considéré, les données climatiques issues de réanalyses couvrent la période s'étalant du 1^{er} août 1958 au 31 juillet 2010. En revanche les données hydrologiques couvrent des périodes variables, mais nous n'avons conservé que celles étant incluses dans la période où les données climatiques sont disponibles ; elles peuvent être toutefois plus courtes que ces dernières.

⁶ <http://www.hydro.eaufrance.fr/>

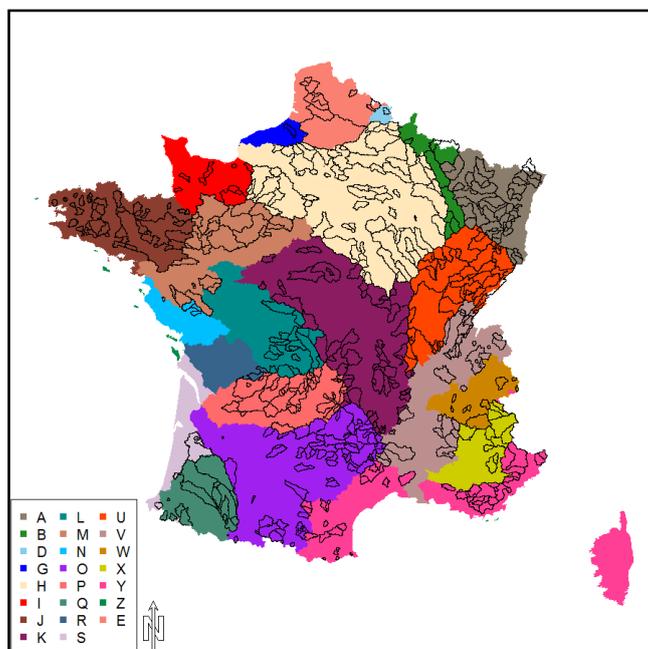


Figure 8 : Localisation des 605 bassins versants étudiés au sein du découpage en zones hydrographiques d'ordre 4.

Avant d'entamer le travail, il faut fixer les variables à étudier. Dans notre cas, nous avons deux variables : la première est le débit observé Q qui a servi à l'étude de l'autocorrélation et la corrélation entre les données des bassins versants, et la deuxième qui est le score de critère d'évaluation pour étudier la différence entre les modèles.

3.2. La chaîne de traitement réalisée

Durant ce stage, une chaîne de traitement composée de trois étapes a été mise en place :

Mise en forme des données d'entrées et analyses descriptives :

Il s'agit de mettre en forme les données pour qu'elles soient proprement organisées. À cette étape, nous avons également un travail exploratoire sur les données en cartographiant les bassins versants de l'échantillon (Figure 8), en traçant les chroniques de pluies et de débits observées à chacune des stations hydrométriques et en réalisant des analyses d'autocorrélation et de corrélation des débits.

Calage des modèles :

Cette étape consistait à caler les modèles hydrologiques utilisés, à savoir GR4J, GR5J et GR6J. Elle est effectuée sous le logiciel de programmation libre R⁷, en utilisant le package airGR⁸. Ce package contient les principaux outils de modélisation hydrologique utilisés par l'équipe Hydrologie des bassins versants. Chaque cœur des modèles est codé en FORTRAN pour assurer un temps rapide de l'exécution. Les autres fonctions du package : les critères d'évaluation (NSE, KGE et RMSE), et l'algorithme de calage, sont codées en R.

Après avoir fait tourner le code sur R, nous avons obtenu des résultats que nous devons analyser. Nous avons étudié la normalité des erreurs des critères d'évaluations.

⁷ <https://www.r-project.org/>

⁸ Téléchargement : <http://webgr.irstea.fr/airgr/>

Tests statistiques :

Dans cette dernière phase de l'étude, nous avons appliqué des tests statistiques pour évaluer la différence des modèles selon les critères.

Le schéma (Figure 9) suivant montre la chaîne d'étude.

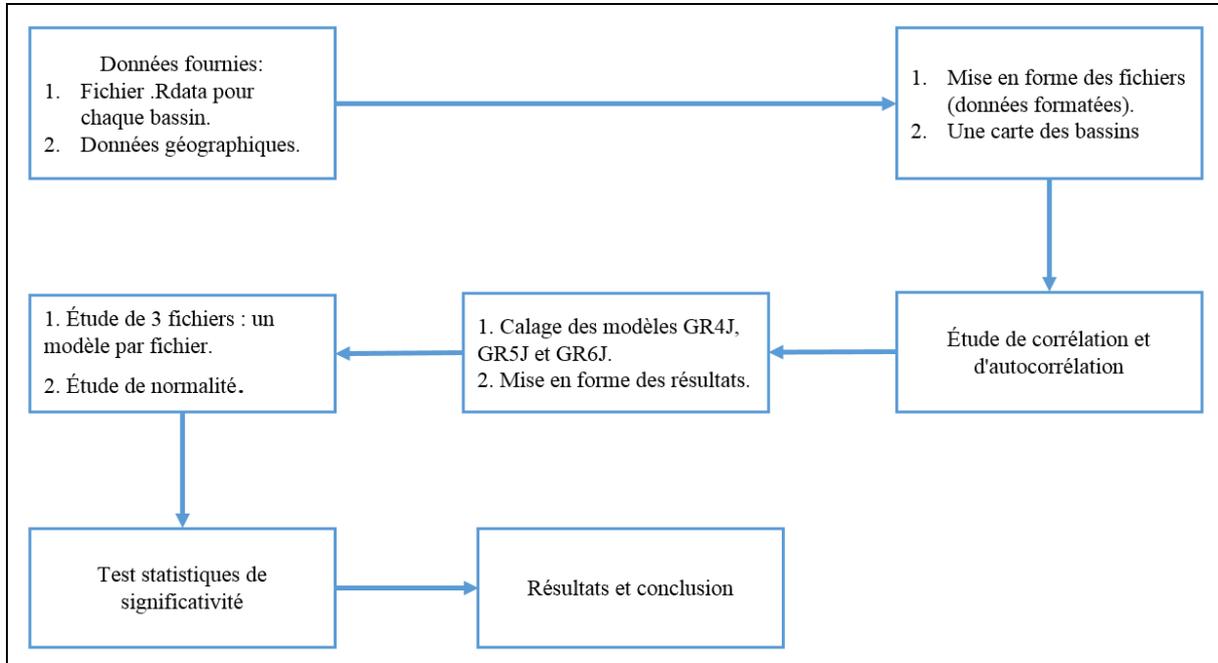


Figure 9 : Schéma des étapes de travail durant le stage.

4. L'étude des variables

Dans cette partie du rapport, nous allons présenter les résultats obtenus en étudiant les variables : les débits observés Q_{obs} et les critères d'évaluations. Ce travail a pour but de vérifier leurs caractéristiques, et donc les conditions d'application des tests statistiques.

Nous avons utilisé la variable Q_{obs} pour étudier l'autocorrélation et la corrélation entre les bassins versants. Cette variable représente les séries temporelles des débits observés. Chaque bassin versant a sa propre période d'étude, mais elles sont toutes comprises dans une période qui s'étale du 1^{er} août 1958 au 31 juillet 2010. Il n'existe cependant pas de période pour laquelle l'ensemble des bassins considérés comprend des données de débits.

Pour chaque bassin versant, nous avons calculé sa propre autocorrélation, et comme nous avons travaillé sur 605 bassins versants donc nous avons 605 graphiques et résultats d'autocorrélations.

Dans l'étape qui suit, nous avons calculé la corrélation entre tous les bassins versants. Deux démarches ont été adoptées, et sont basées sur le découpage en zones hydrologiques :

1. Nous nous sommes intéressés à la mesure de la corrélation entre des bassins versants appartenant aux mêmes zones hydrologiques ; de telle façon pour chaque bassin X d'une zone X, nous avons calculé sa corrélation avec les autres bassins de même zone X.

2. Nous nous sommes intéressés à la mesure de la corrélation entre des bassins versants appartenant à des zones hydrologiques différentes, c'est-à-dire, nous avons calculé la corrélation entre chaque bassin X d'une zone X avec les autres bassins de toutes les autres zones.

Les questions posées sont les suivantes :

- Qu'est-ce qu'une corrélation, et une corrélation croisée des bassins versants ?
- Que représentent les zones hydrologiques ?

La dernière étape consistait à étudier la normalité des critères de performances des débits. Pour chaque bassin parmi les 605, nous avons calculé ses valeurs des trois critères utilisés pour chacun des trois modèles.

Tous ces calculs d'autocorrélation, de corrélation et de normalité, ont été effectués et programmés sur R pour faciliter la tâche.

4.1. L'autocorrélation des bassins versants

Nous avons commencé par étudier l'autocorrélation de chaque bassin versant. Vu que nous avons 605 bassins, il est impossible de tous les analyser et les présenter. Nous avons choisi de les regrouper par surface. Nous avons fait trois classes : les petites surfaces, moyennes surfaces et grandes surfaces. Nous nous sommes basés sur les quantiles ; 5% pour les petits bassins et 95% pour les grands.

La surface maximale des petits bassins est fixée à 31,34 km², et la surface minimale des grands bassins versants est fixée à 2903,81 km². Entre ces deux valeurs, nous avons les bassins de moyennes surfaces. La caractéristique surface des bassins influe sur leurs réactions hydrologiques. Cette réaction est mesurée par l'observation de la quantité d'eau qui s'écoule à l'exutoire du système : plus le bassin est grand, plus le temps de réponse s'allonge.

L'autocorrélation d'une série fait référence au fait que dans une série temporelle ou spatiale, la mesure d'un phénomène à un instant t peut être corrélée aux mesures précédentes (au temps $t-1$, $t-2$, $t-3$, etc.) ou aux mesures suivantes (à $t+1$, $t+2$, $t+3$, ...). Une série autocorrélée est ainsi corrélée à elle-même, avec un décalage (*lag*) donné.

La définition mathématique de l'autocorrélation de la variable $Q_{obs,t}$, de moyenne μ et d'écart-type σ , pour un décalage k est donnée par :

$$\rho_k = \frac{\gamma_k}{\sigma^2} \quad (26)$$

Avec γ_k est l'autocovariance de $Q_{obs,t}$:

$$\gamma_k = E((Q_{obs,t} - \mu)(Q_{obs,t+k} - \mu)) \quad (27)$$

Dans ce qui suit, nous présenterons les graphiques d'acf (*Autocorrelation Function*) nommés « corrélogramme » (à droite) et de nuage de points des débits en fonction des débits pour un décalage de pas de temps (à gauche), lot précédent de débits présente le décalage $t-1$ des débits, et ce pour trois bassins versants de deux classes de surface. Les deux catégories contiennent 31 bassins versants.

Les petits bassins versants :

Les surfaces dans cette classe sont inférieures ou égales à 31,34km².

- Le bassin de la Neste de Cap de Long à Aragnouet :

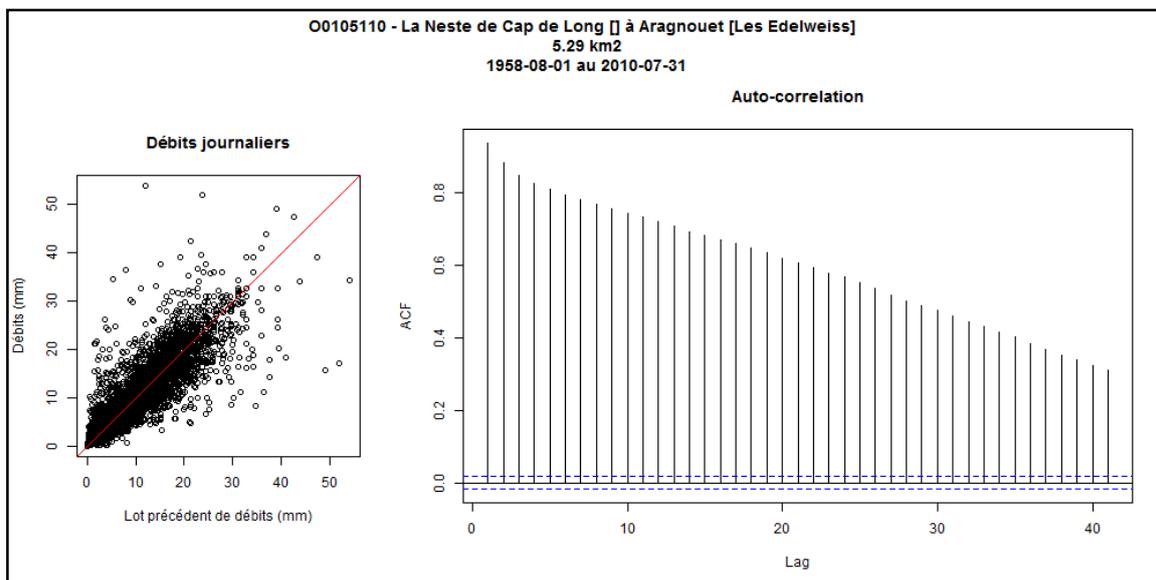


Figure 10 : Autocorrélation du bassin de la Neste de Cap de Long à Aragnouet - 5,29km².

Au décalage de 0 pas ou sans décalage, l'autocorrélation de la série des débits est forcément 1, et plus nous décalons dans le temps plus l'autocorrélation diminue. Elle est particulièrement forte (supérieure à 0,8) pour les 5 premiers décalages. Pour les 8 décalages d'après, l'autocorrélation est comprise entre 0,7 et 0,8. Et à partir de décalage de 29 pas de temps, elle est inférieure à 0,5.

D'après le graphique à gauche, nous remarquons que les débits observés en fonction des débits aux mesures précédentes sont présentés sur la diagonale, avec un peu de dispersion, ce qui signifie que la corrélation des débits n'est pas très élevée.

- Le bassin de l'Aber-Benoit à Plabennec :

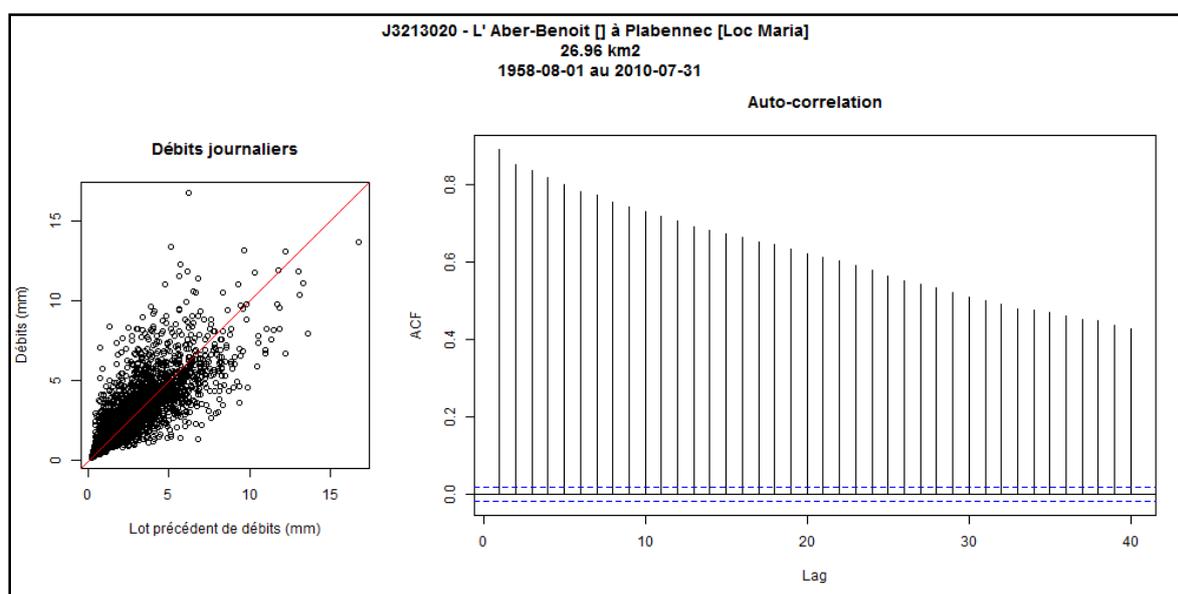


Figure 11 : Autocorrélation du bassin de l'Aber-Benoit à Plabennec - 26,96km².

Aux 5 premiers décalages, la corrélation est forte ; supérieure à 0,8. Elle est comprise entre 0,7 et 0,8 aux 7 décalages d'après. Et à partir du décalage de 31 de pas de temps, l'autocorrélation est inférieure à 0,5.

Sur le graphique à gauche, il existe une dispersion pour les fortes valeurs des débits, par ailleurs les plus faibles sont assez corrélées.

- Le bassin du Strengbach à Ribeauvillé :

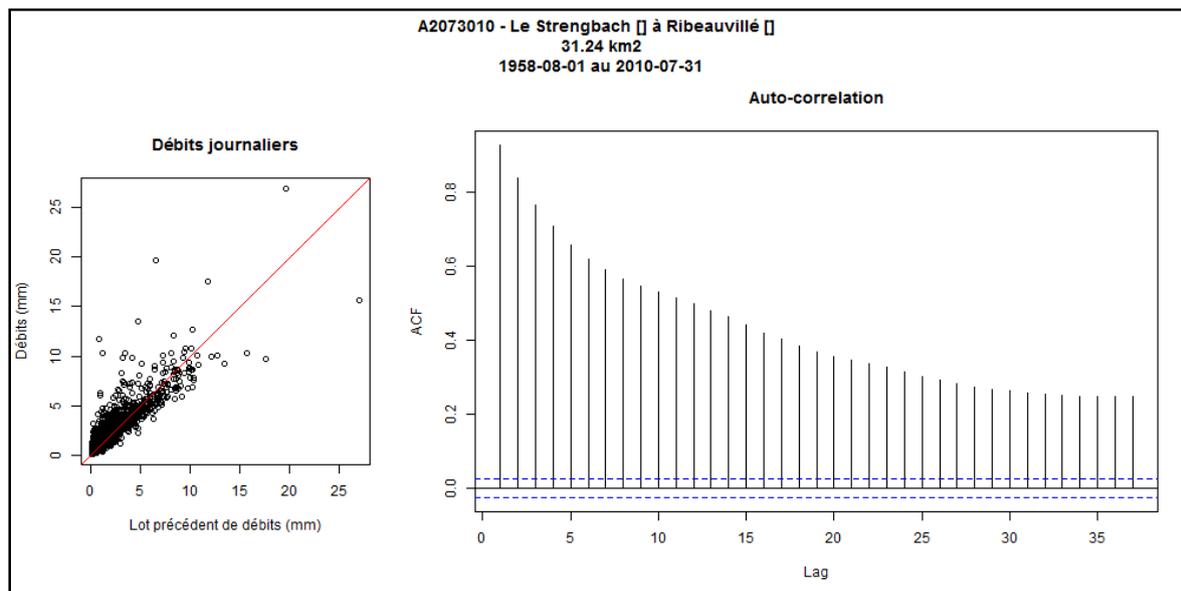


Figure 12 : Autocorrélation du bassin du Strengbach à Ribeauvillé - 31,24km².

Aux deux premiers décalages, nous avons une autocorrélation forte supérieure à 0,8. Du 3^e décalage au 11^e, l'autocorrélation est supérieure à 0,5. Et à partir de décalage de 12 pas de temps, l'autocorrélation est inférieure à 0,5.

Quant au graphique de nuages des points, nous avons une forte corrélation pour les petites valeurs de débits avec un peu de dispersion pour quelques-unes, alors que pour les grandes valeurs supérieures à 15mm, elles ne sont pas présentées sur la diagonale.

Les grands bassins versants :

Les surfaces dans cette classe sont supérieures ou égales à 2903,814 km².

- Le bassin de la Moselle à Toul :

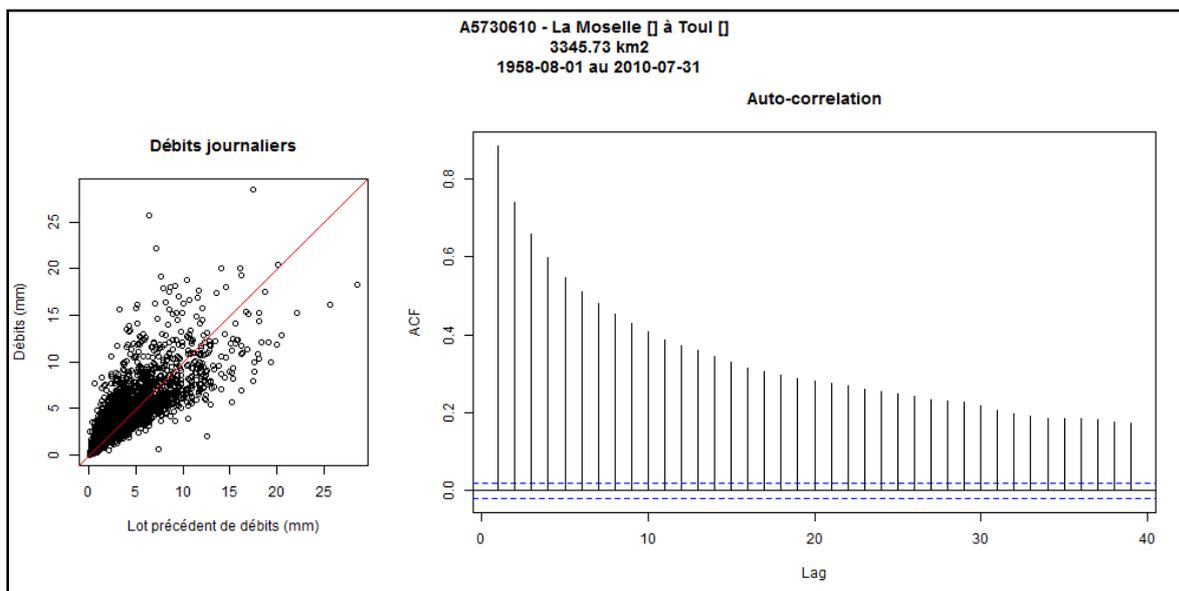


Figure 13 : Autocorrélation du bassin de la Moselle à Toul - 3345,73 km².

L'autocorrélation est particulièrement forte pour le premier décalage avec une valeur égale à 0.88, et le deuxième décalage 0.73. À partir du 7^e décalage, l'autocorrélation est inférieure à 0,5. Sur le graphique à gauche, nous remarquons une dispersion pour les grandes valeurs de débits.

- Le bassin de l'Oise à Pont-Sainte-Maxence :

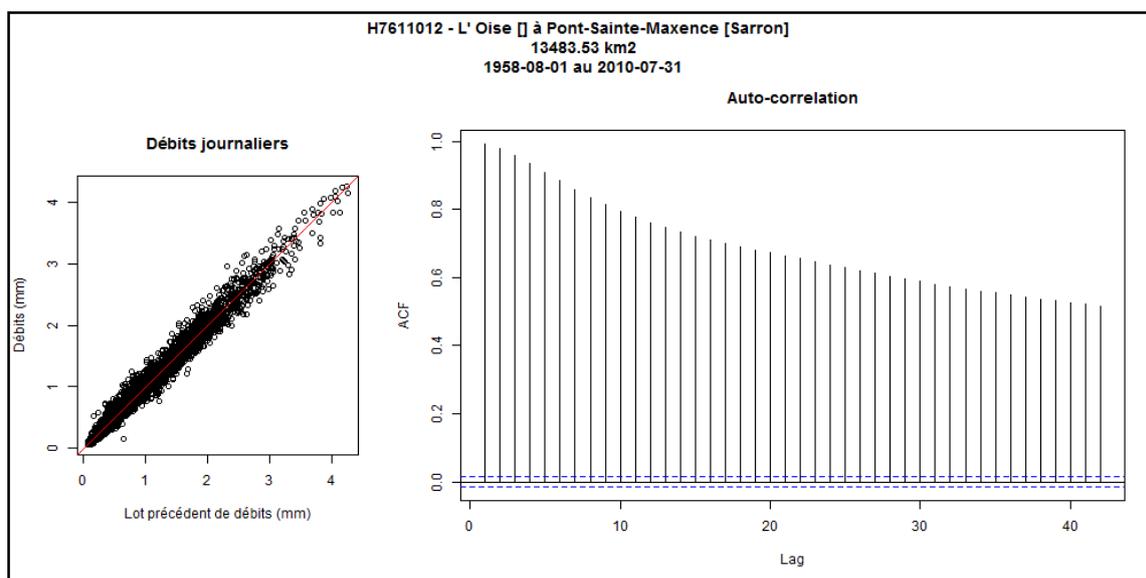


Figure 14 : Autocorrélation du bassin de l'Oise à Pont-Sainte-Maxence – 13483,53 km².

Nous remarquons une forte autocorrélation aux 9 premiers décalages, supérieure à 0,8. À partir du décalage de 45, l'autocorrélation est inférieure à 0,5. Et entre ces deux ensembles, les valeurs sont comprises entre 0.50 et 0.79.

De plus, nous avons une forte corrélation au niveau du graphique à gauche ; les valeurs des débits sont bien présentées sur la diagonale.

- Le bassin de la Saône à Mâcon :

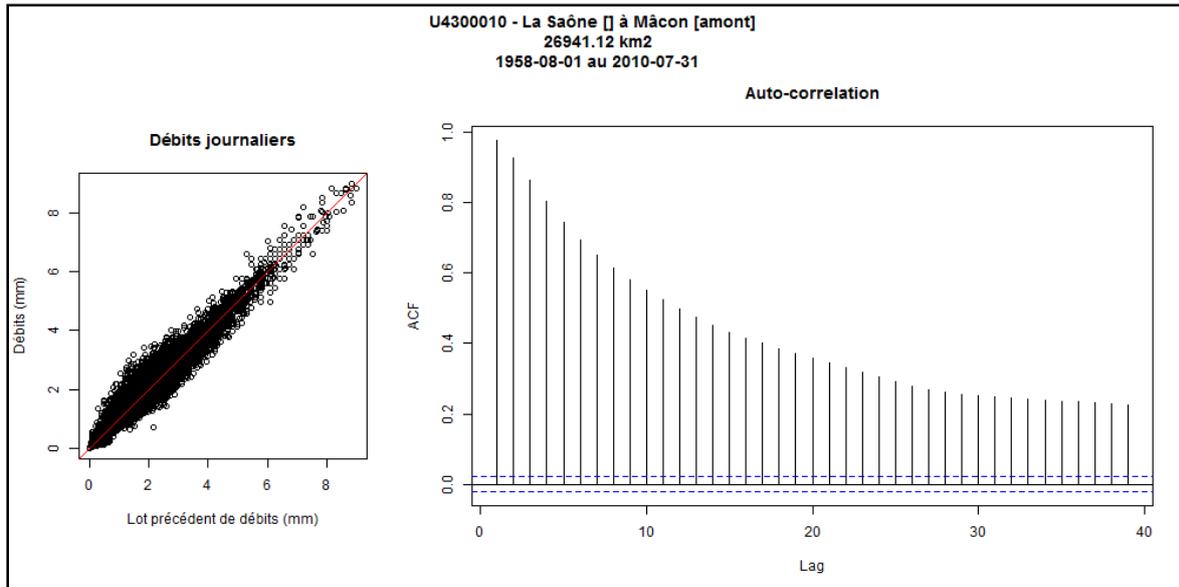


Figure 15 : Autocorrélation du bassin La Saône à Mâcon – 26941,12 km².

Pour ce bassin, les remarques sont les mêmes que pour le bassin de l’Oise : une autocorrélation de la variable débit existe pour ce bassin.

Conclusion :

D’un point de vue hydrologique, les bassins versants de petites surfaces peuvent ne pas garder la quantité d’eau ou de neige reçue à l’exutoire le jour k pour le jour k+1 ou les jours suivants, ce qui implique une plus faible corrélation des débits, d’autant plus que le décalage est important.

Par ailleurs, plus le bassin est grand plus le temps de réponse s’allonge. La quantité des précipitations peut durer des jours et des mois pour disparaître de l’exutoire (c’est le temps de réponse du bassin, qui agit comme un « tampon »). Ceci implique une forte autocorrélation des débits.

Pour conclure, les résultats obtenus valident la vision hydrologique des bassins versants.

4.2. La corrélation des bassins versants

La corrélation :

En statistique, la corrélation entre deux bassins ou plus désigne la liaison qui peut exister entre eux. Son coefficient est compris entre -1 et 1. Plus nous approchons de 1 (ou -1), plus nous avons une forte corrélation (ou anti-corrélation), et une corrélation égale à 0 signifie que les bassins sont non corrélés.

La définition mathématique de la corrélation entre deux bassins X et Y, tel que σ_X et σ_Y présentent leurs écarts-types, est donnée par cette formule :

$$\text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \quad (28)$$

Avec $\text{cov}(X, Y)$ est la covariance entre bassin X et bassin Y, calculée par :

$$\text{cov}(X, Y) = E[(X - E(X))(X - E(X))] = E(XY) - E(X)E(Y) \quad (29)$$

La corrélation croisée :

La corrélation croisée (ou intercorrélacion), ccf (*Cross-Correlation Function*), est utilisée pour designer la relation entre deux séries temporelles.

Soit deux séries temporelles de débits $Q_{A\text{obs}, t}$ et $Q_{B\text{obs}, t}$, correspondantes aux bassins versants A et B. La corrélation croisée est définie comme suit : la série $Q_{B\text{obs}, t}$ dans une période de temps θ_1 doit être en relation avec les décalages précédents de la série $Q_{A\text{obs}, t}$ dans une période de temps θ_2 . θ_1 et θ_2 présentent deux périodes différentes, et cela revient au fait que les moments d'enregistrement des débits pour le bassin A ne sont pas les mêmes pour le bassin B en raison du fait que les précipitations ne tombent pas simultanément sur les deux bassins versants. En revanche, nous pouvons trouver des cas des bassins pour lesquels la corrélation croisée s'applique sur une période de temps commune.

Mathématiquement, la corrélation croisée est donnée par :

$$\rho_{A,B}(\theta) = \int Q_{A\text{obs}}(t)Q_{B\text{obs}}(t - \theta)dt \quad (30)$$

La classification des bassins en zones hydrologiques :

En France, les cours d'eau sont divisés en six grandes zones géographiques. Elles représentent les grands bassins : Seine-Normandie, Artois Picardie, Rhin Meuse, Rhône méditerranée et Corse, Loire-Bretagne et Adour-Garonne, et qui sont présentés sur la Figure 8 par 23 sous-zones : A, B, D, E, G, H, I, J, K, L, M, N, O, P, Q, R, S, U, V, W, X, Y et Z. Dans notre cas, nous disposons de toutes ces sous-zones sauf Z. Chaque sous-zone contient des bassins versants dont leurs codes commencent par la lettre correspondante. Nous avons travaillé sur la corrélation entre les bassins versants de chaque sous-zone.

Vu que tous les bassins versants d'une même zone sont corrélés au seuil 0,5 (ou -0,5), nous avons fixé la corrélation à 0,7 (-0,7). Nous avons défini 5 catégories, représentant la proportion de bassins au sein d'une même zone dont la corrélation est supérieure à 0.7.

Le Tableau 5 présente les proportions de la corrélation égale à 0,7, donc les différents niveaux de la corrélation pour les différentes zones hydrologiques.

Tableau 5 : la proportion de la corrélation égale à 0,7 des bassins versants.

Corrélation Groupes	Faible	Moyenne	Forte	Très forte	Totale
A			× (78 %)		
B			× (65 %)		
D					× (= 1)
E	× (40 %)				
G					× (= 1)
H		× (52 %)			
I			× (79 %)		
J				× (94 %)	
K	× (24 %)				
L				× (97 %)	
M			× (88 %)		

N					× (= 1)
O	× (22 %)				
P			× (82 %)		
Q		× (58 %)			
R					× (= 1)
S			× (68 %)		
U		× (50 %)			
V	× (26 %)				
W	× (30 %)				
X		× (53 %)			
Y	× (14 %)				

Avec :

Tableau 6 : Interprétation des niveaux de la corrélation.

Proportion de « corrélation égale à 0,7 »	Interprétation
< 40 %	Faible
€] 40 %, 60 % [Moyenne
€ [60 %, 90 % [Forte
€ [90 %, 100 % [Très forte
= 100%	Totale

Nous avons conclu que la majorité des bassins versants selon les sous-zones sont corrélés. Et cela nous a posé un problème dans l'application de certains tests statistiques de significativité. En effet, par exemple pour A, la proportion de la « corrélation = 0,7 » est 78 % donc une forte corrélation entre les bassins versants de la sous-zone A. Les D, G, N et R présentent 100 % « corrélation = 0,7 », quant aux sous-zones E, K, O, V, W et Y les proportions sont inférieures à 40 %, et les autres sous-zones sont divisées entre moyenne et forte corrélation.

L'étape qui suit consistait à étudier la corrélation des bassins versants entre les groupes. Pour cela, nous avons combiné tous les bassins, de façon que tous les bassins d'un groupe soient en permutation avec tous les autres bassins des autres groupes.

Les Figure 16, Figure 17, Figure 18 et Figure 19 présentent des exemples de la corrélation croisée de 4 bassins versants.

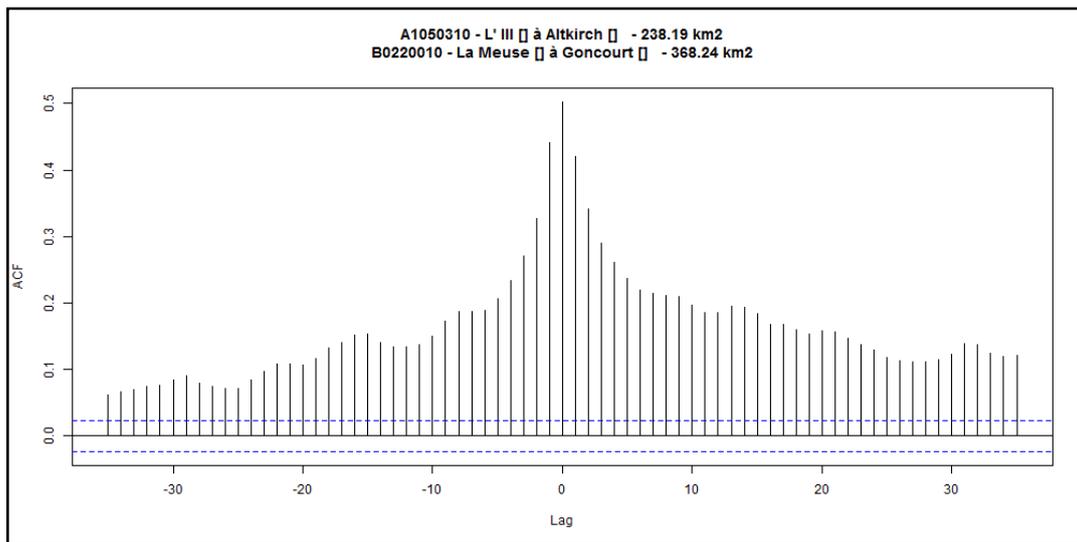


Figure 16 : Corrélation croisée entre deux bassins versants des zones hydrologiques A et B.

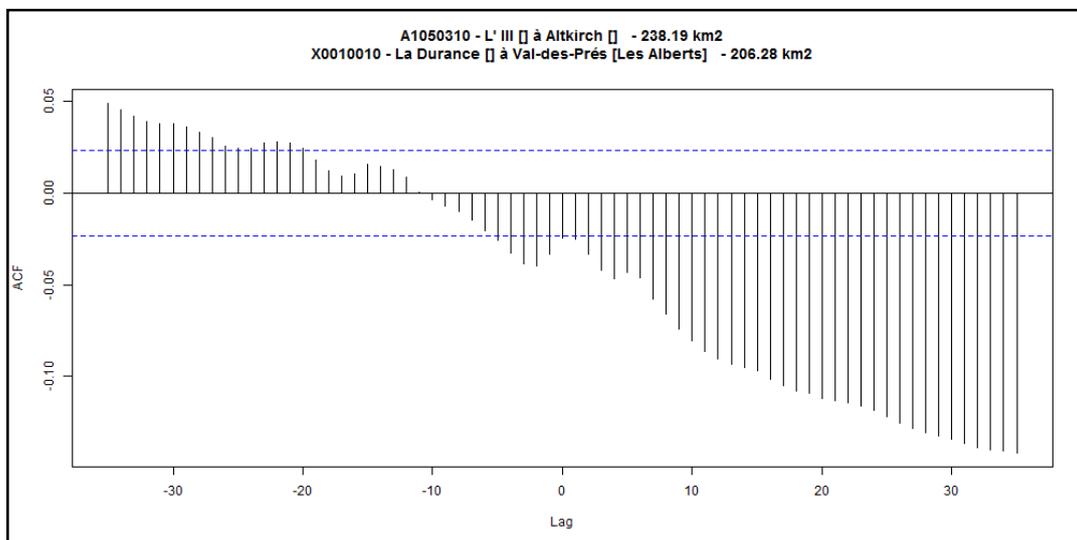


Figure 17 : Corrélation croisée entre deux bassins versants des zones hydrologiques A et X.

La Figure 16 montre une corrélation importante de 0,5 entre le bassin de l’Ile à Altkirch (A1050310) du zone A et le bassin la Meuse à Goncourt (B0220010) du zone B. En parallèle, une corrélation faible de 0,05 et -0,15 entre le bassin de l’Ile à Altkirch (A1050310) et le bassin la Durance à Val des Prés (X0010010) du zone X dans la Figure 17.

Un autre exemple de corrélation croisée (Figure 18 et Figure 19) entre le bassin de la Seine à Nod sur Seine (H0100010) et le bassin de la Soire à Ferrière la Grande (D0206010) sont corrélés avec un coefficient de corrélation égal à 0,63, par contre le bassin de la Seine à Nod sur Seine (H0100010) et le bassin la Roya à Breil sur Roya (Y6624010) ne sont pas corrélés.

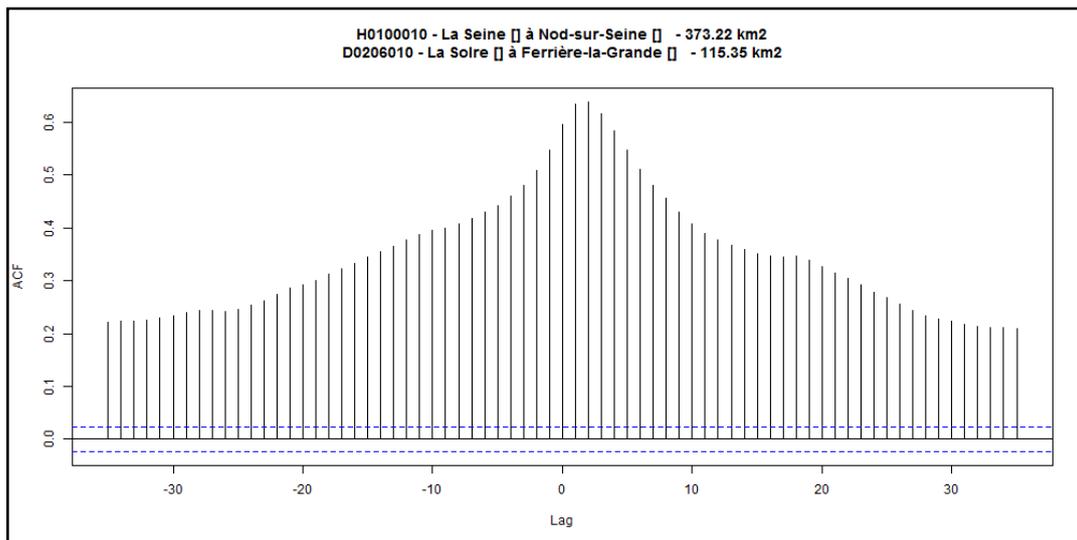


Figure 18 : Corrélation croisée entre deux bassins versants des zones hydrologiques H et D.

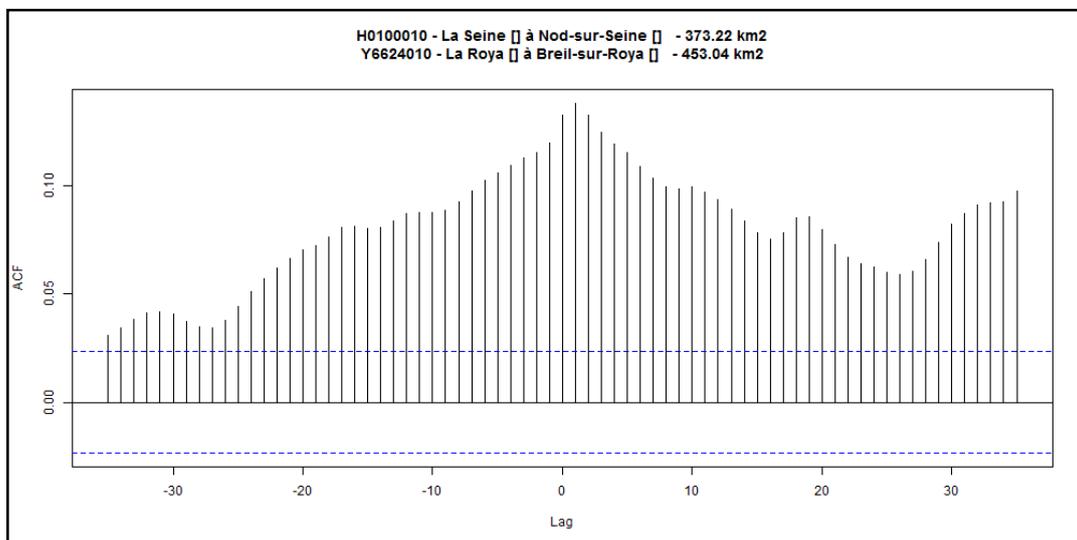


Figure 19 : Corrélation croisée entre deux bassins versants des zones hydrologiques H et Y.

Le résultat que nous avons pu conclure est : plus nous nous éloignons plus la corrélation est faible entre les bassins. Ce résultat est logique car l'éloignement géographique induit des différences météorologiques le plus souvent importantes. A l'opposé, les bassins de l'Il et de la Meuse ont une corrélation élevée sans décalage, car ils subissent la même météorologie.

Comme les bassins versants sont corrélés entre eux et quelques tests demandent la non corrélation des bassins versants, nous avons préparé une liste des bassins non corrélés, basée sur la corrélation croisée. Cette liste contient 22 bassins représentant un bassin par zone.

Nous avons commencé par fixer deux zones, et nous avons choisi un bassin par zone tel que la corrélation croisée entre ces deux bassins soit nulle ou faible. De même le coefficient de corrélation est fixé à 0,7. Par la suite, nous avons choisi un troisième bassin d'une troisième zone de telle façon la corrélation de ce bassin avec les deux premiers soit nulle ou faible. Nous avons répété cette procédure 22 fois afin d'avoir une nouvelle base de donnée de 22 bassins versants non corrélés.

4.3. La normalité des critères d'évaluation

Une distribution normale dite aussi gaussienne, correspond à la distribution de probabilités d'une variable aléatoire continue dont la courbe est parfaitement symétrique, unimodale et en forme de cloche. Formellement, une distribution normale de moyenne μ et d'écart type σ est définie par la formule :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (31)$$

Après avoir calé les modèles hydrologiques GR4J, GR5J et GR6J, nous avons obtenu les scores des 3 critères KGE', NSE et RMSE. Ces scores présentent la variable d'étude de la normalité. Pour cela, nous avons utilisé deux différentes méthodes pour tester la normalité de ces scores : la droite de Henry et les tests statistiques.

La droite de Henry ou le diagramme quantile-quantile, est une méthode pratique pour ajuster la distribution normale à celle d'une série d'observation. Concernant les tests statistiques, nous trouvons par exemple : le test de Shapiro-Wilks pour une taille de données de 3 à 5000, le test Kolmogorov-Smirnov dans le package `fBasic` et `lillietest()` dans le package `nortest`. Nous avons utilisé le test Kolmogorov-Smirnov et le test `lillietest` pour étudier la normalité des erreurs des critères d'évaluation.

Les Figure 20, Figure 21, et Figure 22 présentent les droites de Henry et le Tableau 7 montre les p-values des tests statistiques appliqués aux trois critères pour les trois modèles hydrologiques.

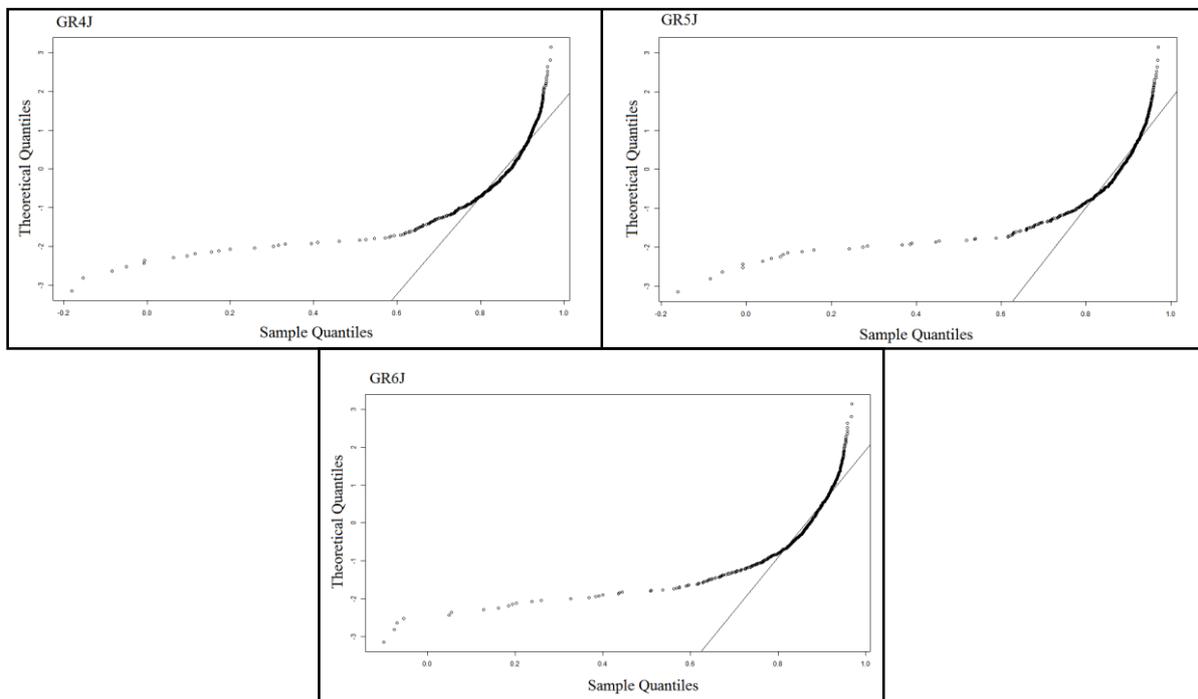


Figure 20 : Q-Q plot des scores de KGE'

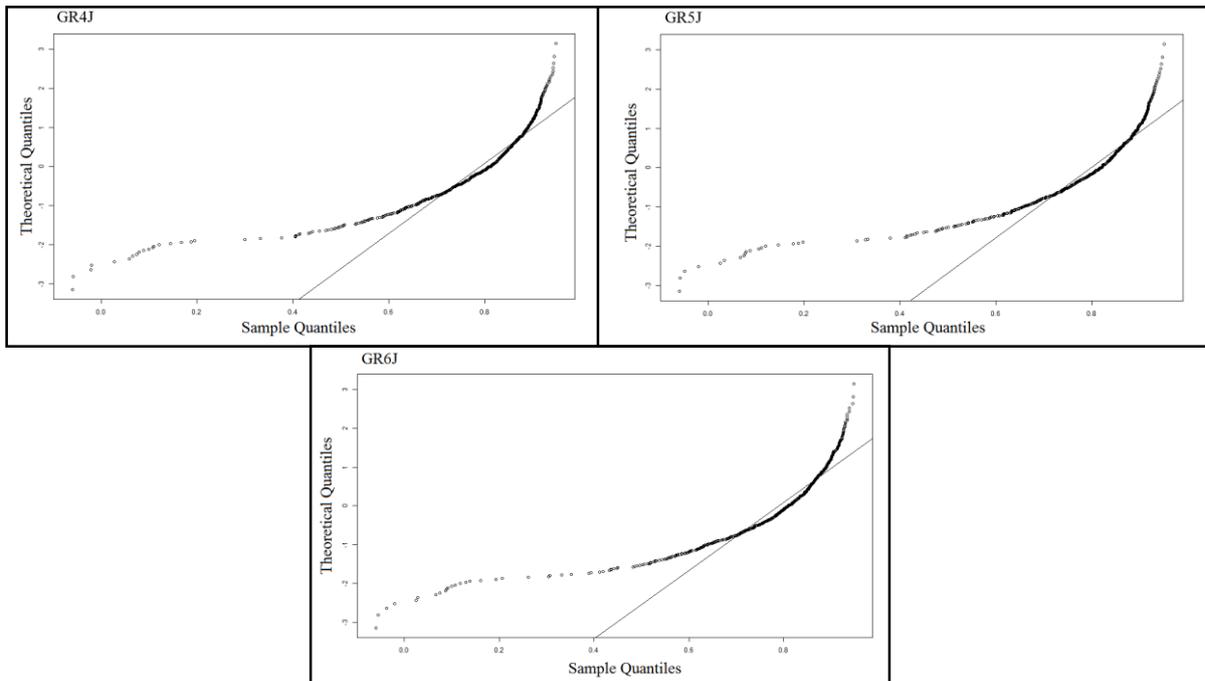


Figure 21 : Q-Q plot des scores de NSE

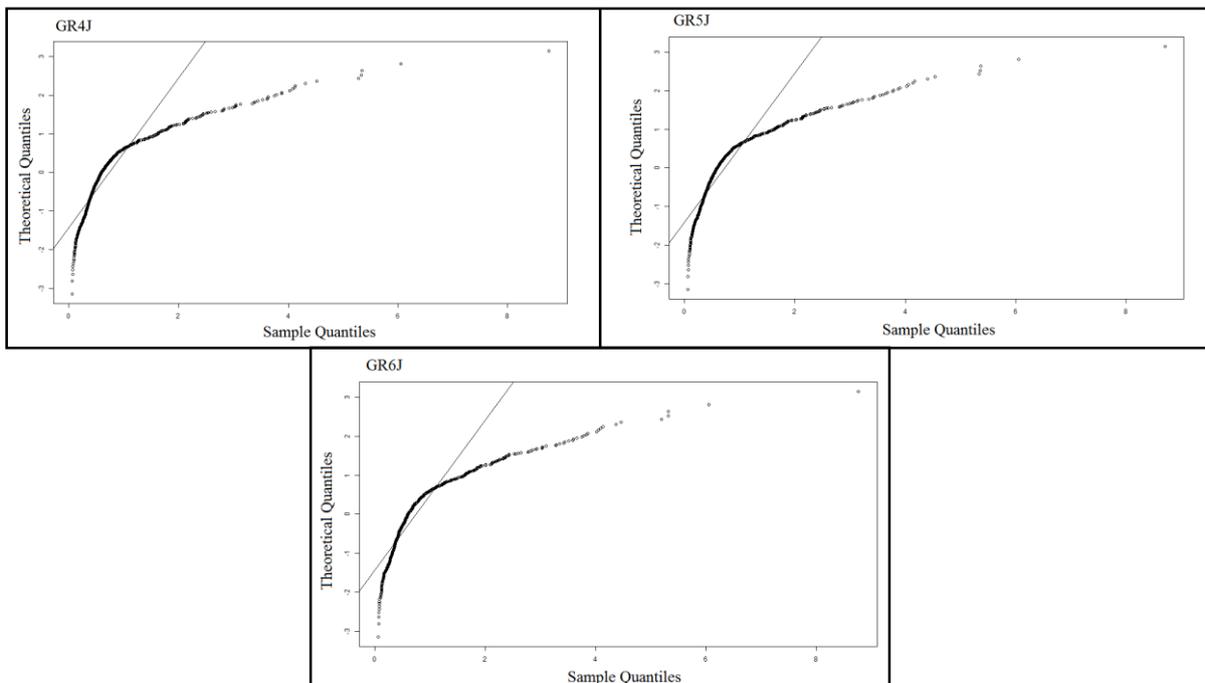


Figure 22 : Q-Q plot des scores de RMSE

Tableau 7 : Les p-values résultant des tests de normalité Kolmogorov-Smirnov et lillietest sur les critères d'évaluation des modèles hydrologiques.

Modèle	GR4J	GR5J	GR6J
Critère			
KGE'	$<2,2.e^{-16}$	$<2,2.e^{-16}$	$<2,2.e^{-16}$
NSE	$<2,2.e^{-16}$	$<2,2.e^{-16}$	$<2,2.e^{-16}$
RMSE	$<2,2.e^{-16}$	$<2,2.e^{-16}$	$<2,2.e^{-16}$

Graphiquement, il est évident que les tracés ne suivent pas la droite. De plus, d'après les tests, les p-values sont toutes inférieures au seuil 0,05. Nous concluons donc que les critères d'évaluation ne suivent pas la loi normale. Dans ce cas, soit nous utilisons les tests non paramétriques pour évaluer la

performance des modèles, soit nous essayons de transformer les scores des critères pour que ces derniers suivent une distribution normale pour appliquer les tests paramétriques.

Pour normaliser, nous pouvons utiliser la fonction logarithme $\ln(x)$, qui est utile pour des données positives ou encore $\ln(x+1)$ si la variable x prends des valeurs positives ou nulles, la fonction racine carré \sqrt{x} aussi pour des variables positives, ce qui n'est pas notre cas. En effet les scores d'erreurs prennent quelques valeurs négatives. Et en appliquant la fonction inverse $\frac{1}{x}$ sur nos données, nous avons trouvé ceci (Figure 23, Figure 24 et Figure 25) :

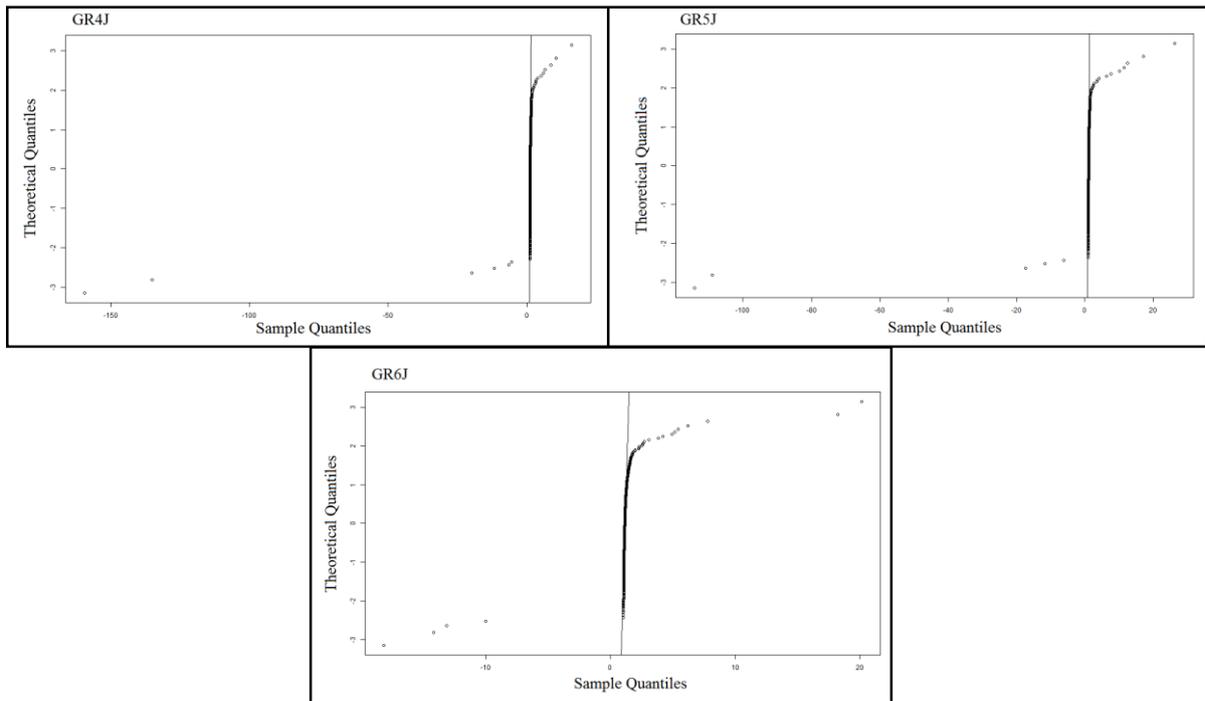


Figure 23 : Q-Q plot des scores normalisés de KGE'

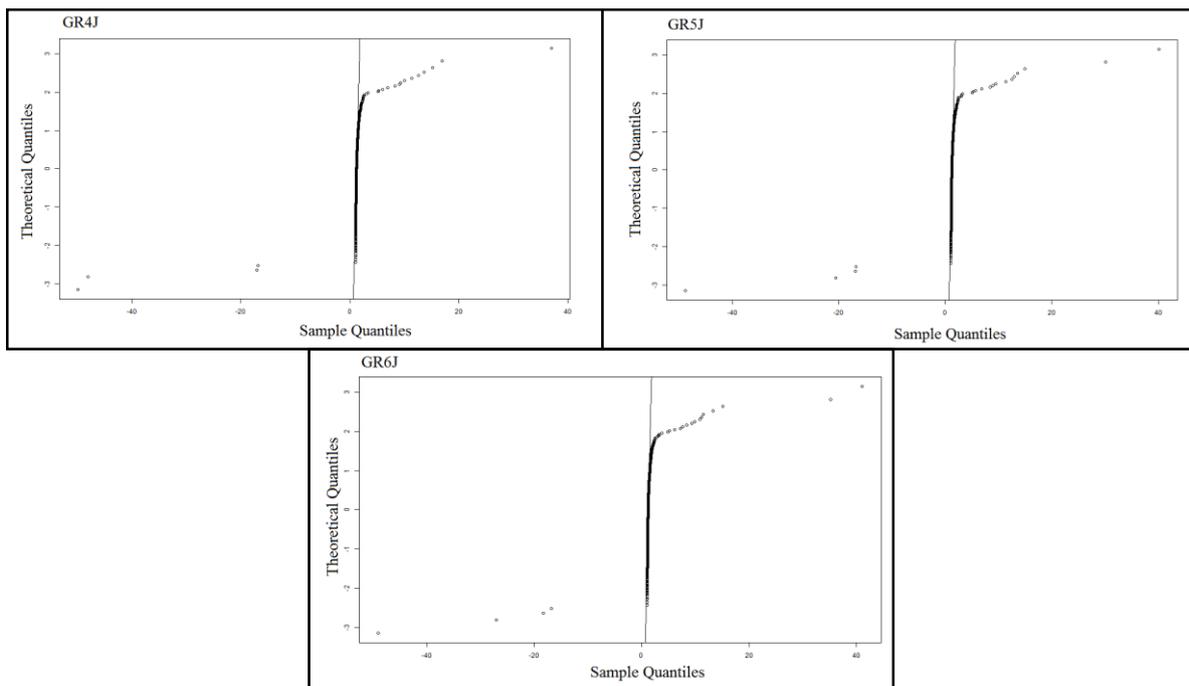


Figure 24 : Q-Q plot des scores normalisés de NSE

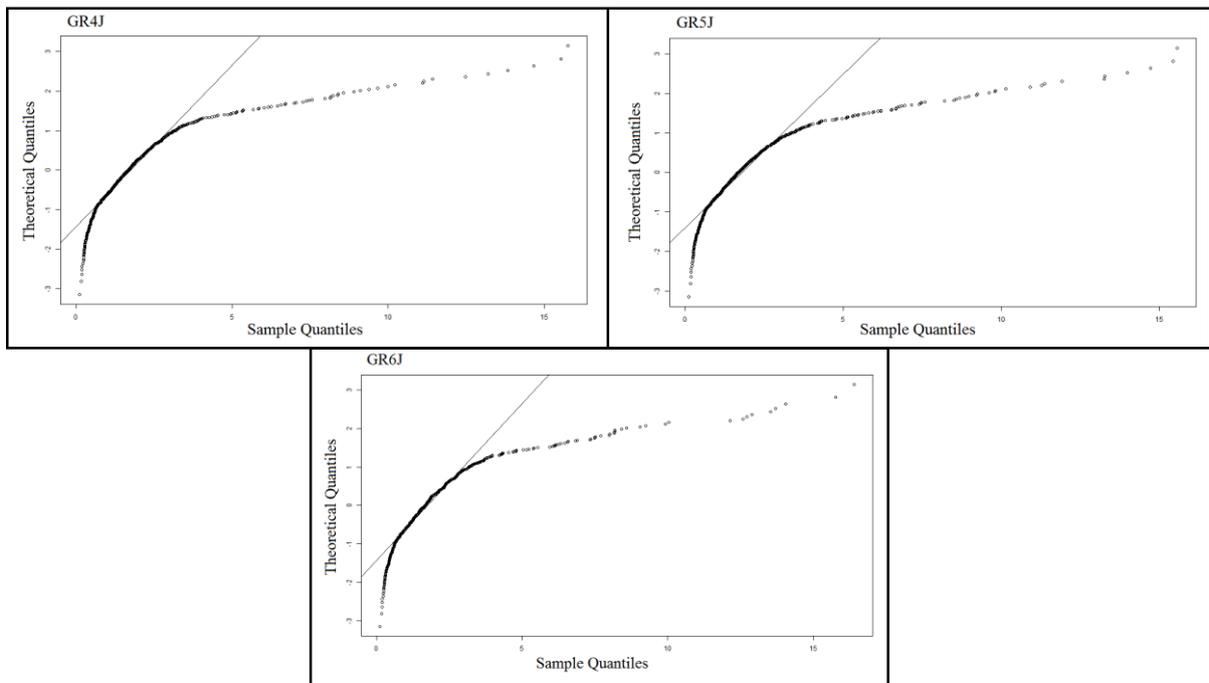


Figure 25 : Q-Q plot des scores normalisés de RMSE

Ces trois figures présentent les droites de Henry des trois critères d'évaluation KGE', NSE et RMSE normalisés par la fonction inverse. Nous remarquons que les critères ne suivent pas la loi normale, le même résultat est obtenu en appliquant les deux tests Kolmogorov-Smirnov et lillietest(). Il faudrait donc chercher une autre fonction croissante et monotone qui doit parfaitement normaliser les données. Nous ne l'avons pas fait faute de temps.

5. Les résultats d'application des tests de significativité

Nous avons étudié les débits et critères de performance des modèles hydrologiques. Nous avons trouvé que les bassins versants sont corrélés lorsque l'on utilise de larges jeux de données et les critères de performance des modèles hydrologiques ne suivent pas la loi normale. D'où la nécessité d'appliquer les tests statistiques non paramétriques.

Le but de ce stage est d'appuyer l'évaluation de la performance des modèles hydrologiques avec des tests statistiques de significativité. Alors, dans cette dernière partie du rapport, nous présenterons les résultats obtenus des performances.

En premier lieu, nous avons tracé des boîtes à moustaches pour comparer les distributions de critères de performances des différents modèles, et ce, sur notre échantillon de 605 bassins versants (Figure 26) :

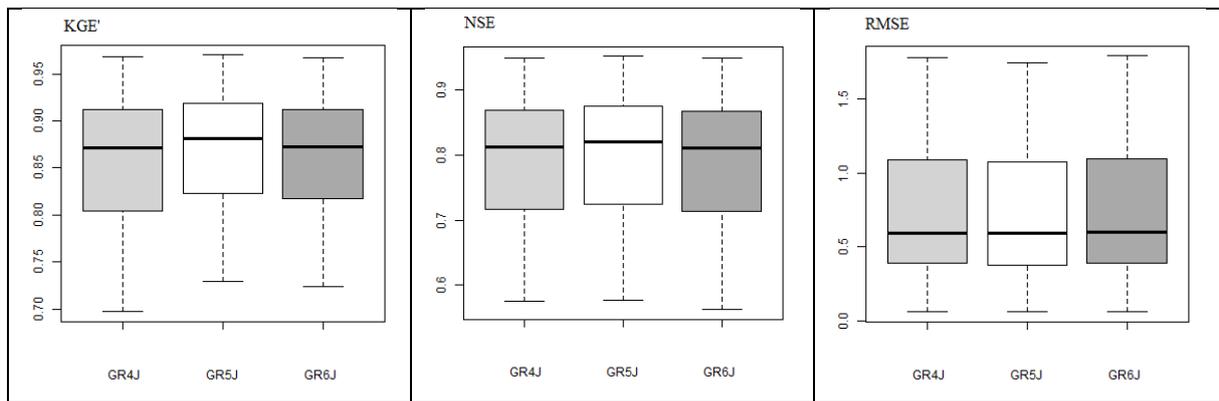


Figure 26 : Distributions de critères de performance en fonction des modèles pour 605 bassins versants.

D'après la Figure 26, nous remarquons que les modèles GR4J et GR6J sont à peu près égaux, alors qu'ils sont différents de GR5J.

En ce qui concerne les tests, nous avons appliqué le test Q de Cochran, le test de Mann-Whitney, et le test de permutation pour un seuil de significativité fixé à 0,05 sur les 605 bassins versants. Le Tableau 8 présente les résultats :

Tableau 8 : Les p-values des tests statistiques de significativité pour 605 bassins versants (en gras, les résultats significatifs)

Test	Permutation			Cochran			Mann-Whitney		
	KGE'	NSE	RMSE	KGE'	NSE	RMSE	KGE'	NSE	RMSE
GR4J-GR5J	0,5092	0,1812	0,9366	-	-	-	< 2,2e ⁻¹⁶	< 2,2e ⁻¹⁶	< 2,2e ⁻¹⁶
GR4J-GR6J	0,7534	0,8828	0,999	-	-	-	0,709	0,4694	0,1821
GR5J-GR6J	0,2973	0,425	0,9355	-	-	-	1,58e ⁻¹⁵	1,82e ⁻⁰⁶	0,0004
GR4J-GR5J-GR6J	0,3681	0,6939	0,9957	0,00051	0,00091	0,4204			

Nous concluons que les résultats du test de Mann-Whitney, et du test Q de Cochran pour les critères KGE' et NSE correspondant à ce que nous avons observé graphiquement, alors que les résultats sont différents pour le test de permutation et le test Q de Cochran pour RMSE.

De plus, nous avons essayé d'appliquer le test de Friedman sur les 605 bassins versants et nous avons trouvé ceci (Tableau 9) :

Tableau 9 : Les p-values du test de Friedman pour 605 bassins versants.

Test	Friedman		
	KGE'	NSE	RMSE
GR4J-GR5J	-	-	-
GR4J-GR6J	-	-	-
GR5J-GR6J	-	-	-
GR4J-GR5J-GR6J	3,43e⁻¹³	3,24e⁻¹¹	3,24e⁻¹¹

Le test de Friedman qui ne s'applique que sur les bassins versants non corrélés, montre une différence significative entre les trois modèles, donc de bons résultats dans le cas des 605 bassins versants. Ce qui n'est pas toujours vrai pour autres bases de données.

Dans ce qui suit, nous présentons les boîtes à moustaches des critères appliqués sur les modèles hydrologiques pour les 22 bassins versants non corrélés (Figure 27), et un tableau qui contient les p-values correspondants aux tests statistiques (Tableau 10).

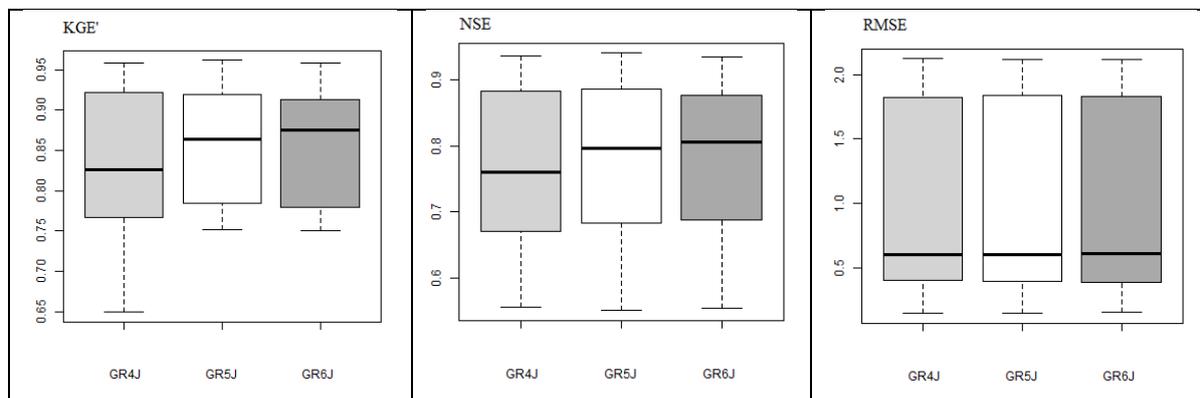


Figure 27 : Distributions de critères de performance en fonction des modèles pour 22 bassins versants.

Cette fois-ci, les modèles marquent une différence pour les critères KGE' et NSE, alors que pour le critère RMSE, les modèles sont un peu similaires.

Tableau 10 : Les p-values des tests statistiques de significativité pour les 22 bassins non corrélés

Test	Permutation			Cochran			Mann-Whitney		
	KGE'	NSE	RMSE	KGE'	NSE	RMSE	KGE'	NSE	RMSE
GR4J-GR5J	0,9267	0,9033	0,9817	-	-	-	0,799	0,05007	0,08541
GR4J-GR6J	0,8959	0,9097	0,9797	-	-	-	0,5449	0,5446	0,5446
GR5J-GR6J	0,9708	0,9932	0,9981	-	-	-	0,8736	0,799	0,6327
GR4J-GR5J-GR6J	0,991	0,9908	0,9996	0,2231	0,2231	-	-	-	-

Tableau 11 : Les p-values du test de Friedman pour les 22 bassins non corrélés

Test	Friedman		
	KGE'	NSE	RMSE
GR4J-GR5J	-	-	-
GR4J-GR6J	-	-	-
GR5J-GR6J	-	-	-
GR4J-GR5J-GR6J	0,9556	0,4216	0,4216

D'après les résultats obtenus, dans tous les cas, aucune différence entre les modèles. Cela peut être dû au fait que le nombre des bassins versants est très petit.

Conclusion

Nous avons montré dans ce stage que les tests statistiques de significativité sont des outils qui servent à étudier la différence entre les modèles hydrologiques. Pour cela, nous avons à disposition les données de débit observé Q_{obs} .

La surface, une caractéristique géométrique d'un bassin versant, joue un rôle important dans l'autocorrélation des débits d'un bassin, de sorte que plus le bassin est grand plus l'autocorrélation est forte. En ce qui concerne la corrélation, nous concluons que plus la distance entre les bassins qui traduit une différence importante de la géographie aux leurs alentours est grande, plus la corrélation des débits est faible entre eux. Quant à la normalité, nous avons trouvé que les critères des performances des modèles ne suivent pas la loi normale.

Ces résultats nous ont emmené à appliquer des tests statistiques non paramétriques pour étudier la différence entre les modèles hydrologiques GR4J, GR5J et GR6J à pas de temps journalier en fonction des critères d'évaluation KGE' , NSE et RMSE.

Nous concluons que les tests statistiques semblent confirmer nos conclusions « visuelles » (avec les boîtes à moustaches), le cas du test Q de Cochran, le test de Mann-Whitney et le test de Friedman mais que dans certains cas ils sont en désaccord entre eux, le test de permutation et le test de Friedman dans le cas des bassins non corrélés, et qu'il n'a pas été possible de comprendre pourquoi dans le cadre de ce stage.

Ce stage a permis de définir le cadre d'application des tests statistiques pour la comparaison de scores d'évaluation de modèles hydrologiques sur de larges échantillons de bassins versants. Le schéma (Figure 28) suivant résume la démarche de choix d'utilisation d'un test, que je propose :

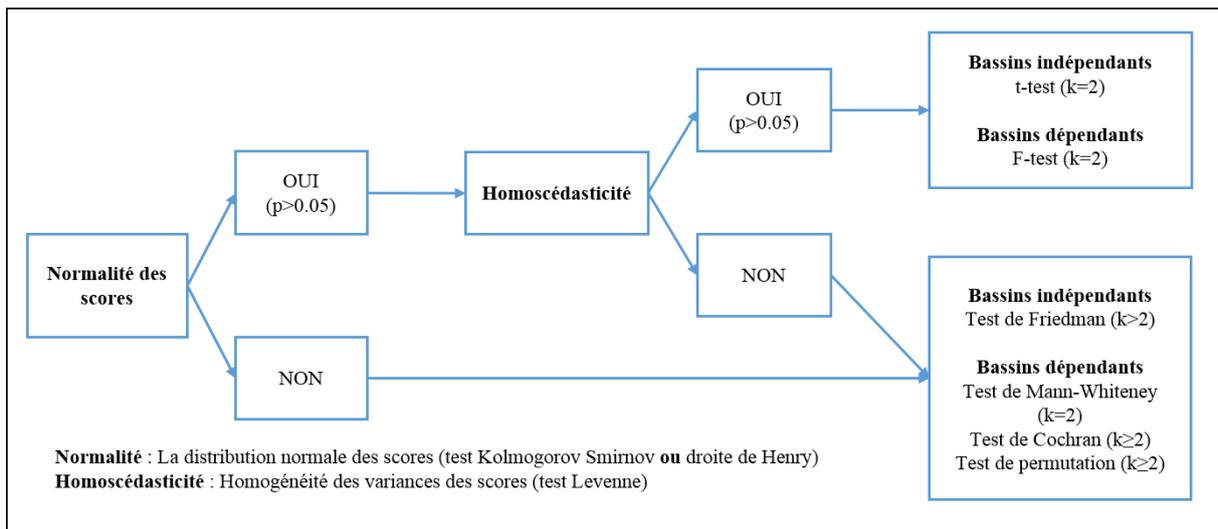


Figure 28 : Démarche d'application d'un test

Nous pouvons aussi utiliser des tests statistiques paramétriques, mais il faut penser à trouver des fonctions convenables pour normaliser les critères d'évaluations.

Finalement, le stage m'a permis vraiment d'approfondir mes connaissances en statistiques et ses applications en hydrologie, d'acquérir des compétences en programmation sur le logiciel R et de découvrir le monde de la recherche qui fait partie de mes ambitions futures.

Références bibliographiques

- Anctil, F., C. Perrin and V. Andréassian (2003). "Ann output updating of lumped conceptual rainfall/runoff forecasting models." *JAWRA Journal of the American Water Resources Association* **39**(5): 1269-1279.
- Andréassian, V., E. Parent and C. Michel (2003). "A distribution-free test to detect gradual changes in watershed behavior." *Water Resources Research* **39**(9): SWC101-SWC1011.
- Cochran, W. G. (1950). "The comparison of percentages in matched samples." *Biometrika* **37**(3/4): 256-266.
- Coron, L., C. Perrin, O. Delaigue, G. Thirel and V. Andréassian (in prep.). "airGR : A suite of lumped hydrological models in an R-package. Environmental Modelling and software." EGU General Assembly 2016.
- Coron, L., C. Perrin and C. Michel (2016). Suite of GR hydrological models for precipitation-runoff modelling. R package version 1.0.1
- Crochemore, L. (2016). Prévission saisonnière des débits pour la gestion de reservoirs. Thèse de doctorat, L'Institut des Sciences et Industries du Vivant et de l'Environnement.
- Ficchì, A., C. Perrin and V. Andréassian (2016). "Impact of temporal resolution of inputs on hydrological model performance: An analysis based on 2400 flood events." *Journal of Hydrology* **538**: 454-470.
- Friedman, M. (1937). "The use of ranks to avoid the assumption of normality implicit in the analysis of variance." *Journal of the American Statistical Association* **32**(200): 675-701.
- Gupta, H. V., H. Kling, K. K. Yilmaz and G. F. Martinez (2009). "Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling." *Journal of Hydrology* **377**(1-2): 80-91.
- Kampf, S. K. and S. J. Burges (2007). "A framework for classifying and comparing distributed hillslope and catchment hydrologic models." *Water Resources Research* **43**(5).
- Le Moine, N. (2008). 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, Université Pierre et Marie Curie.
- Nash, J. E. and J. V. Sutcliffe (1970). "River flow forecasting through conceptual models part I - A discussion of principles." *Journal of Hydrology* **10**(3): 282-290.
- Perrin, C., C. Michel and V. Andréassian (2003). "Improvement of a parsimonious model for streamflow simulation." *Journal of Hydrology* **279**(1-4): 275-289.
- Pushpalatha, R. (2013). Simulation et prévision des étiages sur des bassins versants français : Approche fondée sur la modélisation hydrologique. Thèse de doctorat, L'Institut des Sciences et Industries du Vivant et de l'Environnement.
- R Core Team (2014). "R: A language and environment for statistical computing." R Foundation for Statistical Computing, from URL <http://www.R-project.org/>.
- Rudolph, P. E. (1995). "Good, Ph.: Permutation tests. A practical guide to resampling methods for testing hypotheses." *Biometrical Journal* **37**(2): 150-150.

SANDRE, C. d. a. (2002). Référenciel hydrographique. M. c. d. l'Environnement: p. 49.

Sivapalan, M., G. Blöschl, L. Zhang and R. Vertessy (2003). "Downward approach to hydrological prediction." *Hydrological Processes* **17**(11): 2101-2111.

Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., and Soubeyroux, J.-M. (2010). "A 50-year high-resolution atmospheric reanalysis over France with the Safran system." *International Journal of Climatology*, **30** (11): pp. 1627-1644.

Thirel, G., F. Rousset-Regimbeau, E. Martin and F. Habets (2008). "On the impact of short-range meteorological forecasts for ensemble stream flow predictions." *Journal of Hydrometeorology* **9**(6): 1301-1317.