



HAL
open science

Détection et prise en compte d'éventuels impacts du changement climatique sur les extrêmes hydrologiques en France

Benjamin Renard

► **To cite this version:**

Benjamin Renard. Détection et prise en compte d'éventuels impacts du changement climatique sur les extrêmes hydrologiques en France. Sciences de l'environnement. Thèse de doctorat, INP Grenoble, 2006. Français. NNT: . tel-02588353

HAL Id: tel-02588353

<https://hal.inrae.fr/tel-02588353v1>

Submitted on 15 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.

- INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque
 |_|_|_|_|_|_|_|_|_|_|_|

T H E S E

pour obtenir le grade de

DOCTEUR DE L'INP Grenoble

Spécialité : Océan, atmosphère, hydrologie

préparée dans l'**Unité de Recherche Hydrologie-Hydraulique, Cemagref (Lyon)**

dans le cadre de l'**Ecole Doctorale Terre, Univers, Environnement**

présentée et soutenue publiquement

par

Benjamin RENARD

le 20 septembre 2006

**Détection et prise en compte d'éventuels impacts du
 changement climatique sur les extrêmes hydrologiques en
 France**

DIRECTEURS DE THESE

Michel LANG

Philippe BOIS

JURY

M. Serge PLANTON
 M. Lars GOTTSCHALK
 M. Philippe NAVEAU
 M. Michel LANG
 M. Philippe BOIS
 M. Rémy GARCON
 M. Eric PARENT

Météo-France, CNRM
 Université d'Oslo
 CNRS
 Cemagref
 INPG
 EDF-DTG, Grenoble
 ENGREF

Président
 Rapporteur
 Rapporteur
 Directeur de thèse
 Co-directeur de thèse
 Examineur
 Examineur

RESUME

Le changement climatique est une réalité largement reconnue aujourd'hui dans la communauté scientifique. Néanmoins, son impact sur les extrêmes hydrologiques est délicat à observer et à prévoir. Dans ce contexte, l'objectif de cette thèse est d'apporter des éléments de réponse aux trois questions suivantes : Comment détecter des changements dans les séries hydro-climatiques ? Quelles sont les évolutions observées sur les régimes extrêmes des cours d'eau en France métropolitaine ? Comment prendre en compte ces éventuels changements dans les méthodes de gestion des risques ? Chacun de ces trois objectifs sera traité à l'échelle locale et à l'échelle régionale.

Au niveau local, une analyse par simulation a permis de comparer un grand nombre de tests statistiques dédiés à la détection de tendances ou de ruptures, et de réaliser un schéma de sélection du test le plus approprié à la variable d'étude. Au niveau régional, deux aspects ont été pris en compte : la significativité et la cohérence régionale. Dans les deux cas, des méthodes sont proposées et comparées aux outils existant. L'analyse d'environ 200 stations hydrométriques en France ne montre pas de changement généralisé, ni en crue, ni en étiage, ni en régime. A plus petite échelle, certaines évolutions sont cependant détectées, mais les résultats ne suffisent pas à en certifier la cause : variabilité naturelle, évolution du bassin versant ou changement climatique. Enfin, les méthodes d'analyse fréquentielle ont été adaptées pour prendre en compte d'éventuelles évolutions temporelles. A cet effet, plusieurs outils statistiques se sont révélés particulièrement intéressants. Tout d'abord, l'analyse Bayésienne permet de considérer l'incertitude liée à la stationnarité. A l'échelle régionale, l'estimation des paramètres de modèles régionaux non stationnaires est possible grâce aux méthodes MCMC. Enfin, nous proposons une première piste, basée sur l'utilisation d'une copule Gaussienne, pour prendre en compte formellement la dépendance spatiale.

ABSTRACT

Climate change is widely considered as a reality by scientists. Nevertheless, impacts on hydrological extremes are more difficult to observe and to predict. The aim of this thesis is to answer the following questions: How to detect changes in hydro-climatic series? What are the observed changes for extreme discharges in France? How to take into account possible changes in frequency analysis? These objectives refer to both the local and the regional scales.

A number of at-site tests for change have been compared by simulations, leading to the selection of the more suitable method for a given variable. At the regional scale, both field significance and regional consistence are studied. New approaches are proposed and compared with the available methods. The study of almost 200 hydrometric stations does not show global patterns of change for high flows, low flows or regime. At a smaller scale, some consistent trends are detected, but without conclusion about their cause: natural variability, watershed evolution or climate change. Lastly, frequency analysis methods are modified in order to deal with possible non-stationarity. Several statistical tools were found to be suitable for this purpose. The Bayesian approach allows incorporating the uncertainty related to stationarity. At a regional scale, non-stationary models can be dealt with thanks to MCMC methods. Finally, a first approach for dealing with intersite dependence is proposed, based on a model for dependence arising from a Gaussian copula.

REMERCIEMENTS

Cette thèse n'aurait pu voir le jour sans le soutien scientifique de nombreuses personnes, auxquelles je tiens ici à exprimer ma gratitude.

Je remercie tout d'abord Michel Lang et Philippe Bois, mes directeurs de thèse, pour toutes les connaissances qu'ils m'ont transmises, pour la confiance et la liberté qu'ils m'ont accordées, et pour leur sympathie. Merci également à Michel pour l'aide qu'il continue à m'apporter aujourd'hui.

J'adresse un remerciement tout particulier à Aurélie Muller, ma "jumelle" de thèse, avec qui j'ai partagé durant trois ans beaucoup plus qu'un bureau: les problèmes statistiques ou informatiques, les discussions futiles ou métaphysiques, les doutes, les joies, les moments de folie, les pépito au goûter... Merci de m'avoir supporté pendant tout ce temps. Merci également à Matthieu Ribatet, qui a complété l'équipe de matheux du bureau du fond, et à qui je souhaite bonne chance pour la dernière année...

Je remercie également Vincent Garreta, qui a largement contribué aux études réalisées dans cette thèse grâce à la qualité de son travail de stage.

Durant ces trois années, j'ai bénéficié de l'aide et des précieux conseils d'un comité de pilotage, dont je souhaite remercier tous les membres: Olivier Mestre (Météo France), Anne Dupeyrat, Christel Prudhomme, Joël Gaillard, Emmanuel Paquet, Denis Aelbrecht, Carine Laurent, Sylvie Parey (EDF), Luc Neppel et Hélène Niel (Hydrosciences). Je n'oublie évidemment pas Eric Sauquet, qui fut au quotidien un précieux soutien scientifique (entre autres), malgré une fâcheuse tendance à sécher les réunions...

J'adresse également mes sincères remerciements à toutes les personnes du Cemagref de Lyon que j'ai eu la joie de côtoyer durant cette thèse. Je pense notamment à Jean-Michel Grésillon, directeur de l'U.R. Hydrologie-Hydraulique, Hélène Faurant, Anne Eicholz, Etienne Leblois, Christine Poulard, Gilles Galéa, Sébastien Proust, Alain Recking, Bernard Chastan, Denis Barbet, Eric Héroin, Isabelle Braud, et bien sûr l'incroyable Jean-Pierre.

Cette thèse n'aurait pu voir le jour sans données hydrométriques. Je remercie donc Christian Scherrer, responsable de la banque HYDRO, de nous avoir permis d'accéder aux données. Merci également à tous les gestionnaires de stations hydrométriques, qu'ils soient d'EDF, de la CNR ou des différentes DIREN, de m'avoir toujours reçu avec sympathie lors de mes visites.

Je remercie sincèrement les membres du jury d'avoir accepté d'évaluer mon travail: Serge Planton, directeur du jury, Lars Gottschalk et Philippe Naveau, rapporteurs, ainsi que Rémy Garçon et Eric Parent, examinateurs.

Cette thèse a été co-financée par le Cemagref et par EDF. Je suis donc reconnaissant aux deux organismes pour leur soutien et l'intérêt qu'ils ont manifesté pour ce travail.

Heureusement, trois années de thèse, ce n'est pas uniquement du travail, mais aussi des week-end, des vacances, des voyages, des apéros, des concerts, des mariages, des naissances... Je remercie donc tous mes amis avec qui j'ai partagé ces moments indispensables.

Etant donné que j'écris ces remerciements confortablement installé dans un hamac à Lamentin, Guadeloupe, je vais commencer par mes amis d'outre-mer, et notamment mon jumeau et maître Mitch et sa *doudou* Caroline, ainsi que Cécile, David, Jules, et leur prochain *ti mal*... Je pense également à tous ceux que j'ai eu le bonheur de rencontrer au cours des

deux merveilleuses années que j'ai passées ici, notamment Flo, Maguy, Xav, Mumu, Benj, David, Manou, Nico, Marco, Cécile, Laetitia, Pauline, Isa, Tramby, Fred... Merci également à Nicolas Sauvion, Claude Welcker et Nathalie Boissot, avec qui j'ai effectué mes premiers pas dans le monde de la recherche.

Je remercie également mes amis du Cemagref pour le plaisir que j'ai eu à les côtoyer quotidiennement, ainsi que pour le beau témoignage d'amitié qu'ils m'ont offert après la soutenance... Merci donc à la cartoonesque Elo, la Goldmanesque Anne-Laure, l'étonnante Aurélie, le grognon Eric, Magalitex, Otman le sage, la joyeuse Céline, Fabien le Lyonnais, Guillaume & Caro les Stéphanois, le moustachu Raouf et la douce Delphine, mon humoriste préféré Jean-Phi, la théâtrale Sandra, Jérôme l'amateur d'art abstrait, la puissante Flora, Kamal le tendre, le broussailleux Jean-Marie, les jeunes mariés Judicaël et Marie-Liesse, le cassant Olivier, la skieuse Sandya, Matthieu l'original, la pointeuse Karine, la dangereuse Adeline, le Zouk-lover Matthieu, le Zidanesque Taha, Arnaud le Président, Michel qui n'y connaît rien en foot, Stéphane le fêtard, Loïc le pongiste, Tabernacle Emmanuelle, l'argentine Elodie, le très grand Oldrich, Fred Desproges, Ju et Pauline les voyageurs, Sandy la belle vie, le bavard Manu, les Martiniquais Noémie et Jean-Guillaume, l'historique Robin, obrigado Maria-Hélène, le sportif Matthias, Aurélie the fatal picarde, « real » Carlos, et la plongeante Charlotte.

Je remercie également mes colocataires successifs qui ont partagé avec moi les rudesses des hivers Lyonnais, à commencer par mon Frangin Ben, puis Eugenia, Maxime, Mina, et aujourd'hui Déric Aaaaaa et Doc Simon.

Le torse bombé et le regard fier, je salue respectueusement tous les citoyens de la glorieuse république Zubrovskaienne, à commencer par notre futur ex-président Andouine, et tous les ministres de son gouvernement, re-Déric, Clochetilde, U, Guimtz, Titiou, Bret, Dove, Elotchie, Pat, Nico, Armelle, Ju, Pauline & Julien... & Luis, la Grosse Guillemette, Robert, Papa, Chourles & Agathe, Ben & Steph & Léonard, Marina, Pisha, Matou, Olive & Fred, Tom Cruise, John, Aurélie. En attendant le putsch, je vous adresse à tous mes *dermoutch* les plus sincères.

Merci à Jo le coq et au Gitan pour les bons moments partagés aux quatre coins du monde. Seuwfin'.

Merci à Mo et Charles pour leur amitié, et mes hommages à tous les membres et sympathisants du YFBC.

Merci à mes compagnons musicaux, membres de la Mofo Monk Foundation, de la Lyrical Propagande, d'H.O.S.T., de Chacha Bouna Blast, de Quadricolor ou autres, à savoir Doc, Kéké, Cissou, Thomas, Manu, Dalil, Hervé, K-Mal, Ali, Charly, Yves, Anto, Raph, Olivier, Pierre.

Enfin, je remercie toute ma famille pour sa présence et son soutien depuis plus de 28 ans, plus particulièrement mes parents géniaux, pour tout, mon artiste de frère à qui je souhaite le meilleur pour les années à venir, et mes grands-parents, qui m'ont logé, nourri et supporté durant de nombreux week-end.

L'épisode de la thèse étant à présent terminé, je repars à l'aventure ma serviette autour du cou, en vous disant à tous...

So long... and thank's for all the fish!

CHAPITRES

CHAPITRE 1. Eléments de contexte	17
I. Les débits extrêmes	18
II. Le contexte du changement climatique global	25
CHAPITRE 2. Données et échantillons	53
I. Chroniques de débits journaliers	54
II. Choix de variables descriptives	69
CHAPITRE 3. Comment détecter des changements ?	93
I. Echelle locale : quel test pour quelle variable ?	95
II. Significativité régionale	125
III. Détection de changements régionalement cohérents	133
CHAPITRE 4. Analyse de stationnarité en France Métropolitaine	147
I. Deux exemples commentés	148
II. Analyse locale	154
III. Analyse régionale	167
IV. Lien avec des variables explicatives	180
CHAPITRE 5. La gestion du risque en contexte non stationnaire	199
I. La notion de période de retour dans un contexte non stationnaire	201
II. Echelle locale : l'analyse Bayésienne	206
III. Analyse régionale non stationnaire	220
CHAPITRE 6. Conclusions et perspectives	245
I. Acquis pratiques	246
II. Acquis méthodologiques	247
III. Perspectives	250

TABLE DES MATIERES

INTRODUCTION.....	15
CHAPITRE 1. ELEMENTS DE CONTEXTE	17
I. LES DEBITS EXTREMES	18
I.1. Généralités sur l'hydrologie en France.....	18
I.1.1. L'eau, ressource et risque.....	18
I.1.2. Variabilité spatio-temporelle des écoulements	19
I.1.3. Crues et étiages en France	20
I.2. Débits extrêmes et gestion du risque : l'apport des probabilités.....	21
I.2.1. Gestion préventive du risque : aléa et vulnérabilité	22
I.2.2. Les probabilités comme un outil de gestion du risque	23
II. LE CONTEXTE DU CHANGEMENT CLIMATIQUE GLOBAL	25
II.1. Qu'est-ce que le changement climatique ?	25
II.1.1. Quelques notions de climatologie.....	25
II.1.2. Variabilité temporelle du climat	27
II.1.3. L'impact des activités humaines	28
II.2. Changement climatique et incertitudes.....	30
II.2.1. Fonctionnement d'un modèle de circulation générale (GCM)	30
II.2.2. Les incertitudes dans la modélisation du climat	31
II.3. Ce que l'on a d'ores et déjà observé	32
II.3.1. Températures et cryosphère	33
II.3.2. Précipitations et écoulements.....	35
II.4. Ce que l'on attend	38
II.4.1. Températures.....	38
II.4.2. Précipitations et écoulements.....	40
II.5. Le cas de la France métropolitaine.....	43
II.5.1. Impacts observés	43
II.5.2. Prévisions pour le XXI ^e siècle	46
CHAPITRE 2. DONNEES ET ECHANTILLONS.....	53
I. CHRONIQUES DE DEBITS JOURNALIERS	54
I.1. Les données hydrométriques en France	54
I.1.1. La mesure des débits	54
I.1.2. Quelques repères historiques.....	55

I.1.3. Critères de sélection des stations.....	57
I.2. Description des données utilisées.....	58
I.2.1. Description générale.....	58
I.2.2. Régions hydro-climatiques.....	60
II. CHOIX DE VARIABLES DESCRIPTIVES	69
II.1. Variables de crues	69
II.1.1. Définition des variables	69
II.1.2. Algorithme d'extraction de variables.....	74
II.1.3. Critères d'indépendance.....	75
II.1.4. Calcul d'une durée caractéristique de crue	77
II.1.5. Cas des mélanges de population (régime pluvio-nival).....	80
II.1.6. Procédure complète d'extraction	83
II.1.7. Traitement particulier des crues nivales	84
II.2. Variables descriptives des basses eaux.....	85
II.2.1. Définition des variables	86
II.2.2. Choix d'un seuil de bas débit.....	86
II.2.3. Analyse multi-durées	87
II.3. Variables descriptives du régime des eaux	89
CHAPITRE 3. COMMENT DETECTER DES CHANGEMENTS ?.....	93
I. ECHELLE LOCALE : QUEL TEST POUR QUELLE VARIABLE ?	95
I.1. Présentation des principaux tests	95
I.1.1. Rappel : le principe des tests statistiques	95
I.1.2. Quatre tests classiques.....	95
I.1.3. Autres tests	98
I.1.4. Le test du rapport des vraisemblances maximales	99
I.1.5. Approximation de la loi du rapport des vraisemblances maximales Z_n dans le cas du test LR-rupture.....	100
I.1.6. Quelques modifications du test LR-rupture	102
I.2. Etude de robustesse	104
I.2.1. Robustesse aux petits échantillons	104
I.2.2. Erreur sur la distribution	105
I.2.3. Présence d'autocorrélation	107
I.3. Etude de puissance.....	115
I.3.1. Comparaison paramétrique (P) / non paramétrique (NP).....	115
I.3.2. Recherche des tests les plus puissants.....	117
I.3.3. Influence de la position d'une rupture.....	119
I.3.4. Le cas des données autocorrélées.....	120
I.4. Conclusion sur les méthodes locales de détection	121

II. SIGNIFICATIVITE REGIONALE.....	125
II.1. Quelques considérations théoriques	125
II.2. Les différentes méthodes disponibles	126
II.2.1. Utilisation du nombre de stations équivalent.....	126
II.2.2. Le Bootstrap par année	127
II.2.3. Utilisation d'une copule Gaussienne	127
II.3. Etude par simulation.....	127
II.3.1. Simulation à partir d'une loi normale multivariée.....	128
II.3.2. Simulation par une copule Gaussienne	129
II.3.3. Simulation par un générateur de pluies.....	130
II.4. Conclusion sur la significativité régionale	132
III. DETECTION DE CHANGEMENTS REGIONALEMENT COHERENTS.....	133
III.1. Principales approches utilisées	133
III.1.1. Approches descriptives.....	133
III.1.2. Création de variables régionales.....	134
III.1.3. Création de statistiques régionales	135
III.2. Modèles régionaux non stationnaires	136
III.2.1. Une approche semi-paramétrique.....	137
III.2.2. Approches paramétriques	140
III.3. Conclusion sur la cohérence régionale.....	143
CHAPITRE 4. ANALYSE DE STATIONNARITE EN FRANCE	
METROPOLITAINE.....	147

I. DEUX EXEMPLES COMMENTES	148
I.1. Mise en forme et diffusion des résultats	148
I.2. La Seine à Paris-Austerlitz	150
I.3. La Neste de Cap de Long à Aragnouet.....	152
II. ANALYSE LOCALE	154
II.1. Première analyse sur l'échantillon complet de stations.....	154
II.1.1. Résultats en crues.....	154
II.1.2. Résultats en étiages	156
II.1.3. Résultats en régime	157
II.2. Critique des résultats	158
II.2.1. Organisation de la critique.....	158
II.2.2. Causes supposées des ruptures observées.....	158
II.3. Résultats sur les jeux réduits de stations.....	160
II.3.1. Crues	160
II.3.2. Crues nivales.....	162

II.3.3. Etiages.....	164
II.3.4. Régime.....	166
III. ANALYSE REGIONALE.....	167
III.1. Significativité régionale.....	167
III.2. Cohérence régionale.....	169
III.2.1. Vérification des hypothèses du test LR-Reg.....	169
III.2.2. Crues pluviales.....	170
III.2.3. Crues nivales.....	171
III.2.4. Etiages.....	171
III.2.5. Régime.....	172
III.2.6. Récapitulatif des évolutions régionalement cohérentes.....	173
IV. LIEN AVEC DES VARIABLES EXPLICATIVES.....	180
IV.1. Comparaison qualitative avec les résultats du projet IMFREX.....	180
IV.1.1. Crues pluviales.....	180
IV.1.2. Crues nivales.....	182
IV.1.3. Etiages.....	183
IV.1.4. Régime.....	184
IV.2. Lien avec les pluies et les températures <i>via</i> la modélisation hydrologique.....	185
IV.2.1. Crues pluviales du Nord-Est.....	185
IV.2.2. Crues nivales des Alpes.....	190
CHAPITRE 5. LA GESTION DU RISQUE EN CONTEXTE NON STATIONNAIRE.....	199
<hr/>	
I. LA NOTION DE PERIODE DE RETOUR DANS UN CONTEXTE NON STATIONNAIRE	201
I.1. Données MAXAN.....	201
I.2. Données SUP-SEUIL.....	203
II. ECHELLE LOCALE : L'ANALYSE BAYESIENNE.....	206
II.1. Principe général.....	206
II.1.1. La stationnarité comme une incertitude.....	206
II.1.2. <i>A priori</i> et <i>a posteriori</i>	208
II.1.3. Difficultés numériques.....	209
II.2. Exemple de mise en œuvre pour l'analyse des crues.....	210
II.2.1. Les modèles utilisés.....	210
II.2.2. Spécification des <i>a priori</i>	211
II.2.3. Utilisation de la distribution <i>a posteriori</i>	212
II.2.4. Cas d'étude.....	214
III. ANALYSE REGIONALE NON STATIONNAIRE.....	220

III.1. Comment passer à l'échelle régionale ?	220
III.1.1. Principe général	220
III.1.2. Mise en œuvre dans un cadre non stationnaire.....	221
III.2. Cas d'étude	222
III.2.1. Données et modèles	222
III.2.2. Méthodes MCMC	225
III.2.3. Distribution <i>a posteriori</i>	226
III.2.4. Validation du modèle.....	228
III.2.5. Analyse fréquentielle.....	230
III.3. Le problème de la dépendance spatiale	230
III.3.1. Quels sont les dangers d'ignorer la dépendance spatiale ?.....	230
III.3.2. La copule Gaussienne	231
III.3.3. Exemple sur les pluies	232
III.3.4. Discussion.....	241

CHAPITRE 6. CONCLUSIONS ET PERSPECTIVES245

I. ACQUIS PRATIQUES	246
I.1. Critique de longues séries de débits journaliers	246
I.2. Réflexions sur les techniques d'échantillonnage	246
I.3. Impact du changement climatique sur les crues et les étiages	246
II. ACQUIS METHODOLOGIQUES	247
II.1. Détection de ruptures ou de tendances	247
II.2. Solutions pour la gestion du risque en contexte non stationnaire	248
II.3. Mise en place d'outils statistiques dédiés à l'hydrologie	249
III. PERSPECTIVES	250
III.1. Organisation de la gestion des données hydrométriques en France	250
III.2. Vers une meilleure prise en compte de la variabilité spatio-temporelle des événements extrêmes	251

BIBLIOGRAPHIE255

ANNEXES269

ARTICLE 1 : ANALYSE BAYESIENNE NON STATIONNAIRE271

ARTICLE 2 : METHODES MCMC301

ARTICLE 3 : UTILISATION DE COPULES333

INTRODUCTION

« De toute évidence, le climat de la Terre a évolué à l'échelle régionale et mondiale depuis l'époque préindustrielle ».

Ce constat émane du dernier rapport de synthèse du groupe d'experts intergouvernemental sur l'évolution du climat (IPCC, 2001). Ce qui n'était encore qu'une hypothèse il y a une décennie semble aujourd'hui être un fait avéré : l'augmentation évidente de la concentration des gaz à effet de serre dans l'atmosphère a eu des répercussions sur un certain nombre de variables climatiques. La température mondiale moyenne a ainsi augmenté d'environ $0.6\text{ }^{\circ}\text{C}$ ($\pm 0.2\text{ }^{\circ}\text{C}$) depuis la fin du XIX^e siècle, entraînant une augmentation du nombre de jours chauds et une diminution du nombre de jours de gel. Les précipitations seraient elles aussi en augmentation, au moins dans l'hémisphère Nord, tandis que la couverture neigeuse et l'étendue des glaciers non polaires seraient en régression prononcée. Au delà de ces observations directes, les modèles de circulation générale (GCM) prévoient tous une persistance de ces tendances au cours du siècle. Les incertitudes qui pèsent sur ces projections sont très importantes, car les modèles sont conditionnés par des scénarios sociaux-économiques imaginés pour les années à venir. De plus, les résultats peuvent varier significativement d'un GCM à l'autre, et un certain nombre de phénomènes, agissant à des échelles intermédiaires, sont pris en compte de façon très simplifiée dans les codes de calculs (dont la résolution spatiale est assez grossière) et restent parfois largement méconnus. Ainsi, à l'horizon 2100, l'augmentation de la température globale devrait se situer entre 1.4 et $5.8\text{ }^{\circ}\text{C}$.

Ces perturbations climatiques devraient avoir des répercussions sur le régime hydrologique. Cependant, ces conséquences sont difficiles à évaluer, pour deux principales raisons : tout d'abord, un réchauffement global peut avoir des effets antagonistes sur les écoulements, en augmentant à la fois l'évapotranspiration et les précipitations. L'influence de ces deux facteurs diffère également suivant le régime, nival ou pluvial, du cours d'eau. De plus, les échelles spatiales et temporelles utilisables en sortie des GCM ne concordent pas avec celles des modèles hydrologiques pluie/débit. Toutefois, les études publiées sur ce sujet (Zwiers et Kharin, 1998; Kharin et Zwiers, 2000; Voss *et al.*, 2002) prédisent une aggravation des phénomènes extrêmes, en basses eaux comme en hautes eaux. A ce jour, aucun signal cohérent à grande échelle n'a cependant été détecté pour les écoulements, malgré des changements avérés à l'échelle régionale.

Parallèlement au contexte du changement global, la vulnérabilité face aux événements hydrologiques extrêmes est en constante augmentation, notamment en France. Une urbanisation mal maîtrisée a conduit la population à être de plus en plus exposée au risque d'inondation, tandis que la ressource en eau est de plus en plus sollicitée pour les besoins de l'agriculture et de l'industrie. Ces dernières années, la multiplication et la médiatisation de tels événements, comme les inondations de décembre 2003 ou la canicule de l'été 2003, ont conduit à la perception d'une recrudescence des événements extrêmes en France, qui n'est à ce jour pas établie.

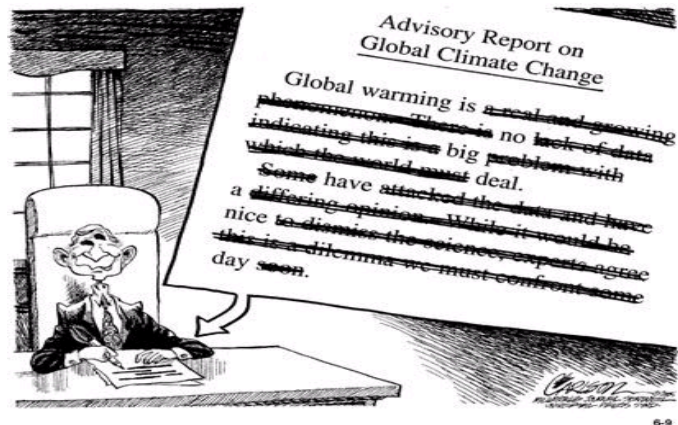
L'absence de réponse claire de la part des scientifiques quant à la réalité d'un changement climatique affectant les événements extrêmes est en partie due à la difficulté méthodologique de détecter des tendances dans ce type de séries hydro-climatiques qui, par nature, sont soumises à une importante variabilité naturelle. Il est toutefois primordial de pouvoir répondre à cette interrogation, car en matière de prévention des risques, la stationnarité des phénomènes dans le temps est une hypothèse qui est presque systématique.

Cette thèse poursuit donc trois objectifs majeurs. Tout d'abord, nous souhaitons progresser sur la **méthodologie de détection de changements** dans des séries hydro-météorologiques qui, bien que très fournie, reste perfectible. Ensuite, nous espérons apporter une réponse quant à l'**évolution du régime hydrologique extrême en France**. Enfin, nous proposerons des outils de **gestion des risques dans un contexte non stationnaire**.

Les éléments de réponse apportés au cours de ces trois années de thèse sont regroupés dans ce mémoire, qui est organisé en six parties. Les deux premiers chapitres ont pour but de mettre en place tous les outils nécessaires à la réalisation du travail, et ne contiennent donc pas de résultats originaux. Le **premier chapitre** rappelle quelques éléments de contexte, ainsi que les principales règles qui gouvernent les événements hydro-climatiques. La **seconde partie** est dédiée à la présentation des données de débits qui seront utilisées par la suite. Les trois chapitres suivants décrivent les principaux résultats obtenus au cours de la thèse. Le **chapitre trois** est dédié aux méthodes de détection de changements. Au niveau local, une étude bibliographique associée à des simulations Monte Carlo permettront de comparer différents tests statistiques et de mettre en place un schéma de sélection du test le plus approprié aux caractéristiques probabilistes de la variable d'étude. Au niveau régional, nous aborderons deux aspects qui ne sont que rarement évalués à l'heure actuelle dans les études de stationnarité : la significativité et la cohérence régionales. L'analyse de stationnarité des débits extrêmes en France est décrite dans le **chapitre quatre**. L'étude concernera encore une fois les échelles locales et régionales, et nous tenterons de faire la part des évolutions qui peuvent être imputées au climat de celles qui sont d'origines anthropiques directes (aménagement de bassin, problèmes météorologiques,...). Quelques pistes pour l'analyse fréquentielle en contexte non stationnaire seront exposées au **chapitre cinq**. La mise à jour de cette famille de méthodes pour permettre la prise en compte d'évolutions temporelles a été réalisée avec deux objectifs sous-jacents. Le premier est de considérer la stationnarité des séries non pas comme une hypothèse à accepter ou à rejeter, mais plutôt comme une incertitude à quantifier. Le second concerne l'échelle régionale, et vise à prendre en compte la dépendance spatiale entre les stations afin de fournir une estimation plus réaliste des incertitudes. Enfin, le **chapitre six** résumera les principales conclusions du travail effectué, et exposera quelques perspectives.

Afin de répondre aux objectifs initiaux, plusieurs outils statistiques ont été utilisés, notamment l'analyse Bayésienne, les méthodes de Monte Carlo par Chaînes de Markov (MCMC) et les copules. Ces méthodes, bien connues des statisticiens, connaissent actuellement un succès grandissant dans la communauté hydrologique. Cette thèse n'étant pas une thèse de statistiques, nous avons préféré regrouper les résultats par objectif hydrologique plutôt que par outil méthodologique. En conséquence, ces méthodes transversales ne sont pas décrites en détails dans le corps du document. Elles ont cependant fait l'objet de publications qui figurent en **annexes**, et qui fournissent toutes les informations nécessaires à leur utilisation.

CHAPITRE 1. ELEMENTS DE CONTEXTE



Le travail présenté dans cette thèse se positionne à la croisée de trois disciplines scientifiques : l'hydrologie, la climatologie et les statistiques appliquées. Bien qu'une grande partie du travail réalisé soit de nature méthodologique, il est capital de ne pas faire abstraction du contexte physique. Comprendre les mécanismes généraux des phénomènes étudiés, ainsi que la nature des données (et leurs incertitudes) est un préalable indispensable à toute analyse. Le but de ce premier chapitre est donc de présenter le contexte physique du travail réalisé. Cette présentation ne se veut pas exhaustive, mais prétend seulement décrire les principaux processus gouvernant les phénomènes climatologiques et hydrologiques.

Dans une première partie, nous exposons quelques fondements de l'hydrologie. Nous nous attachons tout d'abord à décrire la variabilité spatio-temporelle des écoulements, avant d'insister plus particulièrement sur les événements extrêmes. Du point de vue physique, les

processus conduisant à la genèse des crues et des étiages en France seront décrits. Du point de vue de la gestion du risque, les concepts de vulnérabilité et d'aléa seront détaillés. Nous exposerons enfin quelques résultats issus de la théorie des valeurs extrêmes, qui sont très couramment utilisés en pratique dans le cadre de la prédétermination de l'aléa hydrologique.

La seconde partie est consacrée au contexte du changement climatique. Nous exposons tout d'abord le fonctionnement général du climat sur Terre, avant d'aborder l'impact des activités humaines. Les modèles de circulation générale (GCM) seront ensuite brièvement décrits, avec une attention particulière sur la chaîne d'incertitudes qui apparaît lorsque l'on cherche à quantifier le changement climatique. Une revue bibliographique non exhaustive permettra ensuite de décrire les évolutions d'ores et déjà observées, puis celles qui sont attendues dans le futur. Cette revue concernera à la fois l'échelle globale et l'échelle de la France métropolitaine.

I. Les débits extrêmes

Cette première section est dédiée à la présentation des extrêmes hydrométriques. Une description de leur variabilité et leur genèse est tout d'abord effectuée. Nous décrirons ensuite les méthodes utilisées pour gérer les risques associés.

I.1. Généralités sur l'hydrologie en France

I.1.1. L'eau, ressource et risque

L'eau est un élément indispensable à toute forme de vie terrestre. L'homme ne fait pas exception à cette règle, et a donc de tout temps placé la question de l'approvisionnement en eau au centre de sa lutte pour la survie. Au fur et à mesure de son développement technologique, l'humanité a découvert de nouvelles utilisations possibles, pour l'agriculture, le transport et l'industrie notamment. Ainsi, la plupart des grandes civilisations de l'Antiquité se sont développées autour de grands cours d'eau, et aujourd'hui encore, l'accès à l'eau reste une préoccupation géopolitique majeure.

L'élément liquide est pourtant présent en très grande quantité sur notre planète, mais sa disponibilité est plus problématique : environ 97% de l'hydrosphère est en effet constituée d'eau salée, et sur la quantité restante, la plus grande partie est stockée sous forme de glace. Au final, on estime que l'eau contenue dans les réseaux hydrographiques (lacs et cours d'eau) ne représente qu'environ 0.33% de l'eau douce disponible. L'eau est donc avant tout une ressource, et est abondamment étudiée en tant que telle par les hydrologues. Des problématiques telles que la cohabitation des différents usages, la qualité, les aménagements, l'impact des activités humaines ou du réchauffement climatique imposent en effet de comprendre le fonctionnement général des eaux continentales.

Cette proximité et cette dépendance avec la ressource en eau comportent cependant un certain nombre de risques, liés aux événements extrêmes affectant épisodiquement les cours d'eau. Deux phénomènes bien différents, tant au niveau de leur genèse physique que de leurs impacts, doivent être discernés. Tout d'abord, les étiages correspondent à des écoulements anormalement faibles. Les impacts de ces périodes de basses eaux sont multiples. En France, on peut notamment citer les problèmes écologiques liés au biotope ou à la qualité des eaux, les problèmes agricoles liés à la limitation de l'irrigation, le problème de la sécurité de certains ouvrages (refroidissement des centrales nucléaires), voire la difficulté d'approvisionnement en eau potable. Dans les cas les plus graves, des périodes d'étiages prononcées et récurrentes peut avoir des conséquences bien plus dramatiques (Afrique Sahélienne). A l'inverse, le phénomène de crue correspond à une période d'écoulements

anormalement élevés. Lorsque la quantité d'eau dépasse la capacité d'écoulement du lit mineur, l'eau s'écoule dans le lit majeur, provoquant ainsi des phénomènes d'inondations. Bien que généralement plus localisées que les étiages, les crues ont en France des conséquences plus importantes au niveau des pertes de vies humaines. Les conséquences économiques peuvent également être désastreuses, particulièrement pour les bassins versants les plus anthropisés.

I.1.2. Variabilité spatio-temporelle des écoulements

Suivant l'échelle à laquelle on les étudie, les écoulements possèdent des caractéristiques différentes (Figure 1-1). En un point particulier du réseau hydrographique, on peut ainsi s'intéresser aux caractéristiques inter-annuelles des débits, comme par exemple le module, qui est le débit moyen calculé sur toute la durée de données disponible. Cette valeur cache une première variabilité, liée à la répartition non uniforme des débits au cours des douze mois de l'année. Les débits mensuels inter-annuels peuvent ainsi être calculés, et expriment la saisonnalité des écoulements, qui sert de base à la classification des cours d'eau en régimes (Pardé, 1955). Le module peut également être calculé année par année, et on peut alors observer la variabilité inter-annuelle de la quantité d'eau écoulee annuellement. En une année particulière, la répartition mensuelle des écoulements peut également s'éloigner du régime normal de la rivière. Enfin, en observant les débits à un pas de temps plus fin (ici, la journée), une série chronologique complexe apparaît, révélant des variations parfois brusques sur des faibles échelles temporelles. C'est également à cette échelle que l'on voit apparaître les phénomènes extrêmes auxquels nous allons nous intéresser. Signalons que cette multiplicité d'échelle a conduit à une approche qui connaît aujourd'hui un succès grandissant en hydrologie : l'approche multi-fractale, qui étudie la stabilité de certaines caractéristiques au travers des différentes échelles d'observation possibles (Schertzer *et al.*, 2006).

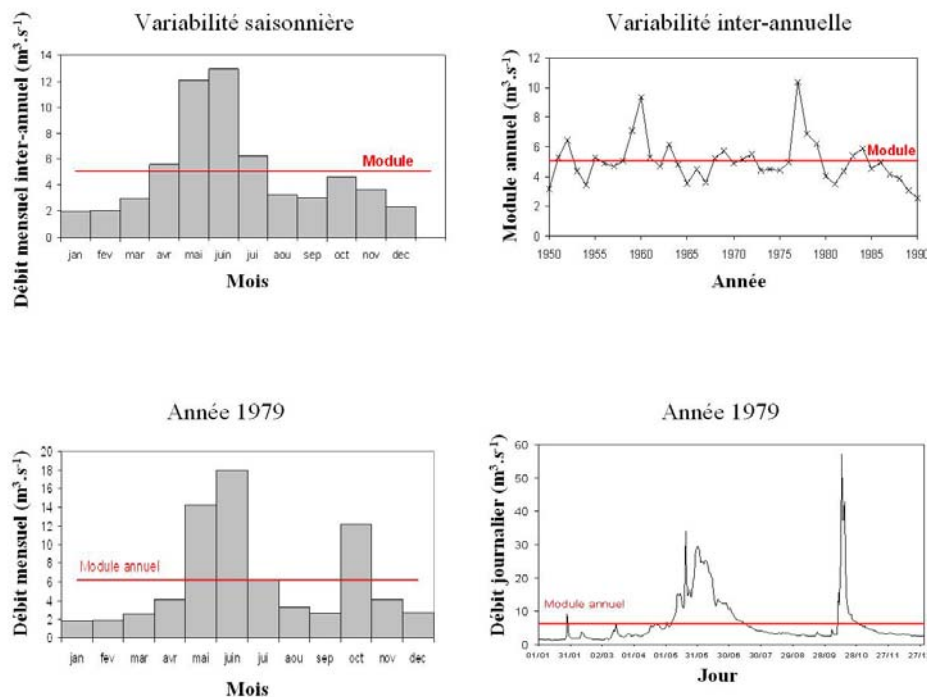


Figure 1-1. Variabilité des écoulements de la Tinée au pont de la Belloire suivant l'échelle temporelle étudiée.

La variabilité spatiale des écoulements est elle assez particulière, puisqu'elle est conditionnée par la structuration du réseau hydrographique. Cependant, la plupart des processus influant sur les écoulements (précipitations, températures, nappes) sont continus dans l'espace, et cette continuité se retrouve en partie dans la variabilité spatiale des débits. La répartition des différents régimes sur le territoire français en constitue un exemple. Les régimes pluviaux océaniques, caractérisés par une saison de hautes eaux hivernales et des basses eaux estivales, sont ainsi répartis sur la partie nord et l'ouest du pays. Les régimes pluviaux méditerranéens, qui connaissent des étiages plus marqués en été et des précipitations automnales importantes, sont répartis dans la partie sud-est du pays. Enfin, les régimes nivaux et glaciaires se retrouvent sur les zones de reliefs marqués. Evidemment, du fait de la continuité amont-aval, des mélanges de ces régimes de base peuvent être observés en un point du réseau hydrographique. Cependant, lorsque l'on s'intéresse à un pas de temps plus fin, la variabilité spatiale des débits devient beaucoup plus difficile à appréhender. En période de basses eaux, une certaine continuité peut toujours être observée, puisque les processus à l'origine des étiages (absence de précipitations ou températures élevées) ont une portée spatiale importante. Par contre, les crues sont généralement des phénomènes beaucoup plus localisés, puisque la grande variabilité spatiale des précipitations peut créer des réponses bien différentes entre deux bassins versants proches. Ceci est particulièrement vrai dans le cas des précipitations convectives à l'origine des plus fortes crues en zone méditerranéenne. Des exceptions existent cependant, et des crues généralisées sur une grande échelle spatiale sont également parfois observées (grandes crues de 1856 en Europe).

I.1.3. Crues et étiages en France

Un phénomène hydrologique extrême peut être considéré comme la réponse du bassin versant à un ensemble de conditions particulières affectant les processus influant sur les écoulements. On peut ainsi rechercher les principales causes de ces phénomènes, et en dresser une typologie.

Ainsi, la principale cause des étiages en France est un déficit de précipitations liquides sur une durée qui dépendra des caractéristiques intrinsèques du bassin. Les écoulements sont alors le résultat de la vidange de la nappe, qui pourra atteindre des niveaux critiques en l'absence d'une recharge significative. Ainsi, pour des cours d'eau de plaine possédant une nappe d'accompagnement importante, l'étiage pourra résulter d'un déficit des cumuls pluviométriques sur une durée de plusieurs mois, voire plus. Inversement, des rivières où la composante souterraine est moins importante réagiront beaucoup plus rapidement à l'absence de précipitations. La présence de fortes températures est également un facteur aggravant, puisque l'augmentation de l'évapotranspiration peut contribuer significativement à la baisse des débits. Ainsi, pour les cours d'eau à régime pluvial, les étiages s'observent généralement en été, période au cours de laquelle les cumuls pluviométriques sont les plus faibles et les températures sont les plus élevées. Dans le cas des rivières avec une composante nivale, le déterminisme des étiages dépendra de l'altitude du bassin versant. Pour les bassins versants les plus élevés, la contribution de l'onde de fonte nivale est généralement suffisante pour assurer des débits relativement élevés en été. L'étiage se situera donc en hiver, période à laquelle les précipitations sont stockées sous forme de neige. A des altitudes plus modérées, les deux phénomènes précédents peuvent parfois coexister. On observe alors une saison d'étiage estivale, durant laquelle la contribution nivale n'est plus suffisante pour compenser l'absence de pluie et les températures élevées, et une saison d'étiage hivernale, liée au stockage des précipitations.

Les crues sont quant à elles le résultat de fortes précipitations sur tout ou partie du bassin versant. Le type de précipitations générant les crues les plus importantes dépend encore une fois des caractéristiques du bassin versant. Le cas le plus particulier concerne les cours d'eau

d'altitude à régime nival. Dans ce cas, les crues résultent de la fonte des précipitations hivernales stockées sous forme de neige, déclenchée par l'augmentation des températures au printemps. Ces crues sont généralement très étendues dans le temps, et leur volume dépendra principalement des cumuls hivernaux. Des épisodes pluvieux peuvent se superposer à cette onde de fonte, ajoutant ainsi une part de ruissellement au débit de base important généré par la fonte. En l'absence de composante nivale, on distingue deux types de ruissellement à l'origine de l'augmentation des débits. Le premier, nommé le ruissellement Hortonien, est lié au dépassement de la capacité d'infiltration des sols du fait de l'intensité des précipitations. La part d'eau qui ne peut s'infiltrer ruisselle donc sur le bassin versant jusqu'à retrouver le réseau hydrographique et, *in fine*, l'exutoire. Il s'agit d'un ruissellement rapide, que l'on observe par exemple sur les bassins cévenols, ou dans les zones urbaines, qui sont fortement imperméabilisées et ont donc une capacité d'infiltration quasi nulle. Le second type de ruissellement est lié à la saturation du sol du fait des précipitations antérieures. On assiste alors à un affleurement de la nappe, et toute l'eau ultérieure reçue par ce profil de sol saturé ne pourra que ruisseler. La gravité de la crue sera donc principalement liée au cumul des précipitations plutôt qu'à leur intensité. Précisons que ces deux phénomènes coexistent souvent au sein d'un même bassin versant. De plus, l'eau ruisselée en un point du bassin peut être infiltrée plus en aval, et inversement, les phénomènes d'exfiltration peuvent libérer de l'eau qui avait rejoint le compartiment souterrain. Enfin, des transferts souterrains s'ajoutent généralement à ces ruissellements en surface. La Figure 1-2 résume l'ensemble de ces mécanismes. Cette complexité physique explique la relation fortement non linéaire entre les pluies et les débits, le rôle intégrateur du bassin versant étant primordial.

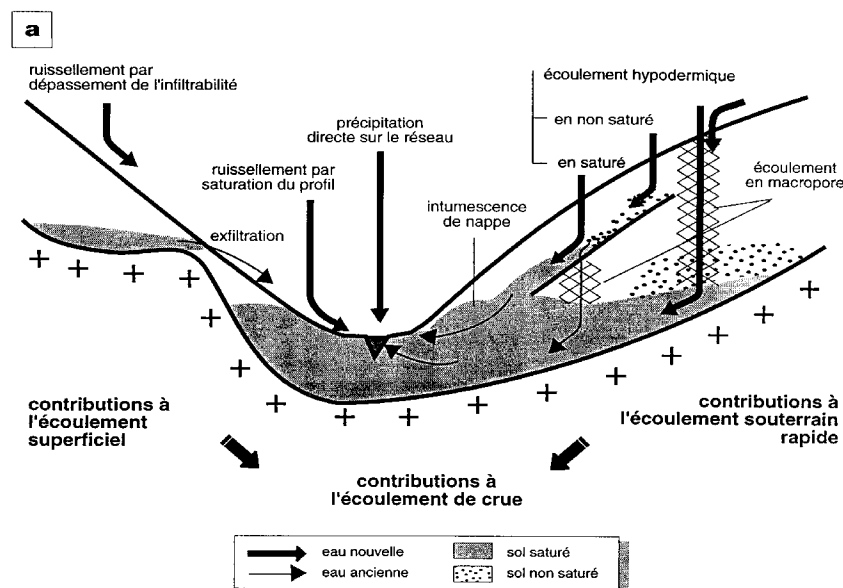


Figure 1-2. Principaux processus conduisant à la genèse des crues. Source : Ambroise (1999)

I.2. Débits extrêmes et gestion du risque : l'apport des probabilités

La gestion des risques liés à l'eau fait intervenir plusieurs niveaux d'action : l'intervention, la prévision et la prévention. L'intervention est constituée de l'ensemble des actions prises pendant l'événement pour tenter d'en limiter les conséquences. Il s'agit donc d'une action essentiellement curative. Dans le cas des crues, on peut citer les plans d'évacuation, l'organisation du secours aux personnes, ou la mise à contribution des barrages écrêteurs de crue. Dans le cas des étiages, on peut citer les arrêtés préfectoraux de limitation de l'usage de l'eau, l'arrêt préventif de certains aménagements sensibles, ou le soutien d'étiage par les

barrages réservoirs. Prévoir un événement hydrologique extrême consiste à anticiper la survenue d'une crue ou d'un étiage, à une échéance plus ou moins importante. Il s'agit donc d'une approche de type météorologique. Dans le cas des crues, on peut sur les grands bassins versants exploiter le délai de propagation amont-aval pour effectuer une prédiction de cote ou de débit à partir d'informations du même type provenant de l'amont. Pour les bassins avec un temps de réaction plus court, la prévision sera basée sur la transformation de pluies observées ou prédites en débit. La capacité de prévision est toutefois limitée par la prévisibilité des pluies et les temps de transfert à l'intérieur du bassin versant. Ainsi, une prévision de qualité acceptable est possible avec une échéance de plusieurs jours sur les grands bassins fluviaux comme la Seine, puisque les quantités de précipitations liées aux grands systèmes dépressionnaires océaniques peuvent être convenablement anticipées, et qu'une onde de crue mettra plusieurs jours à se propager le long du réseau hydrographique. A l'inverse, les crues en région cévenole ne peuvent être anticipées au-delà d'une échéance de quelques heures, puisque les systèmes convectifs méditerranéens sont beaucoup moins prévisibles, tant en terme de quantité que de localisation précise. Enfin, la prévention se place dans une optique plus climatologique, et tente de décrire les territoires les plus fréquemment soumis aux événements extrêmes, afin de réduire l'exposition au risque. On parlera, pour bien la différencier de la prévision, de prédétermination. C'est cette approche que nous allons décrire dans les sections suivantes.

I.2.1. Gestion préventive du risque : aléa et vulnérabilité

Le fait que les cours d'eau connaissent de manière occasionnelle des épisodes extrêmes est un phénomène totalement naturel, qui plus est nécessaire à l'équilibre écologique des milieux aquatiques. Ces phénomènes ne sont problématiques que parce que nous nous y exposons. Ainsi, il est d'usage de définir le risque associé à une entité de territoire comme le croisement de deux informations : la vulnérabilité et l'aléa.

La vulnérabilité peut être définie comme la sensibilité, en terme de conséquences éventuelles (économiques et/ou humaines), à la survenue d'un événement extrême. Elle est essentiellement fonction de l'occupation du sol. Par exemple, dans le cas des inondations, un champ en jachère est peu vulnérable, puisque les conséquences d'une inondation n'y seraient que marginales, alors qu'une zone d'habitation dense inondée conduirait à des dommages économiques ou humains plus conséquents. Par contre, elle ne dépend pas de l'endroit où se situe l'entité que l'on cherche à caractériser : deux hôpitaux auront une vulnérabilité équivalente qu'ils soient situés au bord d'une rivière ou en haut d'une colline. La quantification de cette vulnérabilité peut se faire de différentes manières. Outre les approches économiques (Grelot, 2004) permettant d'évaluer un coût moyen annuel des dégâts ou un consentement à payer le remboursement des dommages, on peut citer la méthode INONDABILITE (Gilard et Gendreau, 1998). Le principe y est de décider, dans le cadre d'une négociation entre les différents acteurs locaux, avec quelle fréquence on peut accepter de subir une inondation donnée, en terme de hauteur et de durée de submersion notamment. Par exemple, on peut accepter qu'une petite route de campagne soit totalement inondée pendant une semaine tous les ans, mais on n'acceptera pas qu'un axe routier majeur soit inondé pendant un jour plus d'une fois tous les dix ans. La principale difficulté réside bien sûr dans cette quantification de la vulnérabilité. Tout d'abord, elle reste extrêmement subjective : le propriétaire d'un champ n'a pas la même vision de sa vulnérabilité que le maire de la commune. De plus, s'il est possible dans certains cas d'essayer d'optimiser l'aménagement sur des critères économiques, il en va différemment dès que des vies humaines sont en jeu : peut-on accepter de faire courir des risques mortels avec une fréquence donnée ? De plus, la vulnérabilité n'est souvent définie qu'au niveau local, alors que les risques sont répétés au niveau d'un territoire donné. Ainsi, si les fréquences sont interprétées en terme de probabilité,

décréter que l'on peut accepter une inondation donnée avec une fréquence d'une fois tous les mille ans pour un aménagement sensible peut donner une impression de sécurité acceptable. Cependant, si le même type d'aménagement est répété vingt fois sur un territoire donné, cela signifie que l'on accepte d'être confronté à un tel événement environ une fois tous les cinquante ans sur au moins un de ces aménagements ! (sous hypothèse qu'ils soient exposés aux événements de manière indépendante).

Parallèlement et indépendamment de la vulnérabilité, on peut définir l'aléa en un point du territoire. Il s'agit de la probabilité de voir survenir un événement extrême donné. La quantification de cette probabilité est encore une fois la principale difficulté. Dans le cas des inondations, on utilise en général le couplage de deux approches, l'une statistique, pour étudier la distribution des valeurs fortes de débit, et l'autre hydraulique, pour transformer ces débits en hauteur et calculer ainsi les limites des champs d'inondations relatifs à une probabilité donnée. La caractérisation de l'aléa pourra être faite d'après la hauteur de submersion, la vitesse d'écoulement ou la durée de submersion. Dans la pratique, on utilise soit des hydrogrammes de crues historiques, soit des hydrogrammes de projet associés à une période de retour donnée. Parmi ces derniers, citons les hydrogrammes synthétiques monofréquence (Le Clerc, 2004), qui ont la particularité de présenter une probabilité de dépassement identique en débit de pointe et en débit moyen sur plusieurs durées.

En un point du territoire, le risque est alors défini en comparant vulnérabilité et aléa : si la vulnérabilité est supérieure à l'aléa, alors l'occupation du sol n'est pas compatible avec l'aléa, et le risque est trop important. Inversement, si la vulnérabilité est inférieure à l'aléa, alors l'occupation du sol est acceptable. Ce croisement de deux notions bien différentes peut être la cause d'un certain nombre d'ambiguïtés. Ainsi, il semble exister dans l'opinion publique la perception d'une certaine aggravation des phénomènes hydrométriques extrêmes. La réponse à cette question dépend alors de ce que l'on cherche à caractériser : la vulnérabilité aux crues comme aux étiages a effectivement augmenté de manière évidente en France, de sorte que les dommages liés à ces phénomènes extrêmes se sont aggravés. Par contre, l'aggravation du phénomène physique n'est pas clairement établie à ce jour. Apporter des premiers éléments de réponse constitue justement l'un des objectifs principaux du travail présenté dans cette thèse.

I.2.2. Les probabilités comme un outil de gestion du risque

Les méthodes probabilistes apparaissent naturellement appropriées pour la quantification de l'aléa. Ce dernier est en effet généralement exprimé, pour une variable donnée, en terme de probabilité au non-dépassement. Ces probabilités peuvent être calculées empiriquement à partir d'un échantillon d'observations, mais ces calculs ne permettent pas d'extrapoler au-delà du domaine observé. Il est donc nécessaire d'utiliser un modèle probabiliste pour décrire la distribution de la variable d'intérêt.

La théorie des valeurs extrêmes est une branche des Statistiques qui s'intéresse à la caractérisation asymptotique des maxima d'une variable aléatoire. Le théorème des valeurs extrêmes fournit le comportement limite du maximum d'un ensemble de variables indépendantes et identiquement distribuées (*iid*). Il s'agit en quelque sorte de l'équivalent du théorème central limite, qui concerne la moyenne, pour la description du maximum. Plus précisément, notons X_1, \dots, X_n une suite de variables *iid*, et $M_n = \max\{X_1, \dots, X_n\}$ le maximum de ces variables. Le théorème suivant peut alors être démontré (Fisher et Tippett, 1928):

Théorème 1. S'il existe des suites de constantes $a_n > 0$ et b_n , et une distribution non dégénérée G telles que :

$$\Pr\left((M_n - b_n) / a_n \leq x\right) \xrightarrow{n \rightarrow \infty} G(x), \quad (1-1)$$

alors G appartient à l'une des trois familles suivantes (avec $a > 0$ et $\gamma > 0$) :

Distribution de Gumbel : $G(x) = \exp\{-\exp(-(x-b)/a)\}$

Distribution de Fréchet : $G(x) = \begin{cases} 0 & \text{si } x \leq b \\ \exp\left\{-((x-b)/a)^{-\gamma}\right\} & \text{sinon} \end{cases}$

Distribution de Weibull : $G(x) = \begin{cases} \exp\left\{-\left[-((x-b)/a)^\gamma\right]\right\} & \text{si } x \leq b \\ 1 & \text{sinon} \end{cases}$,

Le choix de l'une de ces trois familles peut s'avérer problématique à partir des observations. Jenkinson (1955) a donc proposé une paramétrisation qui englobe ces trois familles, connue sous le nom de distribution généralisée des valeurs extrêmes (GEV) :

$$G(x) = \exp\left\{-\left(1 - \xi \left(\frac{x - \beta}{\alpha}\right)\right)^{1/\xi}\right\}, \quad (1-2)$$

$$\text{avec } \alpha > 0 \text{ et } 1 - \xi \left(\frac{x - \beta}{\alpha}\right) > 0$$

Dans cette expression, les paramètres β , α et ξ sont respectivement appelés les paramètres de position, d'échelle et de forme de la loi GEV. Le cas $\xi > 0$ correspond à une distribution de Weibull, le cas $\xi < 0$ à une distribution de Fréchet, et on obtient la distribution de Gumbel par passage à la limite quand $\xi \rightarrow 0$.

Appliquée à l'hydrologie, la théorie justifie donc l'utilisation d'une loi GEV pour la modélisation des valeurs maximales annuelles de débit. Précisons cependant que les données de débit couramment utilisées au pas de temps journalier ne répondent pas aux hypothèses du théorème, notamment celle du caractère *iid* des variables dont on prend le maximum. Il existe notamment une autocorrélation importante au pas de temps journalier. Le théorème peut cependant être étendu au cas de données non indépendantes (Leadbetter, 1983), en l'absence de dépendances à longue portée. De plus, cette distribution est asymptotique, et la taille des blocs sur lesquels les maxima sont calculés (typiquement, une année) pourrait être insuffisante. En pratique, la distribution GEV se révèle cependant bien souvent adaptée à la modélisation des maxima annuels, et est couramment utilisé en hydrologie comme dans d'autres domaines.

Une autre caractérisation des valeurs extrêmes d'une variable peut être obtenue non pas en considérant les valeurs maximales par bloc, mais les valeurs dépassant un seuil. Il est encore une fois possible de fournir une caractérisation asymptotique de la distribution de ces dépassements de seuil (Pickands, 1975):

Théorème 2. Soit X_1, \dots, X_n une suite de variables *iid* répondant au théorème des valeurs extrêmes (Théorème 1). Alors :

$$\Pr(X - u \leq z | X > u) \xrightarrow{u \rightarrow \infty} H(z), \text{ avec } H(z) = 1 - \left(1 - \xi \frac{z}{\sigma}\right)^{1/\xi}, 1 - \xi \frac{z}{\sigma} > 0 \quad (1-3)$$

Cette distribution est connue sous le nom de distribution de Pareto généralisée (GPD). Il existe de plus une dualité, pour des variables aléatoires données, entre la distribution GEV issue des maxima par bloc et la distribution GPD issue des dépassements de seuil. Plus précisément, les paramètres de forme de ces deux distributions sont identiques, et la relation suivante associe les autres paramètres : $\sigma = \alpha - \xi(u - \beta)$.

Un des avantages de l'approche des dépassements de seuil est qu'elle permet de s'intéresser au processus d'occurrence des événements extrêmes. Ce processus est en général décrit à l'aide d'une loi de Poisson. Plus précisément, si N_t désigne le nombre de dépassements du seuil u sur un intervalle de temps $[0 ; t]$, alors pour u suffisamment grand :

$$\Pr(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \quad (1-4)$$

La considération simultanée des processus d'occurrence et de dépassement décrits ci-dessus fournit une description équivalente à celle du modèle GEV. Les paramètres de la GEV peuvent en particulier être déduits de ceux des distributions GPD et de Poisson (Smith, 1989; Lang, 1995). L'avantage de l'approche par dépassement de seuil est la possibilité de pouvoir étoffer l'échantillon en choisissant un seuil suffisamment bas (tout en vérifiant qu'il permette d'utiliser la distribution asymptotique, cf. Coles (2001)). Cependant, le problème de la dépendance entre valeurs successives demeure. En pratique, une procédure, nommée « déclusterisation » par les statisticiens, consiste à appliquer des contraintes d'indépendance afin de ne pas sélectionner plusieurs valeurs successives dépendantes.

Les quelques résultats présentés ici constituent la base de la théorie des valeurs extrêmes, qui servira de cadre théorique général au travail présenté dans ce manuscrit. De nombreux développements complémentaires existent, concernant par exemple les méthodes d'estimation ou les extensions multivariées. Une présentation plus complète de la théorie peut être trouvée dans les ouvrages d'Embrechts *et al.* (1997) ou de Coles (2001).

II. Le contexte du changement climatique global

Nous allons à présent décrire de manière très générale le contexte du changement climatique. Tout d'abord, quelques principes fondamentaux permettront de comprendre les principaux mécanismes qui régissent le climat. La modélisation du climat par les modèles de circulation générale (GCM) sera ensuite décrite, ainsi que les incertitudes qui y sont associées. Enfin, nous dresserons une vue d'ensemble de nos connaissances actuelles sur le réchauffement global, concernant à la fois les changements observés et les changements prédits. Cette présentation sera loin d'être exhaustive, étant donné le volume de publications traitant de ces questions. Le Groupement Intergouvernemental d'Experts sur l'Evolution du Climat (IPCC) fournit régulièrement une vision beaucoup plus complète de l'état de l'art en la matière, son prochain rapport étant prévu pour l'année 2007. Nous nous contenterons donc d'exposer les publications les plus marquantes.

II.1. Qu'est-ce que le changement climatique ?

II.1.1. Quelques notions de climatologie

Le fonctionnement général du climat sur Terre est basé sur la notion d'équilibre du bilan énergétique : le rayonnement global reçu par la planète doit équilibrer le rayonnement global émis. Dans le cas contraire, la planète se refroidirait ou se réchaufferait continuellement. A l'échelle de la Terre, cet équilibre peut être décrit de la manière suivante : une partie de l'énergie reçue est directement reflétée dans l'espace par l'atmosphère et ses composants

(nuages, aérosols), une autre partie est absorbée par l'atmosphère. Au final, seule la moitié de l'énergie atteint la surface terrestre. Cette énergie est alors lentement réémise sous l'action de divers mécanismes (évapotranspiration, radiation thermique). L'atmosphère redistribue alors cette énergie, en renvoyant une partie vers l'espace, et en piégeant une autre, sous l'effet notamment des gaz à effet de serre, qui rejoint donc la surface terrestre.

Cette distribution de l'énergie sortante et entrante n'est évidemment pas uniforme, ni temporellement, ni spatialement. Elle varie tout d'abord annuellement en fonction des saisons. De plus, elle varie fortement en fonction de la latitude, du fait de l'angle avec lequel arrivent les rayonnements. Cet angle est beaucoup plus important aux pôles qu'à l'équateur, de sorte que l'énergie reçue par unité de surface y est bien moins grande. Parallèlement, l'énergie émise n'est pas non plus uniformément répartie. Ainsi, lorsque l'on effectue le bilan énergétique en différentes latitudes, on s'aperçoit que les tropiques absorbent plus d'énergie qu'ils n'en émettent, alors que les pôles ont le comportement inverse (Figure 1-3). Ce déséquilibre est la cause du phénomène de circulation générale, atmosphérique et océanique, que l'on observe sur Terre : la chaleur excédentaire est transportée des basses latitudes vers les hautes latitudes, où le bilan énergétique est déficitaire.

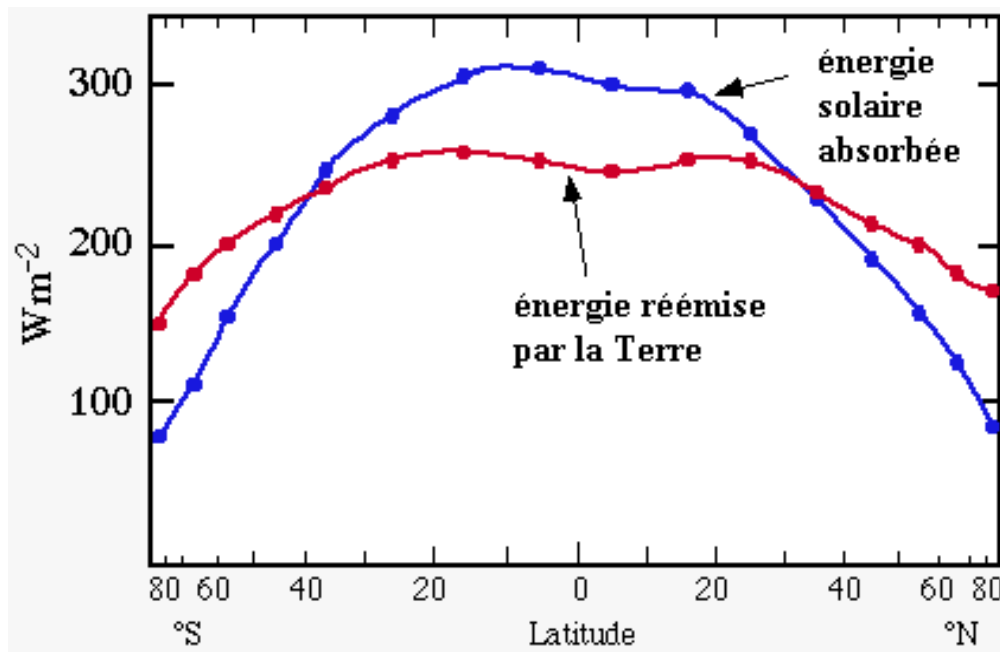


Figure 1-3. Evolution de l'énergie reçue et émise par la Terre suivant la latitude. Source : *climateprediction.net*.

La force de Coriolis, liée à la rotation de la Terre sur elle-même, complique ce schéma général. La circulation atmosphérique générale ne consiste en effet pas en une unique cellule qui lierait l'équateur aux pôles, mais en trois cellules connues sous les noms de cellule de Hadley, cellule de Ferrel et cellule Polaire. C'est l'existence de ces circulations préférentielles qui explique la présence de régimes dominants de vent à la surface de la Terre (régime d'alizés d'est au tropiques, d'ouest aux latitudes moyennes, et d'est aux pôles), ainsi que la présence de zones alternées de basses et hautes pressions.

De même, la circulation océanique générale transporte globalement la chaleur vers les pôles et les eaux froides vers l'équateur, mais présente un certain nombre de circuits préférentiels, liés notamment à la limite que constituent les continents. De plus, cette circulation est également gouvernée par des gradients de salinité de l'eau de mer : on parle donc de circulation thermohaline. La Figure 1-4 présente un schéma simplifié de cette circulation globale, qui reste bien plus méconnue que la circulation atmosphérique. De plus,

cette circulation est beaucoup plus lente que celle observée dans l'atmosphère, et de manière générale, l'inertie thermique des océans est bien plus importante que celle de l'atmosphère.

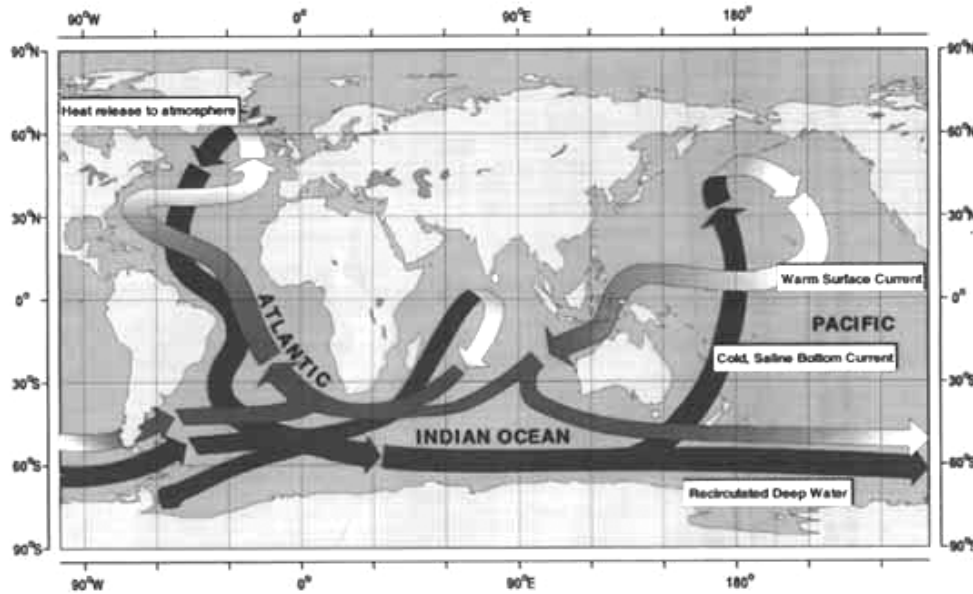


Figure 1-4. Représentation simplifiée de la circulation thermohaline. Source : climateprediction.net.

II.1.2. Variabilité temporelle du climat

Il est d'usage de parler de la stationnarité des phénomènes hydro-climatiques, voire du climat en général. Il convient de définir ce que l'on entend par le terme « stationnarité ». Au sens statistique, un processus temporel est dit stationnaire (au sens strict) si sa distribution est invariante temporellement. Pour pouvoir appliquer cette définition théorique au climat, il convient de préciser l'échelle temporelle à laquelle on s'intéresse. En toute généralité, le climat n'est de toute évidence pas stationnaire : il a en effet fortement évolué depuis la naissance de la Terre, et évoluera encore d'ici sa disparition, ne serait-ce que parce que le forçage radiatif solaire évoluera au cours des prochains millions d'années. De même, les mécanismes d'alternance de périodes glaciaires et interglaciaires sont bien connus, et sont liés à des forçages astronomiques, notamment la position de la terre sur son orbite. Ce mécanisme possède une échelle temporelle de l'ordre de plusieurs dizaines de milliers d'années. Nous nous intéressons ici à une échelle de temps bien plus faible, de l'ordre de quelques centaines d'années au maximum. Les variations qui ont une échelle temporelle géologique y sont donc négligeables. Inversement, lorsque l'on s'intéresse à une échelle de temps plus faible, inférieure à l'année, une variabilité temporelle apparaît également, liée aux comportements saisonniers. Cependant, ces comportements restent relativement stables d'une année sur l'autre, si bien que lorsque l'on décrit le climat sur la base de variables annuelles, la stationnarité de ces variables peut *a priori* être acceptée (ou plutôt pourrait l'être en l'absence d'influences anthropiques).

Cependant, une des difficultés majeures dans l'étude du climat est l'existence de variabilités qui possèdent des échelles temporelles intermédiaires, de l'ordre de la dizaine d'années. Cette variabilité à basse fréquence est particulièrement problématique dans l'optique de la recherche d'évolutions temporelles, puisqu'elle peut pour certaines variables faire apparaître des tendances, surtout si la durée d'observation n'est pas assez importante. Ce type de phénomène se traduit notamment par l'existence des oscillations australes (ENSO, plus connu comme « El Niño ») et Nord-Atlantique (NAO). La NAO se caractérise par exemple par le gradient de pression entre la zone dépressionnaire d'Islande et la zone

anticyclonique des Açores. En phase positive, ce gradient est élevé, favorisant les régimes d'ouest. Il en résulte un déplacement du « rail des tempêtes » qui traverse l'Atlantique vers le nord, et une douceur et une humidité accrue sur l'Europe du Nord (Cassou, 2004). Inversement, en phase négative, les tempêtes seront moins importantes, et auront tendance à suivre une trajectoire plus sud-ouest, apportant plus d'humidité autour de la Méditerranée, et moins en Europe du Nord. Globalement, les deux dernières décennies sont caractérisées par une phase positive de la NAO, et des relations avec certaines variables climatiques ont pu être mises en évidence (Shorthouse et Arnell, 1999). Cependant, l'impact de la NAO sur les événements extrêmes est bien moins évident, et à notre connaissance, aucune relation claire n'a pu être établie en France.

D'autres phénomènes présentent également ce type de comportement oscillatoire à une échelle régionale. Ainsi, il semble que la pluviométrie en Afrique de l'Ouest se caractérise par un fonctionnement par phases, la période actuelle étant marquée par d'importants déficits pluviométriques (Hubert *et al.*, 1989; Paturol *et al.*, 1997; Servat *et al.*, 1997a; Paturol *et al.*, 2003). De même, l'activité cyclonique dans l'Atlantique semble connaître une alternance de périodes de faible et de forte activité (Landsea *et al.*, 1999). Llasat *et al.* (2002) observent également des oscillations sur les fortes pluies en Espagne et dans le Sud de la France, en utilisant des données historiques. Tous ces exemples montrent que le climat présente un certain nombre de comportements oscillatoires à basse fréquence que l'on peut observer, mais dont le mécanisme reste encore très largement méconnu. Il conviendra de considérer cette variabilité comme une cause possible d'apparentes évolutions dans l'étude de stationnarité.

Enfin, des mesures indirectes permettent de remonter de plus en plus loin dans le temps, et d'avoir ainsi une vision plus globale de la variabilité climatique. Ces mesures sont principalement issues de l'étude de la composition des glaces polaires, de données dendroclimatiques ou de l'étude des coraux. Mann *et al.* (1998) ont ainsi effectué, sur la base de plusieurs de ces indicateurs, une reconstitution des températures de surface globales depuis plusieurs siècles. Certaines carottes de glaces polaires permettent de remonter bien plus loin (plusieurs milliers d'années), au prix évidemment d'une incertitude plus importante (Dansgaard *et al.*, 1993). Ces informations modifient continuellement la connaissance du climat, qui semble bien plus capricieux que ce que pensaient les climatologues il y a encore quelques années. A titre d'illustration, la dernière déglaciation, qui a précédé la période climatique qui perdure depuis environ 10000 ans (l'Holocène), a été marquée par des accidents climatiques très importants, induisant au moins localement une augmentation de la température très brutale, de l'ordre de 7°C en quelques décennies (Severinghaus *et al.*, 1998). D'autres phénomènes brutaux ont été relevés durant l'Holocène, mais semblent moins importants et plus localisés. Il semble donc que le climat puisse connaître des modifications majeures de circulation générale et de températures en l'absence de toute influence humaine. Ceci ne suffit pas à remettre en cause l'attribution d'au moins une partie de changements que nous connaissons actuellement à l'influence anthropique, mais illustre le fait que certains aspects du fonctionnement climatique restent largement incompris.

II.1.3. L'impact des activités humaines

Comment l'homme peut-il influencer sur ce système complexe et qui mobilise une énergie considérable qu'est le climat ? La réponse se trouve dans les gaz à effet de serre, qui ont déjà été évoqués précédemment. La particularité de ces gaz est qu'ils laissent passer les radiations solaires entrantes, qui possèdent une petite longueur d'onde, mais absorbent une partie de la chaleur réémise par la surface terrestre à une longueur d'onde plus importante. L'énergie ainsi absorbée est en partie renvoyée dans l'espace, mais une autre partie retourne à la surface terrestre, contribuant ainsi à son réchauffement.

La présence de ces gaz dans l'atmosphère est une composante majeure du climat tel que nous le connaissons sur Terre : en leur absence, la température y serait d'environ -17°C ! Or, l'activité humaine a très fortement modifié la composition chimique de l'atmosphère depuis la révolution industrielle (Figure 1-5), et notamment la concentration des gaz à effet de serre. L'augmentation de cette concentration devrait alors résulter en un réchauffement de la surface de la Terre, puisque plus d'énergie réémise devrait être absorbée et renvoyée vers la surface.

La réalité est cependant bien plus complexe, de sorte que ce raisonnement n'est pas suffisant pour garantir, et encore moins pour quantifier l'impact de l'augmentation de la concentration des gaz à effet de serre dans l'atmosphère. Les différentes composantes du climat ne peuvent en effet pas être dissociées les unes des autres, puisqu'elles sont liées par un grand nombre d'interactions. Ainsi, la modification d'une de ces composantes se répercutera sur les autres par un des phénomènes de rétroaction. Par exemple, une augmentation du taux de CO_2 favorisera la croissance végétale, et donc l'absorption de ce même CO_2 par les plantes (rétroaction négative). De même, dans un monde plus chaud, la couverture nuageuse devrait augmenter, favorisant ainsi la réflexion de l'énergie solaire entrante. Des rétroaction positives, amplifiant le réchauffement, peuvent également être attendues. Par exemple, si la surface des glaciers diminue, alors la quantité d'énergie réfléchie devrait également diminuer, contribuant ainsi au réchauffement de la surface. Ces phénomènes ne constituent que quelques exemples des nombreuses interactions qui existent dans le fonctionnement du climat. Afin de tenter de quantifier quels peuvent être les impacts de l'augmentation de la concentration des gaz à effet de serre, il est ainsi nécessaire de considérer le système climatique dans son ensemble. C'est le rôle des modèles de circulation générale, qui ont pour but de recréer une évolution climatique réaliste sous l'influence de différents forçages (naturels et/ou anthropique).

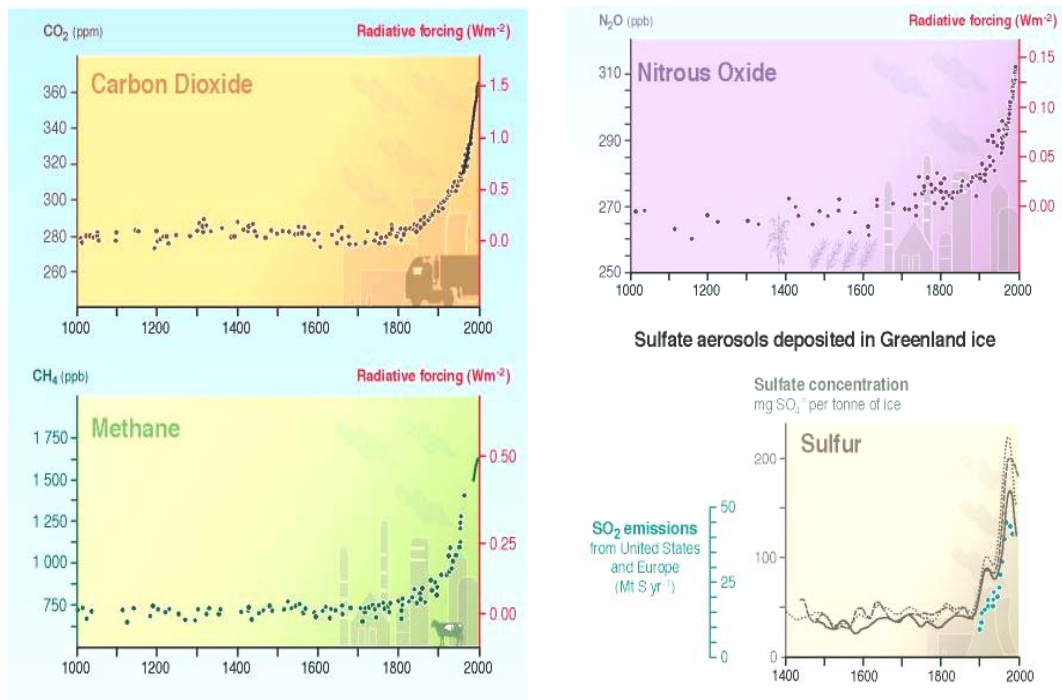


Figure 1-5. Evolution de la concentration de quelques gaz dans l'atmosphère (IPCC, 2001).

II.2. Changement climatique et incertitudes

II.2.1. Fonctionnement d'un modèle de circulation générale (GCM)

Un GCM peut être considéré comme une représentation numérique de la Terre et des processus qui gouvernent son climat. A cet effet, la planète est discrétisée sous la forme d'un assemblage de cubes. Au niveau horizontal, la résolution de la grille ainsi définie est typiquement de l'ordre de la centaine de kilomètres pour les modèles globaux. Des modèles dits régionaux peuvent utiliser une résolution plus fine sur une partie du globe. Au niveau vertical, la grille n'est pas régulièrement espacée, et compte typiquement une vingtaine de couches pour la représentation de l'atmosphère, et une trentaine pour celle des océans. Chacun de ces cubes est caractérisé par un ensemble de variables qui décrivent les différentes composantes du climat : pression, température, précipitations, humidité, végétation, etc. Toutes ces composantes interagissent les unes avec les autres, à la fois à l'intérieur d'un cube et avec les cubes voisins, et évoluent temporellement et spatialement sous l'effet du forçage radiatif notamment. Les lois physiques qui régissent ces évolutions sont connues de manière plus ou moins précises, et sont donc appliquées aux variables contenues dans chaque cube, avec une résolution temporelle qui est de l'ordre de la demi-heure. Ainsi, à partir de conditions initiales données, l'évolution des conditions climatiques sous l'effet de différents forçages peut être simulée.

Une des difficultés majeures des GCM est liée à la discrétisation spatiale : en effet, de nombreux phénomènes possèdent une échelle spatiale très largement inférieure à celle de la grille (les nuages, par exemple), et ne peuvent donc pas être représentés fidèlement, bien que leur rôle dans le fonctionnement du climat soit primordial. Pour remédier à cette difficulté, des formules simplifiées sont utilisées afin de déterminer, à partir de certaines variables, des valeurs « moyennes » de ces phénomènes à petite échelle sur la grille. Par exemple, la couverture nuageuse peut être déduite de la température et de l'humidité d'un cube. L'utilisation de tels schémas simplifiés est connue sous le nom de paramétrisation du GCM. Certains de ces schémas sont bien connus et validés par les observations, mais d'autres liés à des phénomènes plus méconnus sont largement empiriques.

Les GCM sont utilisés pour évaluer l'impact de certains forçages, et notamment le forçage anthropique lié à l'augmentation de la concentration des gaz à effet de serre, sur le climat. Deux types d'analyse sont en général effectuées. La première consiste à quantifier la sensibilité du climat au forçage anthropique. Pour ce faire, un GCM est utilisé pour simuler plusieurs années de données climatiques, avec une concentration atmosphérique en CO₂ représentative de la période actuelle (la période 1961-1990 est généralement utilisée). Cette première étape constitue la phase de « contrôle », sur laquelle on peut vérifier que le modèle a correctement décrit le climat observé. Puis, un doublement de cette teneur en CO₂ est simulé, et le modèle génère alors de nouvelles données, qui évoluent jusqu'à atteindre un nouvel état d'équilibre. La différence entre ce nouvel état et l'état observé avant doublement de la teneur en CO₂ définit la sensibilité du climat au forçage anthropique simulé. L'autre approche utilise des scénarios d'émission de gaz à effet de serre. Ces scénarios utilisent des modèles socio-économiques, qui tentent de prédire l'évolution au cours du siècle d'indicateurs tels que la population, le PIB, les besoins énergétiques, etc. L'impact des choix politiques en matière de développement sont aussi intégrés, et l'ensemble de ces évolutions est traduit en terme d'émissions de gaz à effet de serre. Quelques scénarios types sont ainsi retenus, et décrivent les différentes directions que pourraient prendre nos sociétés dans le futur. Ces scénarios d'émission sont alors utilisés en forçage des GCM, qui décrivent la manière dont réagira le climat suivant le scénario adopté. Une des intérêts de cette approche est qu'elle a une portée

plus « prédictive », ce qui est un avantage en terme de communication, notamment auprès des décideurs politiques.

II.2.2. Les incertitudes dans la modélisation du climat

Tenter de décrire le fonctionnement du climat et encore plus sa possible évolution sous l'effet de différents forçages est une tâche extrêmement complexe, étant donné le grand nombre de phénomènes et d'échelles impliqués. Tous les facteurs ne sont évidemment pas contrôlés, de sorte que l'utilisation de GCM s'accompagne de plusieurs sources d'incertitude.

La première incertitude concerne la mesure des phénomènes climatiques, puisque aucun GCM ne peut fonctionner sans données. Les observations, affectées par divers degrés d'incertitude suivant la variable mesurée, interviennent à plusieurs niveaux : tout d'abord, comme données initiales des modèles. Il est important à ce propos d'avoir conscience que la modélisation climatique est moins sensible à ce problème de données initiales que la modélisation météorologique, tout simplement car les objectifs sont différents. Par exemple, il est difficile de prévoir quelle sera précisément la température en un point du globe à une échéance qui dépasse quelques jours, car de faibles variations des données initiales peuvent se traduire par de fortes variations des prédictions. Par contre, dans la plupart des cas, les modèles parviennent à respecter la climatologie de ce point du globe, c'est à dire que les prédictions sont conformes à la distribution des températures observées. Les observations sont également utilisées en validation, c'est à dire pour vérifier que les simulations issues des GCM sont en accord avec le climat observé. L'incertitude des mesures joue également un grand rôle dans la description de la variabilité climatique, qui permet de relativiser la significativité de certaines évolutions. En effet, les scientifiques recherchent des indicateurs permettant de reconstituer le climat qui régnait sur Terre il y a quelques centaines, voire quelques milliers d'années. Ces indicateurs sont évidemment des mesures indirectes, issues par exemple de l'analyse chimique des glaces polaires, qui sont bien plus incertaines que les mesures directes dont nous disposons depuis une durée de l'ordre de la centaine d'années. Enfin, les observations interviennent dans la paramétrisation des modèles, pour prendre en compte des phénomènes qui ont une échelle spatiale inférieure à la résolution.

La structure même des GCM constitue une source importante d'incertitude. Les différents GCM disponibles actuellement diffèrent sur plusieurs aspects, parmi lesquels la résolution, la paramétrisation, les algorithmes utilisés, etc. En conséquence on peut observer des différences notables entre les simulations issues de plusieurs GCM. Cette différence ne s'observe pas seulement en prévision, mais également en reconstitution du climat actuel. Ainsi, Kharin et Zwiers (2005a) montrent que les températures (moyennes et extrêmes) sont reconstituées de manière cohérente par un ensemble de 15 GCM, mais que la simulation des précipitations varie bien plus fortement d'un modèle à l'autre, plus particulièrement pour les précipitations extrêmes en zone inter-tropicale.

Les données de forçage sont également incertaines, plus particulièrement celles liées aux scénarios sociaux-économiques envisagés pour le futur. Tout d'abord, pour un scénario donné, les évolutions démographiques et économiques ne sont que des estimations. De plus, les formules utilisées pour lier ces évolutions à celle des émissions de gaz à effet de serre sont également des simplifications de la réalité (Hourcade, 2005). Ensuite, les différences entre scénarios se répercuteront également sur l'évolution climatique, ce qui crée une importante fourchette d'incertitude pour les projections à l'horizon 2100, par exemple. De plus, il est fort probable que le futur ne corresponde à aucun des scénarios retenus par la communauté internationale. Il n'est donc pas assuré que cette fourchette, certes basée sur des scénarios tranchés, reste valable pour une réalité qui serait un mélange des différentes évolutions sociaux-économiques envisagées. D'ailleurs, l'IPCC a refusé d'assigner une probabilité à cette fourchette.

Enfin, une dernière incertitude prépondérante apparaît lorsque l'on cherche à estimer l'impact du changement climatique sur une variable qui n'est pas directement prédite par les GCM. C'est le cas notamment des débits, puisque les écoulements dans le réseau hydrographique ne sont pas pris en compte dans les modèles. Pour tenter d'évaluer la manière dont réagiront les régimes hydrologiques au réchauffement global, des modèles pluie-débit sont donc utilisés. Plusieurs nouvelles sources d'incertitudes apparaissent alors. La première est celle liée à ce type de modèle : des erreurs peuvent en effet affecter la structure du modèle, qui n'est qu'une représentation très simplifiée de la réalité, mais également les observations utilisées pour le calage des paramètres. Enfin le calage en lui-même est entaché d'incertitudes. La seconde source d'erreurs est liée à la discordance des échelles spatiales et temporelles utilisées par les modèles pluie-débit par rapport aux sorties fournies par les GCM. Hormis dans le cas de très grands bassins versants, la plupart des systèmes hydrologiques réagissent à des pluies localisées, et concentrées sur un pas de temps assez court. Des procédures dites de « downscaling » sont alors utilisées pour transformer, typiquement, des pluies mensuelles simulées sur une grille très étendue en des pluies journalières calculées sur des pas d'espace plus réduits, en prenant notamment en compte l'effet orographique. Les relations existant entre les caractéristiques pluviométriques à petite et grand échelle spatio-temporelle sont cependant loin d'être évidentes, ce qui crée une incertitude forte sur ces reconstitutions.

Cette problématique des incertitudes liées à la modélisation climatique est unanimement reconnue comme centrale par les climatologues. De nombreuses études ont donc été entreprises pour tenter sinon de les prendre en compte formellement, au moins d'en évaluer l'ampleur et de les hiérarchiser. Parmi celles-ci, on peut citer le projet *climateprediction.net* (Stainforth *et al.*, 2005), qui a pour but de quantifier l'incertitude liée à la paramétrisation des modèles sur la détermination de la sensibilité du climat à un doublement de la teneur en CO₂. Pour ce faire, plus de 2000 simulations climatiques ont été effectuées sur des PC de particuliers, chacune des simulations correspondant à une légère modification des paramètres des modèles. Sur ces 2000 séries simulées, celles qui sont parvenues à recréer correctement le climat observé ont alors été utilisées pour estimer la distribution des réponses au doublement de la teneur en CO₂. La gamme de sensibilités obtenues varie de +1.9°C à +11.5°C, ce qui est bien supérieur à la gamme qui était jusqu'alors communément admise. Bien que cette étude soit critiquable sur certains aspects (utilisation d'un modèle océanique simplifié, et validation incomplète du modèle), elle témoigne d'une volonté de mieux prendre en compte les incertitudes dans le processus de modélisation. D'un point de vue plus hydrologique, Prudhomme *et al.* (2003) ont tenté de hiérarchiser les incertitudes émanant de la chaîne de traitements nécessaire à l'évaluation de l'impact du réchauffement sur les régimes hydrologiques (en crue notamment). Plus précisément, les auteurs ont utilisé plusieurs GCM, plusieurs scénarios d'émission et plusieurs sensibilités du climat afin d'évaluer la réponse hydrologique au réchauffement. Les résultats de cette étude montrent que le choix du GCM constitue la source d'incertitude prépondérante, la réponse hydrologique semblant moins sensible au scénario d'émission ou à la sensibilité du climat. Cette étude met en avant la difficulté d'anticiper l'évolution des régimes hydrologiques, d'autant plus que les incertitudes liées au downscaling et au modèle pluie-débit n'ont pas été pris en compte.

II.3. Ce que l'on a d'ores et déjà observé

Dans cette section, nous allons décrire les principales évolutions climatiques qui ont été observées depuis le début du XX^e siècle, en privilégiant les études concernant l'échelle globale. La bibliographie comprend de très nombreuses études plus localisées, qu'il est impossible de décrire de manière exhaustive. Le cas de la France métropolitaine sera traité en détail dans une section ultérieure (II.5.). De plus, nous nous contentons de décrire les changements, sans aborder le problème de l'attribution à une cause anthropique, bien que

pour certaines variables, cette attribution soit possible (Stott *et al.*, 2000; Spagnoli *et al.*, 2002; Tank *et al.*, 2005). Nous nous restreignons également aux changements affectant les températures, les précipitations et les écoulements. Il convient cependant de préciser que les impacts, avérés ou potentiels, du changement climatique global sont bien plus larges, et peuvent par exemple affecter le niveau des océans, la circulation océanique ou atmosphérique, les phénomènes cycloniques, la biodiversité, les aires de répartition des espèces vivantes et des maladies, les économies locales, etc.

II.3.1. Températures et cryosphère

Si l'on parle de réchauffement global pour désigner le phénomène de changement climatique, c'est parce que la température est la variable sur laquelle les changements les plus marqués apparaissent. L'estimation de l'ampleur de ce réchauffement est réalisée en utilisant des bases de données de température de surface, à la fois au-dessus des terres et des océans (Peterson et Vose, 1997; Hansen *et al.*, 1999; Jones *et al.*, 2001; Smith et Reynolds, 2005). Les quantifications du réchauffement peuvent légèrement différer suivant la base de données utilisée, mais les conclusions restent identiques : la Terre se réchauffe à un rythme anormalement élevé depuis le début du siècle dernier (Figure 1-6). Ce réchauffement a connu une première phase entre 1900 et 1945, suivi d'une stabilisation, voir d'une légère diminution entre 1946 et 1975, avant de reprendre à un rythme plus soutenu depuis 1976. Globalement, l'IPCC (2001) estime que le réchauffement global depuis la fin du XIX^e siècle a atteint environ $+0.6 \pm 0.2$ °C.

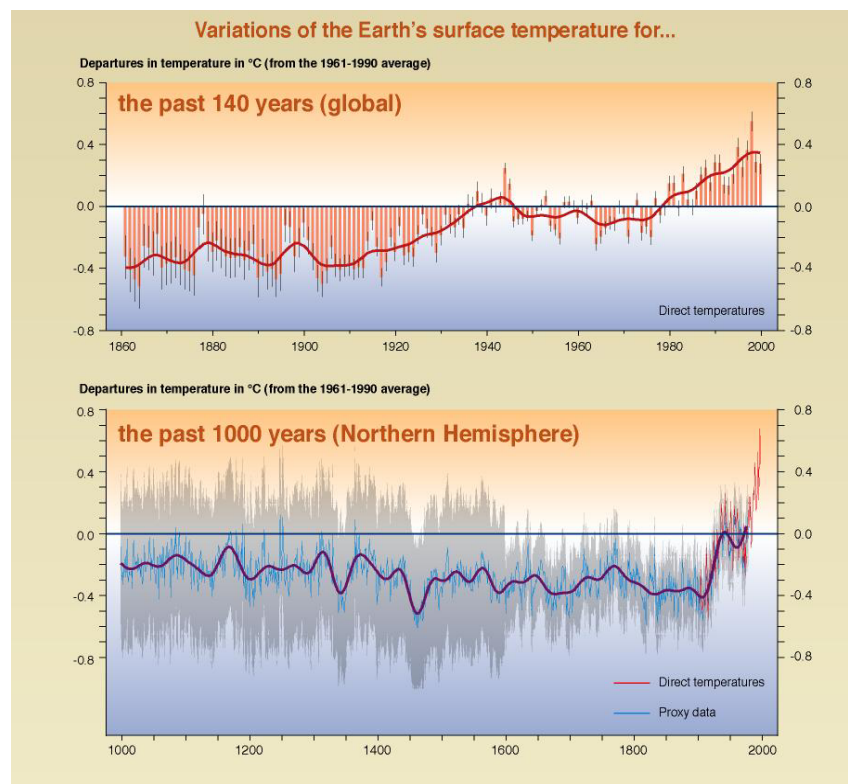


Figure 1-6. Anomalies des températures annuelles globales par rapport à la période 1961-1990 (IPCC, 2001).

Globalement, on observe que depuis 1976, l'hémisphère nord (+0.24 °C/décennie) se réchauffe plus vite que le sud (+0.11 °C/décennie), et que les terres (+0.22 °C/décennie) se réchauffent plus vite que les océans (+0.14 °C/décennie) (Figure 1-7). Ces évolutions de surface ont des répercussions sur les températures des hautes couches de l'atmosphère et des profondeurs des océans. Ainsi, l'énergie calorifique contenue dans les 300 premiers mètres des océans présente globalement un changement similaire à celui des températures de surface (Levitus *et al.*, 2000), bien que moins marqué (+0.037°C/décennie). Les couches atmosphériques sont plus délicates à étudier, car des biais métrologiques liés à l'évolution des outils de mesure peuvent perturber les observations (Sherwood *et al.*, 2005). Il semble cependant que le réchauffement de la troposphère suive le schéma observé en surface depuis 1979 (Mears et Wentz, 2005), alors que plus en altitude, la basse stratosphère a tendance à se refroidir.

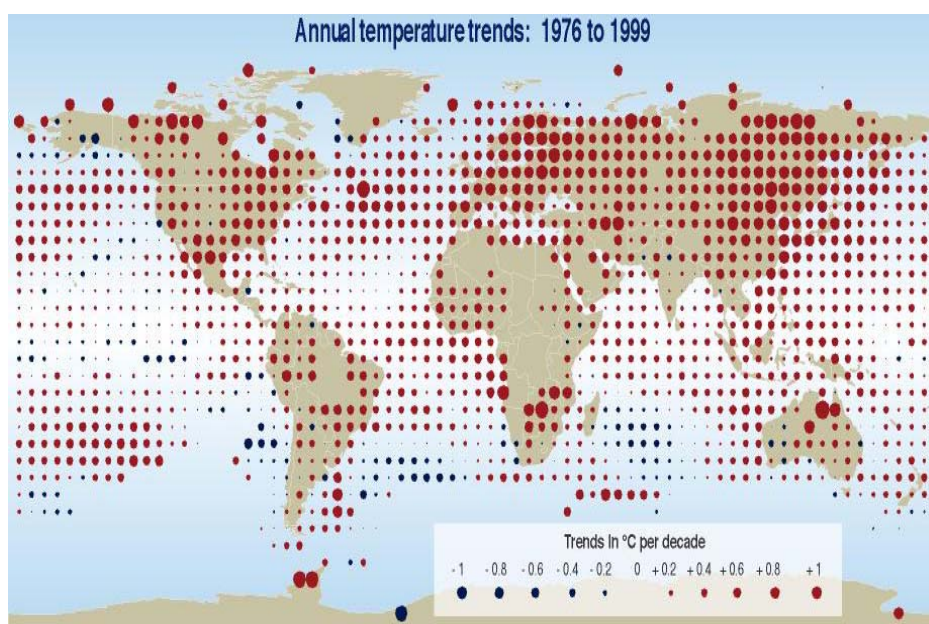


Figure 1-7. Tendances sur les températures annuelles de surface entre 1976 et 1999 (IPCC, 2001).

Certaines répercussions sont également observées sur des phénomènes liés aux extrêmes de température (chauds et froids). La couverture spatiale de données journalières adéquates pour l'étude de ces phénomènes est cependant bien moins importante que celle des températures moyennes. Parmi les études les plus marquantes, on peut citer celle de Frich et Alexander (2002), qui recouvre sur la seconde moitié du XX^e siècle l'Amérique du Nord et l'Europe, ainsi qu'une partie de l'Asie et de l'Océanie. L'Amérique du Sud et l'Afrique ne sont cependant pas incluses dans l'analyse. Les changements les plus marqués et les plus cohérents spatialement concernent le nombre de jours de gel, avec une diminution significative au niveau global, et le nombre de « nuits chaudes » (pourcentage de jours avec une minimale journalière supérieure au dernier décile de la distribution des minimales journalières), qui augmente significativement. Les minima de températures semblent globalement présenter plus de changements que les maxima. D'autres variables, notamment un index décrivant la durée des vagues de chaleur, montrent des changements plus régionaux, qui ne sont pas significatifs globalement. Ces résultats à grande échelle ont été complétés par des études plus régionales (par exemple, Tank et Konnen (2003) ou Moberg et Jones (2005) en Europe, Easterling *et al.* (2000) aux Etats-Unis, Plummer *et al.* (1999) en Océanie, ou Manton *et al.* (2001) dans le Sud-Est Asiatique et le Pacifique Sud). La plupart de ces indices décrivent cependant des extrêmes modérés : les conclusions quant aux extrêmes plus rares

sont encore incertaines à l'échelle globale, mais commencent à apparaître localement (par exemple, augmentation en fréquence mais pas en intensité des extrêmes chauds dans l'Atlantique Nord (Nogaj *et al.*, 2006)).

L'état de la cryosphère constitue également un indicateur intéressant du réchauffement climatique, puisque les températures constituent un des facteurs prépondérants de l'évolution des glaciers, des banquises, etc. De nombreuses tendances viennent confirmer le réchauffement global. Tout d'abord, la couverture neigeuse annuelle semble avoir diminué d'environ 10% depuis les années 60 (Robinson *et al.*, 1993). De même, l'étendue de la banquise a régressé depuis les années 70 en Arctique, à un rythme d'environ -2.8% par décennie (Parkinson *et al.*, 1999). Dans l'hémisphère nord, les limites des terres gelées (permafrost) ont également migré au nord, entraînant une baisse de la surface concernée (Wang *et al.*, 2000). Enfin, le changement le plus spectaculaire affecte l'étendue des glaciers non polaires. Le World Glacier Monitoring Service centralise aujourd'hui des informations, parfois très anciennes, sur l'étendue de nombreux glaciers répartis tout autour du globe. La tendance générale est à un fort recul de la plupart des glaciers, à deux exceptions notables près, les glaciers de Norvège et de Nouvelle Zélande. Ces exceptions pourraient s'expliquer par l'augmentation des précipitations dans ces régions. Cependant, dans de nombreux cas, la récession semble avoir commencé au début du vingtième siècle, alors que les séries de températures disponibles n'indiquent pas de réchauffement net en cette période. Ce décalage temporel est, à notre connaissance, toujours incompris.

II.3.2. Précipitations et écoulements

Le changement climatique devrait en toute logique affecter les précipitations. De manière extrêmement simplifiée, on s'attend à une augmentation des précipitations globales dans un monde plus chaud, du fait d'une augmentation de l'évaporation de l'eau de mer. Au-dessus des terres émergées, on observe effectivement à l'échelle globale une légère augmentation des précipitations au cours du siècle dernier, mais cette tendance n'est pas répartie uniformément au niveau du globe (Figure 1-8).

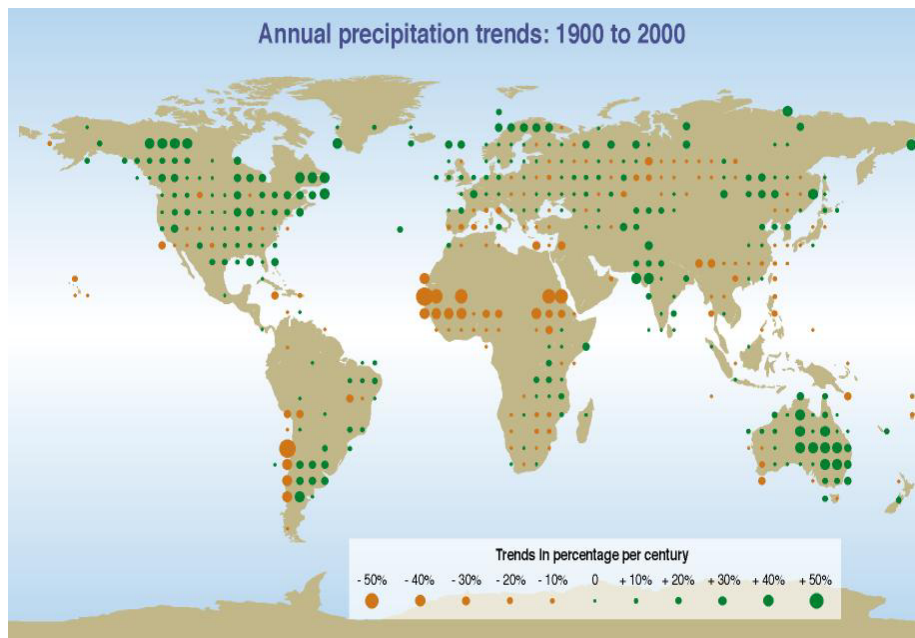


Figure 1-8. Tendence sur les précipitations annuelles entre 1900 et 2000 (IPCC, 2001).

L'augmentation des précipitations semble plus marquée aux moyennes latitudes qu'en zone tropicale. Les Etats Unis (Groisman *et al.*, 2001), le Canada (Zhang *et al.*, 2000), une grande partie Est de l'Australie (Hennessy *et al.*, 1999), l'Europe du Nord et de l'Ouest (Schönwiese et Rapp, 1997) ont connu une tendance à l'augmentation au cours du siècle dernier. Dans ce dernier cas, la phase positive persistante de la NAO qui perdure depuis les deux dernières décennies peut expliquer cette tendance. En Inde, l'augmentation apparente des précipitations semble être liée à une variabilité à grande échelle temporelle (Kumar *et al.*, 1999). A l'inverse, des diminutions ont été reportées au sud-ouest de l'Australie (Haylock et Nicholls, 2000), en Chine (Zhai *et al.*, 1999), sur le bassin Méditerranéen (Pier Vitali *et al.*, 1997) et surtout en Afrique de l'Ouest (Servat *et al.*, 1997b; Paturel *et al.*, 2003; Tapsoba *et al.*, 2004). Là encore, des oscillations peuvent souvent être mises en avant, les changements ne se manifestent donc pas sous la forme d'une simple tendance linéaire.

Les études concernant les valeurs extrêmes de précipitations ne couvrent qu'une part limitée du globe, pour les mêmes raisons que celles évoquées pour les extrêmes de température. L'étude de Frich *et al.* (2002) possède une bonne couverture spatiale dans l'hémisphère nord, ainsi qu'en Australie et en Afrique du Sud. La cohérence spatiale des résultats est assez faible, mais des tendances globalement significatives sont mises en avant sur le nombre de jours de pluies supérieures à 10 mm (augmentation), la durée des périodes sèches les plus longues (diminution) et les maxima annuels des cumuls sur cinq jours (augmentation). Ces résultats ont été étendus récemment à des extrêmes de précipitations plus intenses par Groisman *et al.* (2005). Des études régionales ont également permis d'observer cette augmentation des forts épisodes pluvieux, notamment aux Etats-Unis (Karl et Knight, 1998), en Russie (Gruza *et al.*, 1999) et en Europe (Tank et Konnen, 2003; Moberg et Jones, 2005). *A contrario*, l'aggravation de la sécheresse en Afrique de l'Ouest est accompagnée d'une diminution des forts événements pluvieux (Easterling *et al.*, 2000; Groisman *et al.*, 2005).

Ces évolutions pluviométriques devraient avoir un certain nombre d'impacts sur les écoulements. Encore une fois, les données sont cependant bien plus éparées que dans le cas des températures, ce qui rend problématique la recherche d'évolutions à grande échelle. Les seules analyses globales ont, à notre connaissance, été effectuées par Yoshino (1999) et Kundzewicz *et al.* (2005). Dans la première étude, on remarque que les résultats sont généralement cohérents avec les tendances observées sur les précipitations. Ainsi, l'auteur rapporte une décroissance des débits annuels au Sahel, ainsi qu'une légère augmentation pour les rivières d'Europe de l'Ouest et d'Amérique du Nord. La seconde étude est dédiée aux crues, et les maxima annuels des débits journaliers de 195 stations hydrométriques, principalement situées en Amérique du Nord, en Europe et en Australie, sont utilisés. Il apparaît que la majorité (70% environ) des séries étudiées sont stationnaires, et qu'il y a autant de tendances à la hausse que de tendances à la baisse.

A l'exception des deux études précédentes, l'échelle spatiale d'analyse est généralement plus réduite pour les écoulements. La comparaison entre les résultats obtenus sur plusieurs régions peut être délicate, car les indices utilisés pour caractériser le régime moyen ou extrêmes des cours d'eau diffèrent suivant les auteurs. A titre d'exemple, nous présentons un résumé des résultats obtenus aux Etats-Unis et en Europe. La Figure 1-9 résume plusieurs études de grande ampleur qui ont été menées aux Etats-Unis (Lettenmaier *et al.*, 1994; Hubbard *et al.*, 1997; Zeng et Kelly, 1997; Lins et Slack, 1999; Douglas *et al.*, 2000; Groisman *et al.*, 2001). Leurs résultats sont assez cohérents, et peuvent être résumés de la manière suivante : dans une bande centrale allant de la frontière mexicaine à la région des grands lacs, les étiages deviendraient moins sévères. En Floride et sur la côte du Pacifique Nord, un certain nombre de stations présentent néanmoins une aggravation des étiages. Dans le reste du pays, les changements n'ont guère de cohérence spatiale, ni en basses eaux, ni en

hautes eaux. Ces résultats sont à nuancer par des études régionales plus ciblées, qui démontrent le comportement particulier de certains bassins versants. Ainsi, Olsen *et al.* (1999) détectent une aggravation des phénomènes de crues sur le Mississippi et le Missouri, alors que dans le bassin semi-aride du Rio Puerco, Molnar et Ramirez (2001) observent que les débits d'étiages sont en augmentation, tandis que les hauts débits diminuent.

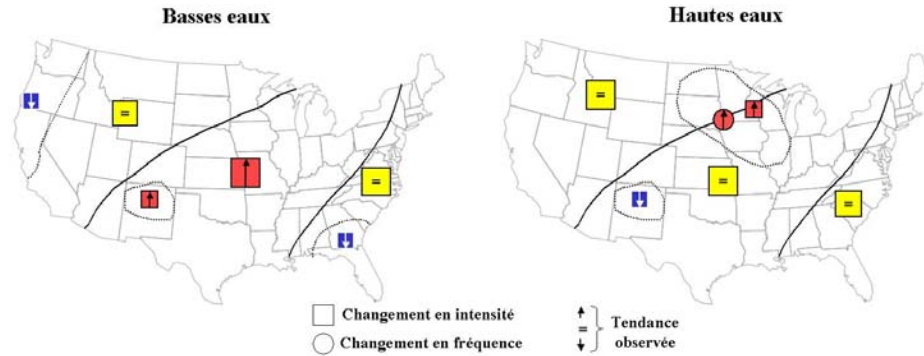


Figure 1-9. Schématisation des principaux changements détectés sur les écoulements extrêmes aux Etats-Unis.

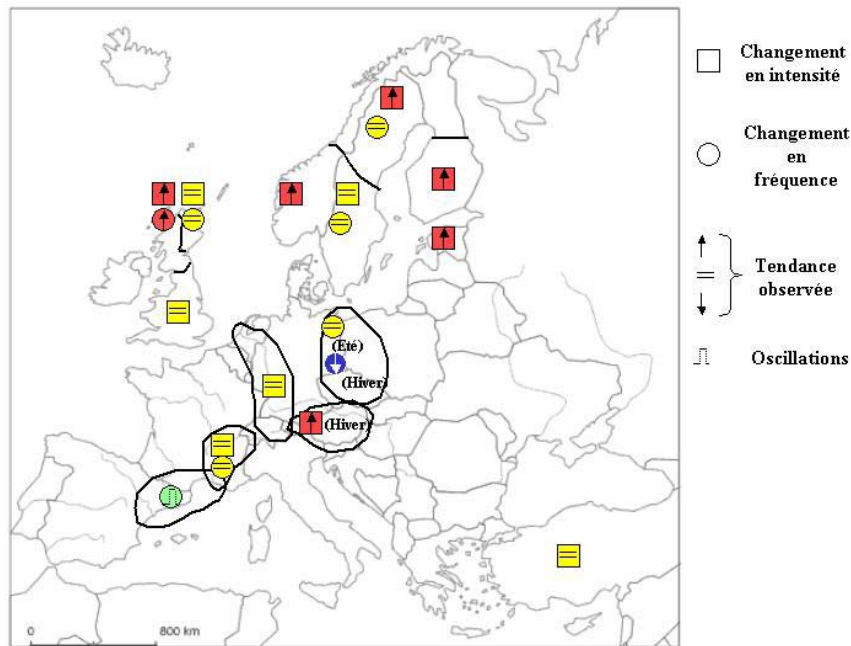


Figure 1-10. Schématisation des principaux changements détectés en hautes eaux sur les écoulements extrêmes en Europe.

En Europe, de nombreuses études ont été réalisées à l'échelle nationale sur les hautes eaux (Black, 1996; Bronstert, 1996; Robson et Reed, 1996; Robson *et al.*, 1998; Kiely, 1999; Black et Burns, 2002; Llasat *et al.*, 2002; Robson, 2002; Werritty, 2002; Mudelsee *et al.*, 2003; Sauquet, 2003; Lindstrom et Bergstrom, 2004; Birsan *et al.*, 2005; Cigizoglu *et al.*, 2005). La Figure 1-10 en résume les principaux résultats. Il est difficile de dégager une tendance générale à si grande échelle. Cependant, il semblerait que les pays d'Europe du Nord ainsi que les îles Britanniques (hors Angleterre) subissent une augmentation des écoulements en crue, qui serait due à l'augmentation des précipitations issues des systèmes dépressionnaires de l'Atlantique Nord. Ce signal semble d'ailleurs en partie concordant avec l'oscillation Nord Atlantique (NAO). Précisons également que l'analyse effectuée sur le pourtour Méditerranéen par Llasat *et al.* (2002) possède l'originalité d'utiliser des données historiques remontant

jusqu'au XIV^e siècle. Grâce à ce recul temporel, les auteurs parviennent à mettre en évidence une variabilité climatique à l'échelle de quelques dizaines d'années, qui est susceptible d'interférer avec une tendance monotone à long terme. Concernant les basses eaux, Hisdal *et al.* (2001) ont étudié l'évolution des étiages en utilisant plus de 600 stations hydrométriques sur la période 1962-1990. Les résultats montrent que sur la plupart des stations, les tendances ne sont pas significatives. Néanmoins, des cohérences régionales semblent apparaître, notamment en Espagne, en République Tchèque, en Slovaquie et en Angleterre, où les déficits de volumes augmentent. Inversement, les étiages semblent moins sévères en Europe de l'Est et en Europe Centrale (Russie, Biélorussie, Ukraine, Allemagne) ainsi que sur les côtes Atlantiques de la Norvège.

II.4. Ce que l'on attend

Nous décrivons dans cette section les évolutions climatiques attendues au cours du siècle, telles que prédites par les GCM. Ces prédictions sont entachées d'importantes incertitudes, détaillées en section II.2. C'est la raison pour laquelle la plupart des résultats présentés sont issus de simulations effectuées à partir de plusieurs modèles et de plusieurs scénarios d'émission de gaz à effet de serre, ce qui permet d'avoir une vision plus exacte de la gamme des climats futurs possibles. Les principaux scénarios d'émission sont résumés dans la Figure 1-11. Nous nous restreindrons encore une fois à l'étude des températures, des précipitations et des écoulements.

II.4.1. Températures

La Figure 1-12 décrit l'évolution attendue des températures journalières annuelles à l'horizon 2100, pour un scénario d'émission donné (A2). L'élévation moyenne de température pour ce scénario devrait se situer aux alentours de 3°C, avec de forts contrastes spatiaux. Ainsi, les terres se réchauffent bien plus que les océans, et l'hémisphère Nord bien plus que l'hémisphère Sud. Ce sont donc les hautes latitudes du nord qui devraient connaître le réchauffement le plus marqué. Inversement, d'autres régions devraient moins se réchauffer que la moyenne mondiale (sud de l'Amérique du Sud, Asie du Sud-Est). Généralement, ces caractéristiques sont globalement observées avec d'autres scénarios, même si l'ampleur du réchauffement est directement liée à l'évolution des gaz à effet de serre dans l'atmosphère. De plus, les minima journaliers devraient augmenter plus rapidement que les maxima, entraînant une diminution de l'amplitude diurne.

La Figure 1-13 permet d'évaluer les incertitudes sur l'évolution des températures liées à l'évolution des gaz à effet de serre et à la modélisation climatique. Etant donné le grand nombre de scénarios proposés par l'IPCC (35) et le coût calculatoire d'une simulation climatique pour un GCM complet océan-atmosphère, des modèles climatiques simplifiés ont ici été utilisés. Ils ont cependant été calibrés sur des GCM complets pour reproduire la réponse globale en terme de température (et donc en terme de sensibilité du climat). La fourchette de l'augmentation des températures globales à l'horizon 2100 s'étend de 1.4 à 5.8°C. Environ la moitié de cette incertitude est liée au choix du scénario d'émission, l'autre moitié résultant du choix du modèle climatique. Pour un scénario d'émission donné, la largeur des barres verticales montre que l'incertitude sur la quantification précise du réchauffement reste très importante. Néanmoins, l'hypothèse d'absence de réchauffement n'est envisagée par aucun des modèles et avec aucun des scénarios.

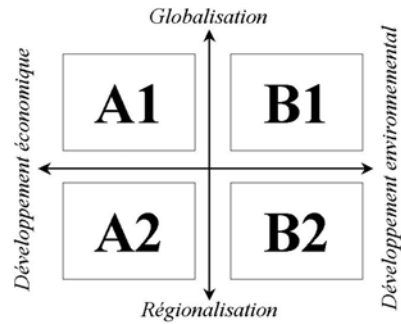


Figure 1-11. Schématisation des principaux scénarios d'émission de l'IPCC.

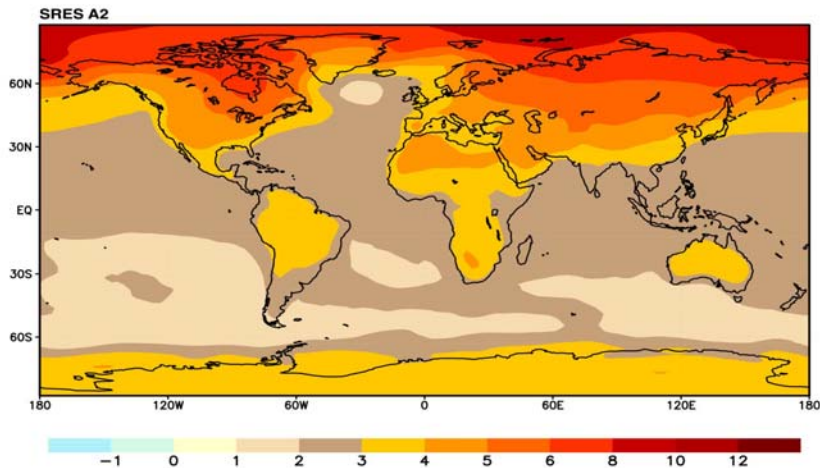


Figure 1-12. Evolution en °C des températures annuelles, calculée comme la différence entre la moyenne des périodes 2071-2100 et 1961-1990. Les projections sont moyennées sur plusieurs GCM et obtenues avec le scénario d'émission A2. Source : IPCC (2001).

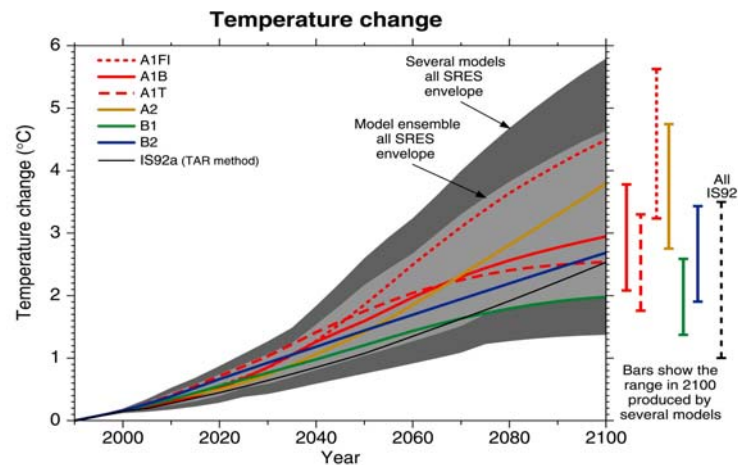


Figure 1-13. Evolution de la température globale annuelle suivant sept scénarios d'émission. Un modèle climatique simple, calibré sur 7 GCM dont la sensibilité du climat varie entre 1.7°C et 4.2°C, est utilisé. La zone claire représente l'étendue des évolutions moyennées sur les 7 modèles pour l'ensemble des 35 scénarios d'émission de l'IPCC. La zone foncée représente l'étendue des évolutions obtenues avec les 35 scénarios et les 7 modèles. Source : IPCC (2001).

L'évolution attendue des extrêmes de températures a été étudiée par Zwiers et Kharin (1998) et Kharin *et al.* (2000; 2005a; 2005b; 2005). Un exemple de résultat est fourni par la Figure 1-14 pour le scénario A1B. Les modèles utilisés simulent des extrêmes globalement cohérents avec les observations. On observe une augmentation des quantiles de période de retour 20 ans pour les maxima annuels de température, plus marqué aux latitudes moyennes qu'aux hautes latitudes. Cette évolution est du même ordre que l'augmentation attendue des températures moyennes estivales. Les quantiles de période retour 20 ans pour les minima annuels devraient présenter des changements plus marqués, principalement aux hautes latitudes. Le réchauffement est cette fois plus marqué pour les extrêmes froids que pour les températures moyennes hivernales.

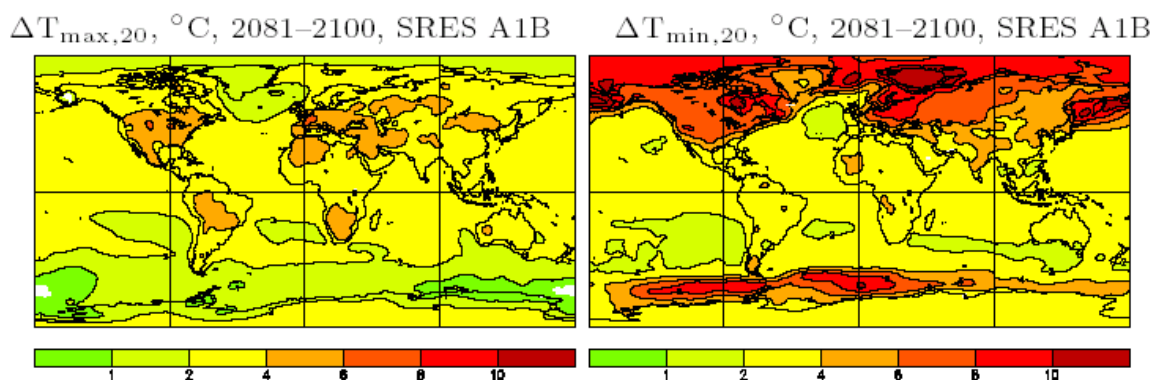


Figure 1-14. Evolution des quantiles de période de retour 20 ans des températures maximales et minimales annuelles. Les quantiles sont estimés sur les périodes 2081-2100 et 1981-2000. Les projections sont moyennées sur 15 GCM et obtenues avec le scénario d'émission A1B. Source : Kharin et Zwiers (2005a).

II.4.2. Précipitations et écoulements

Comme cela avait été le cas pour les évolutions observées sur les précipitations annuelles, les tendances prévues à l'horizon 2100 sont assez contrastées spatialement (Figure 1-15). Globalement, les précipitations devraient augmenter, mais certaines régions devraient connaître une nette diminution de la pluviométrie. Ainsi, les précipitations augmentent aux moyennes et hautes latitudes, ainsi qu'en zone tropicale, alors qu'elles diminuent dans la plupart des régions subtropicales. Cette diminution serait particulièrement marquée dans la zone méditerranéenne. De plus, les évolutions diffèrent suivant les saisons : ainsi, l'augmentation de la pluviométrie concernerait l'été comme l'hiver aux hautes latitudes, mais ne s'exprimerait qu'en hiver aux latitudes moyennes. En Australie, en Amérique Centrale et en Afrique Australe, on s'attend plutôt à une diminution des précipitations hivernales.

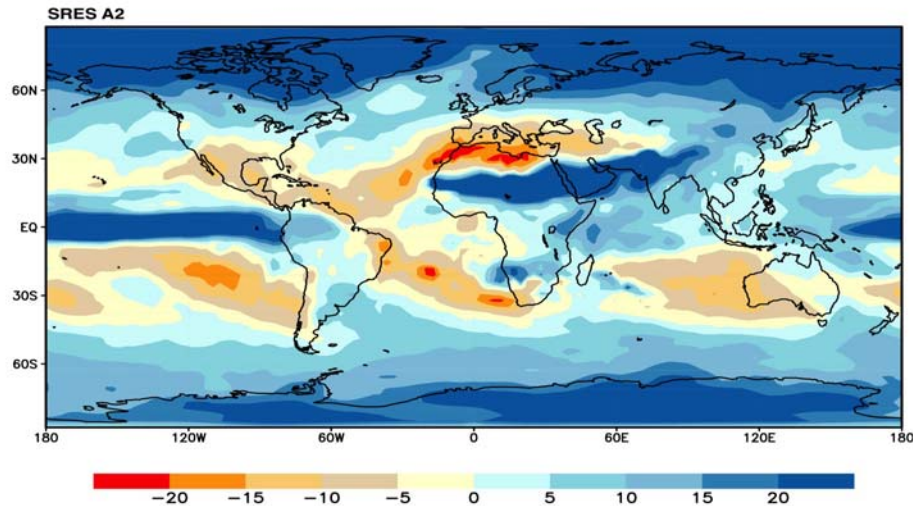


Figure 1-15. Evolution relative (en %) des précipitations annuelles, calculée entre la moyenne des périodes 2071-2100 et 1961-1990. Les projections sont moyennées sur plusieurs GCM et obtenues avec le scénario d'émission A2. Source : IPCC (2001).

L'évolution attendue des extrêmes de précipitations est décrite dans la Figure 1-16. La confiance que l'on peut accorder à ces projections est cependant limitée, car les simulations ne recréent qu'imparfaitement les extrêmes observés, et la variabilité inter-modèles est très importante, notamment dans les zones tropicales. Globalement, on observe une augmentation des fortes précipitations qui est plus importante que l'évolution des précipitations moyennes. Ainsi, même les régions affectées par une baisse des précipitations annuelles pourraient connaître une hausse des fortes précipitations. Ce résultat semble indiquer que l'impact du réchauffement sur les précipitations ne se traduirait pas par une simple translation de la distribution des pluies, mais par un changement plus complexe de forme.

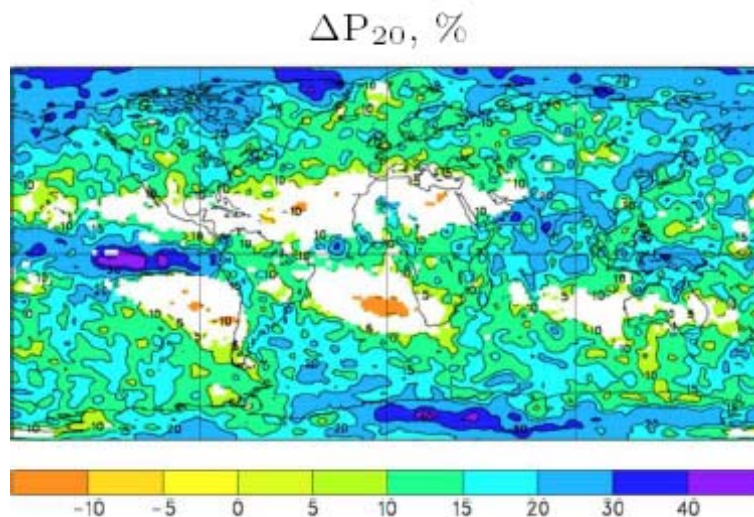


Figure 1-16. Evolution relative des quantiles de période de retour 20 ans des précipitations maximales annuelles. Les quantiles sont estimés sur les périodes 2081-2100 et 1981-2000. Les projections sont moyennées sur 15 GCM et obtenues avec le scénario d'émission A1B. Les zones figurant en blanc correspondent à des évolutions non significatives à 10%. Source : Kharin et Zwiers (2005a).

Les impacts de ces évolutions thermiques et pluviométriques sur les écoulements sont difficiles à évaluer. Tout d'abord, ces évolutions peuvent avoir des effets antagonistes : une augmentation pluviométrique peut par exemple être compensée par une augmentation de l'évapotranspiration. La réaction hydrologique au réchauffement devrait donc être très variable d'un bassin versant à l'autre, mais également entre les saisons. De plus, les échelles spatio-temporelles manipulées par les GCM sont bien plus importantes que celles qui conditionnent le comportement hydrologique des cours d'eau, notamment pour les phénomènes de crue.

Ainsi, la plupart des études effectuées se sont soit intéressées à une échelle globale, au détriment de la précision de la quantification des écoulements, soit à l'échelle d'un bassin versant particulier, permettant une évaluation plus réaliste des débits. Dans la première catégorie, on peut citer par exemple les études de Voss *et al.* (2002) ou Arnell (1999; 2003). En première approximation, Voss *et al.* (2002) estiment que les écoulements devraient globalement augmenter d'environ 6% à la surface des terres à l'horizon 2090 (doublement de la concentration en CO₂), compte tenu des évolutions prédites sur l'évapotranspiration et les précipitations. Arnell (1999; 2003) utilise quant à lui un modèle hydrologique pour transformer les variables prédites par les GCM en débits, chaque carré de la grille étant considéré comme un bassin versant. Le routage de l'eau dans les rivières n'est donc pas pris en compte. Les écoulements annuels semblent augmenter aux hautes latitudes, en Afrique Equatoriale et dans le Sud-Est Asiatique, et diminuer pour la plupart des latitudes moyennes et tropicales (Figure 1-17). Certaines différences notables apparaissent cependant régionalement suivant le modèle climatique utilisé. De plus, la diminution des écoulements est amplifiée sur les bas débits pour les régions où les écoulements annuels diminuent. La saisonnalité des écoulements semble également changer aux latitudes moyennes, à cause de la diminution des précipitations tombant sous forme de neige. Les évolutions sur les forts écoulements sont moins nettes, comparées à la variabilité naturelle.

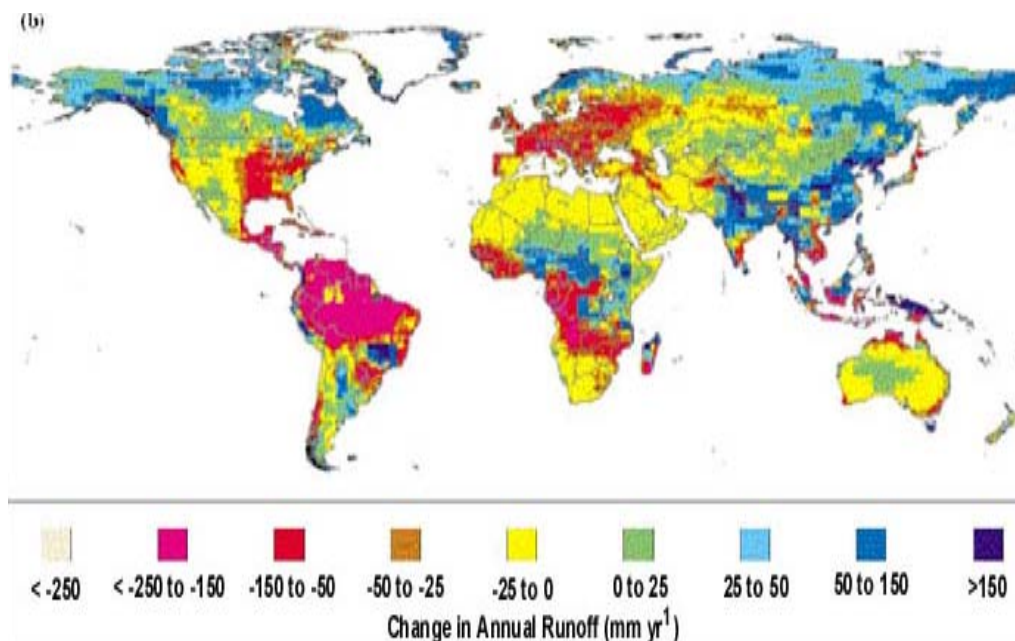


Figure 1-17. Evolution des écoulements annuels calculés sur les périodes 2021-2050 et 1961-1990. Le modèle climatique utilisé est HadCM3, avec une augmentation annuelle de 1% du CO₂. Source: IPCC (2001).

Des études plus proches de la réalité hydrologique ont également été effectuées à l'échelle des bassins versants, notamment pour des bassins de grande taille, pour lesquels les échelles des GCM sont compatibles avec celles des modèles hydrologiques (Lettenmaier *et al.*, 1999; Arora et Boer, 2001; Voss *et al.*, 2002). Il ressort notamment de ces études une très forte variabilité entre les réponses de différents bassins versants. A titre d'illustration, Voss *et al.* (2002) estiment que le débit journalier décennal devrait augmenter d'environ 26% à l'exutoire du Yangtze Kiang (Chine), mais devrait diminuer de 42% pour le bassin de la Saskatchewan et du Nelson (Canada). Etant donné l'importance du rôle intégrateur joué par le bassin versant, il semble donc difficile de quantifier des tendances sur les écoulements en rivière qui seraient valables à l'échelle globale.

II.5. Le cas de la France métropolitaine

Nous concluons cet exposé du contexte du changement climatique en rappelant les principaux acquis à l'échelle de la France métropolitaine, qui correspond à l'échelle de l'analyse que nous souhaitons effectuer sur les débits.

II.5.1. Impacts observés

Les principaux résultats concernant l'évolution des températures en France au cours du siècle dernier ont été obtenus par Mestre (2000), Moisselin *et al.* (2002), Spagnoli *et al.* (2002) ainsi que par les acteurs du projet IMFREX (IMPact des changements anthropiques sur la FRéquence des phénomènes EXtrêmes de vent, de température et de précipitations). Les figures présentées dans cette section sont issues du site Internet du projet (<http://medias.obs-mip.fr/imfrex/web/index>), avec l'aimable autorisation de Michel Déqué.

Le réchauffement observé au cours du XX^e siècle atteint environ 1°C pour les températures annuelles. Ce réchauffement semble s'être accéléré depuis le milieu des années 1970, comme cela a pu être observé à l'échelle globale, pour atteindre un rythme d'environ 0.6°C par décennie sur la période 1976-2003. On peut également observer des disparités géographiques importantes (cf. Figure 1-18, gradient est-ouest pour les températures minimales journalières et nord-sud pour les maximales journalières). Les minimales ont globalement plus augmenté que les maximales, notamment en été. Ces évolutions ont eu des répercussions sur la plupart des indices décrivant les températures, comme le nombre de jours chauds ($T_x > 25^\circ\text{C}$) ou le nombre de jour de gel ($T_n < 0^\circ\text{C}$).

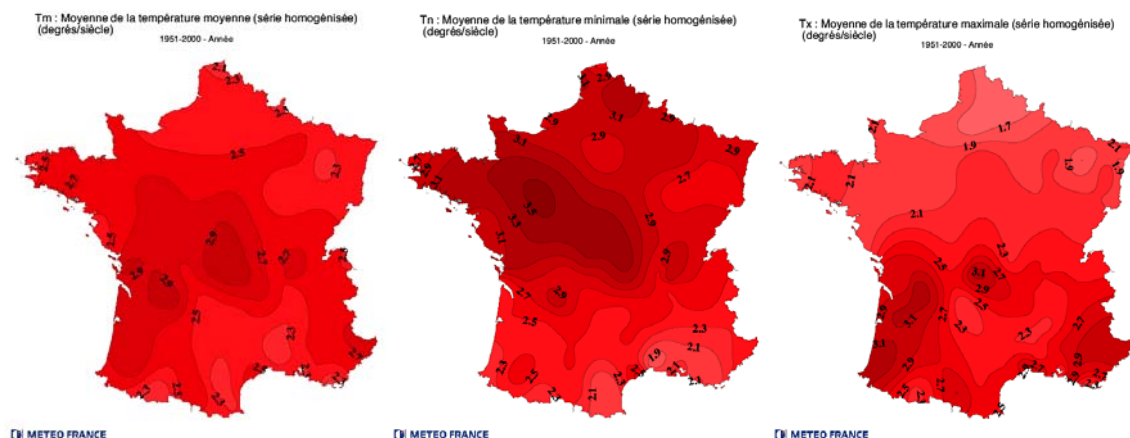


Figure 1-18. Evolution observée des températures journalières moyennes, minimales et maximales. Source : Projet IMFREX (<http://medias.dsi.cnrs.fr/imfrex/web/>).

Le réchauffement peut également être observé sur les indices décrivant les températures parmi les plus froides ou les plus chaudes de l'année (Figure 1-19). Il semble cependant que les extrêmes froids aient moins augmenté que les extrêmes chauds : le réchauffement ne consisterait donc pas en une simple translation de la distribution des températures. Le comportement des extrêmes chauds de température a également été analysé par Parey *et al.* (2006), en utilisant des épisodes de températures supérieures à un seuil chaud. Il semble que la fréquence des dépassements de seuil soit en augmentation assez nette, surtout depuis les 30 dernières années, alors que l'intensité de ces dépassements semble plus stationnaire. Ces résultats sont cohérents avec ceux obtenus par Nogaj *et al.* (2006) à une échelle plus importante (Atlantique Nord).

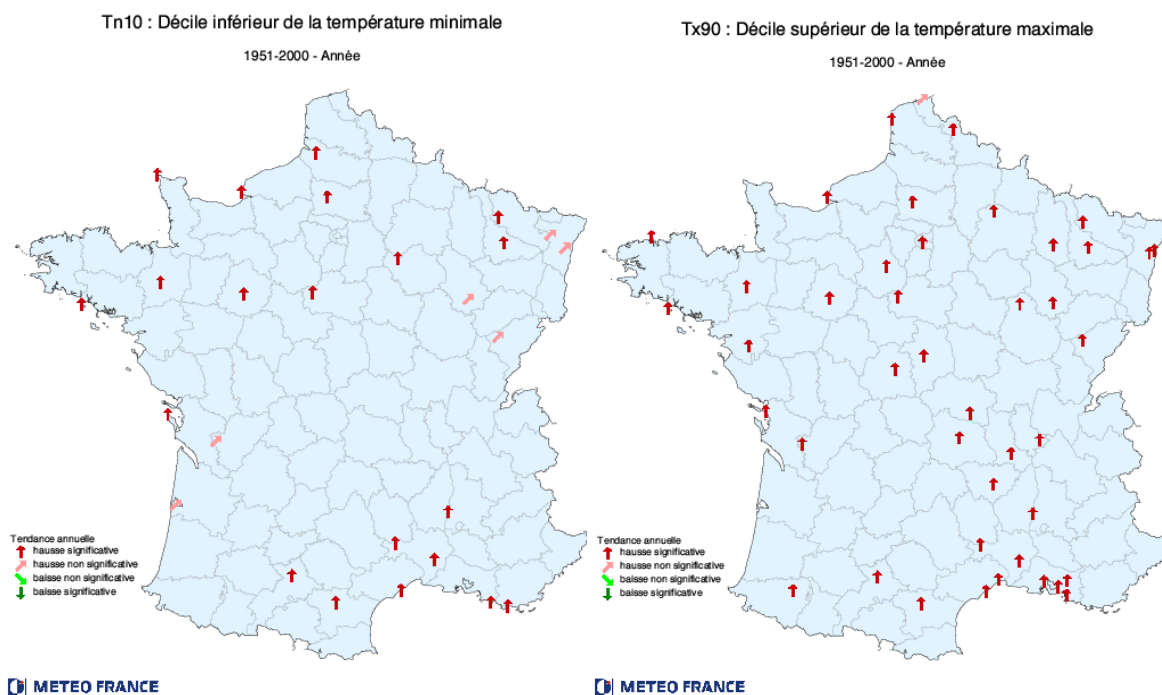


Figure 1-19. Evolution observée des extrêmes annuels de températures. Source : Projet IMFREX (<http://medias.dsi.cnrs.fr/imfrex/web/>).

L'évolution des précipitations, présentée en Figure 1-20, est moins marquée que celle des températures. Les cumuls annuels sont globalement en augmentation, surtout sur la partie nord du pays, mais la plupart de ces tendances ne sont pas significatives localement. L'augmentation la plus marquée se situe au printemps, alors qu'en été, la quasi-totalité du pays subit une chute des cumuls.

Les indices décrivant les extrêmes de précipitations ont globalement tendance à augmenter, mais la significativité des changements dépend fortement de l'indice. Ainsi, on peut observer dans la Figure 1-21 de nombreuses augmentations significatives sur les périodes pluvieuses les plus longues. Précisons cependant qu'une longue période pluvieuse ne correspond pas nécessairement à un extrême pluviométrique en termes de cumuls. D'ailleurs, les évolutions sont bien moins nettes pour les maxima annuels des cumuls sur trois jours. Il semble donc difficile de discerner, à l'heure actuelle, une évolution sur les précipitations intenses.

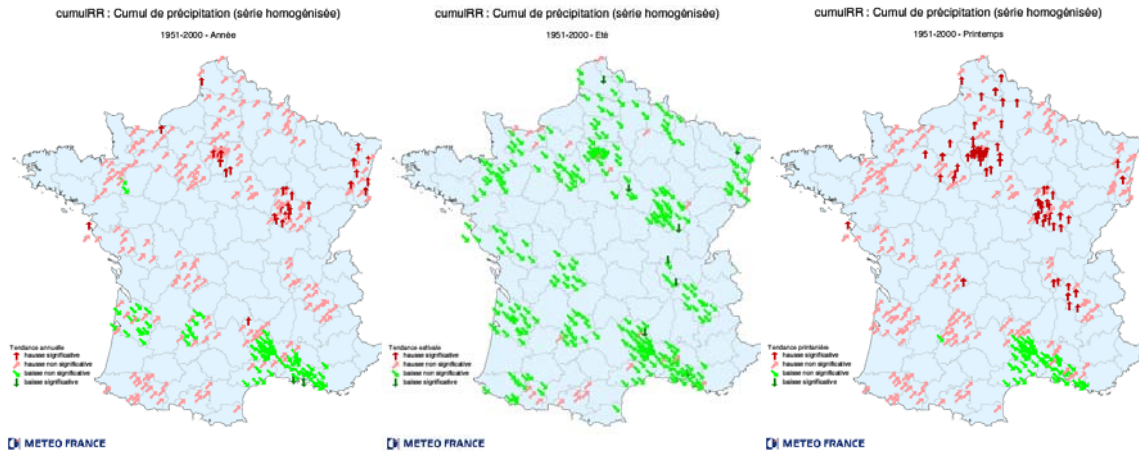


Figure 1-20. Evolution observée des cumuls de précipitations annuels, estivaux et printaniers. Source : Projet IMFREX (<http://medias.dsi.cnrs.fr/imfrex/web/>).

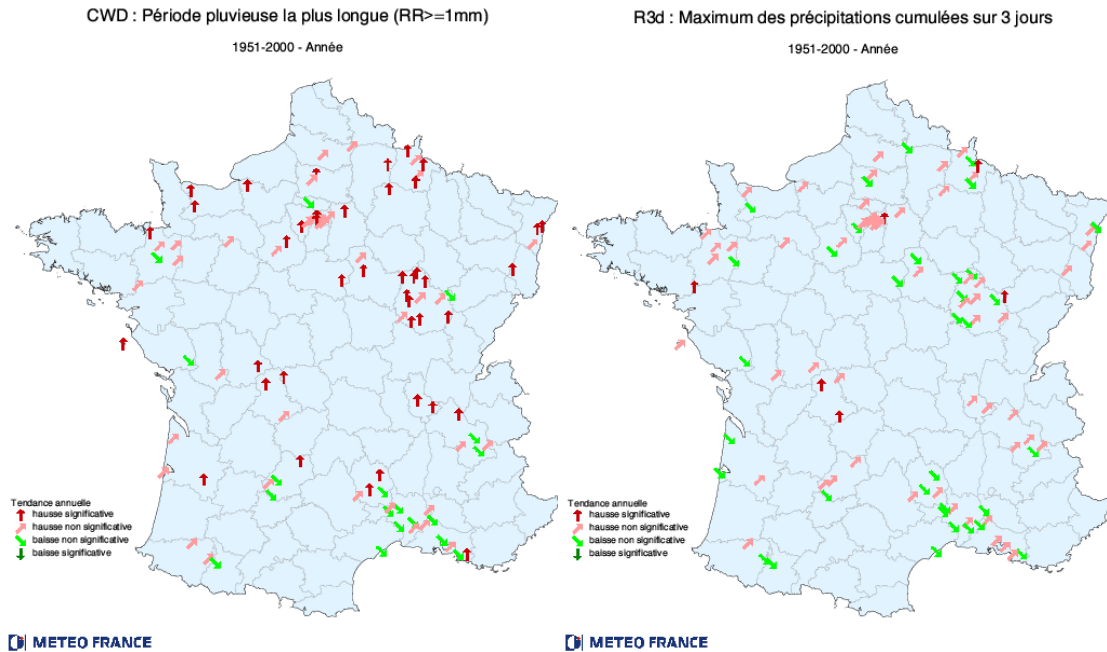


Figure 1-21. Evolution observée sur des indices décrivant les extrêmes de précipitations. Source : Projet IMFREX (<http://medias.dsi.cnrs.fr/imfrex/web/>).

Quelques études se sont également intéressées aux évolutions hydrologiques au cours du siècle dernier. Krasovskaia *et al.* (2002) ont ainsi étudié le régime des bassins versants de la zone Rhône-Méditerranée-Corse, et ont mis en évidence une tendance vers une plus grande variabilité dans la répartition des écoulements au cours de l'année, sans pour autant que cette variabilité puisse être attribuée formellement au changement climatique. Sauquet *et al.* (2003) ne discernent quant à eux aucune évolution liée au climat sur la partie française du bassin du Rhône. Une étude portant sur le bassin du Têt (Ludwig *et al.*, 2004), dans le sud-est de la France, montre une augmentation des débits de crue dans la partie aval du bassin. Enfin, une récession des glaciers Alpins français est observée par Vincent *et al.* (2002; 2004), et semble au moins en partie liée à l'augmentation observée sur les températures.

II.5.2. Prévisions pour le XXI^e siècle

Les prévisions présentées ici sont issues du modèle climatique régional ARPEGE de Météo France. Elles correspondent aux données simulées avec le scénario A2 à l'horizon 2100. Le réchauffement attendu est de l'ordre de 3 à 3.5°C pour les températures annuelles (Figure 1-22), et est plus marqué en été (4 à 5 °C). L'augmentation des températures estivales serait également plus marquée dans la partie sud du pays. Encore une fois, ce réchauffement se traduira par des évolutions sur un certain nombre d'indices dérivés. Ainsi, le nombre de jours de gel devrait connaître une diminution nette (Figure 1-23). Les valeurs extrêmes de températures devraient également suivre cette tendance, notamment les extrêmes chauds. Ainsi le nombre de jours très chauds (température journalière supérieure à 35°C) pourrait passer d'environ 1 jour par an à l'échelle de la France à 14 jours en moyenne. La distribution des températures extrêmes devrait également connaître d'importants bouleversements, comme l'illustre la Figure 1-24. Ainsi, la canicule de l'été 2003 pourrait préfigurer ce que sera un été médian à la fin du siècle. De plus, la moyenne des plus fortes températures augmente nettement, mais la variance semble diminuer, ce qui pourrait avoir d'importantes conséquences en terme de gestion des risques.

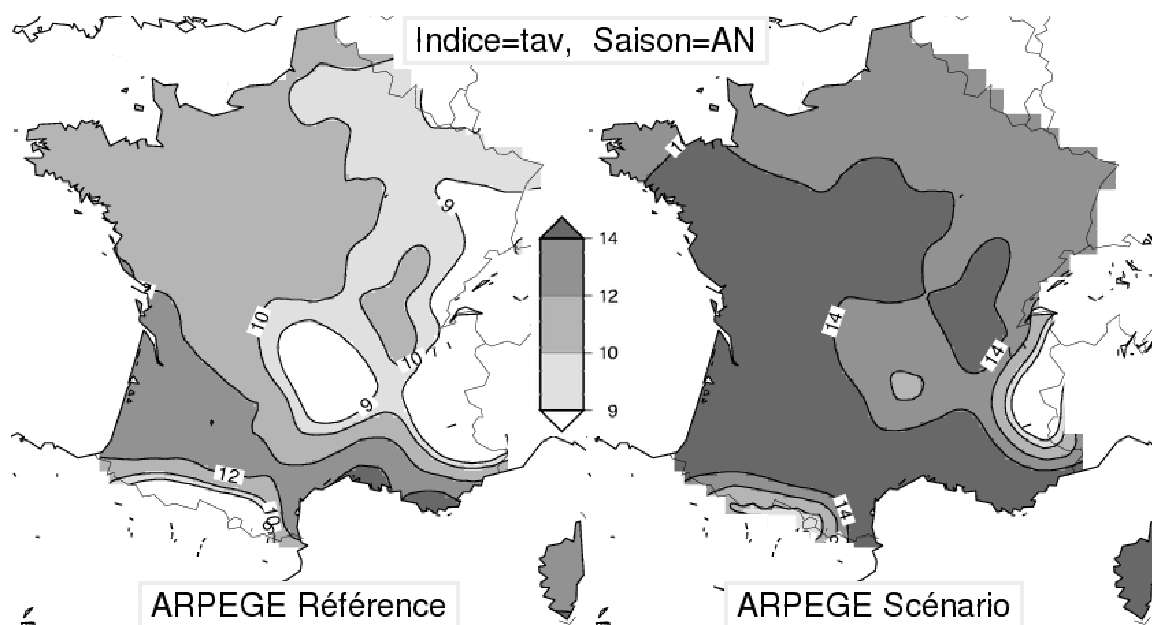


Figure 1-22. Comparaison entre les températures moyennes annuelles des périodes 1960-1989 et 2070-2099 issues d'un modèle climatique régional, avec le scénario d'émissions A2.

Source : Projet IMFREX (<http://medias.dsi.cnrs.fr/imfrex/web/>).

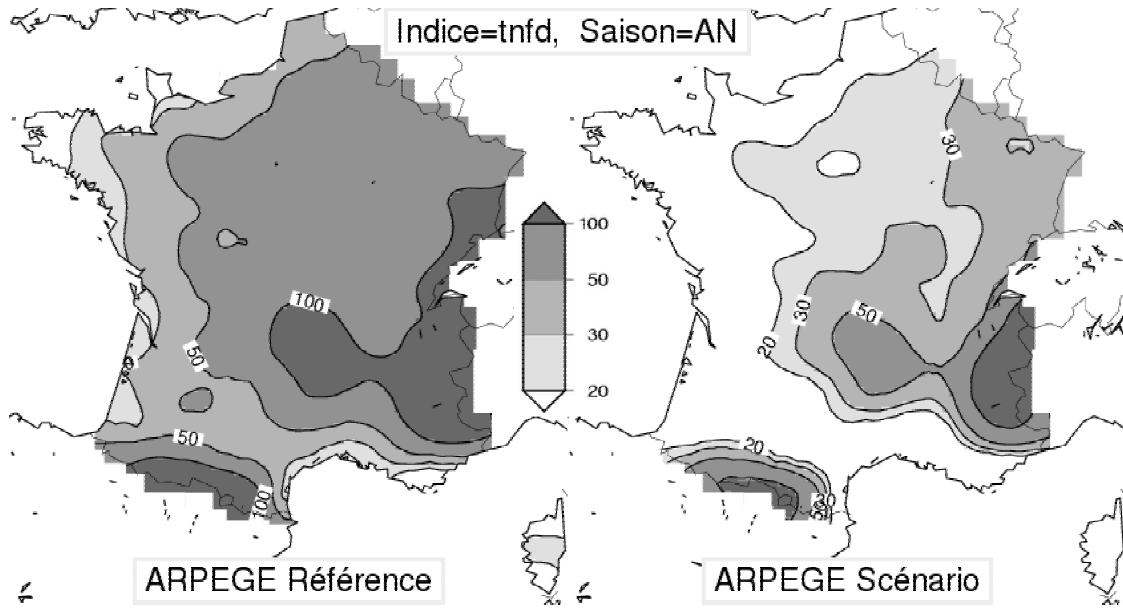


Figure 1-23. Comparaison entre le nombre de jours de gel des périodes 1960-1989 et 2070-2099 issues d'un modèle climatique régional, avec le scénario d'émissions A2. Source : Projet IMFREX (<http://medias.dsi.cnrs.fr/imfrex/web/>).

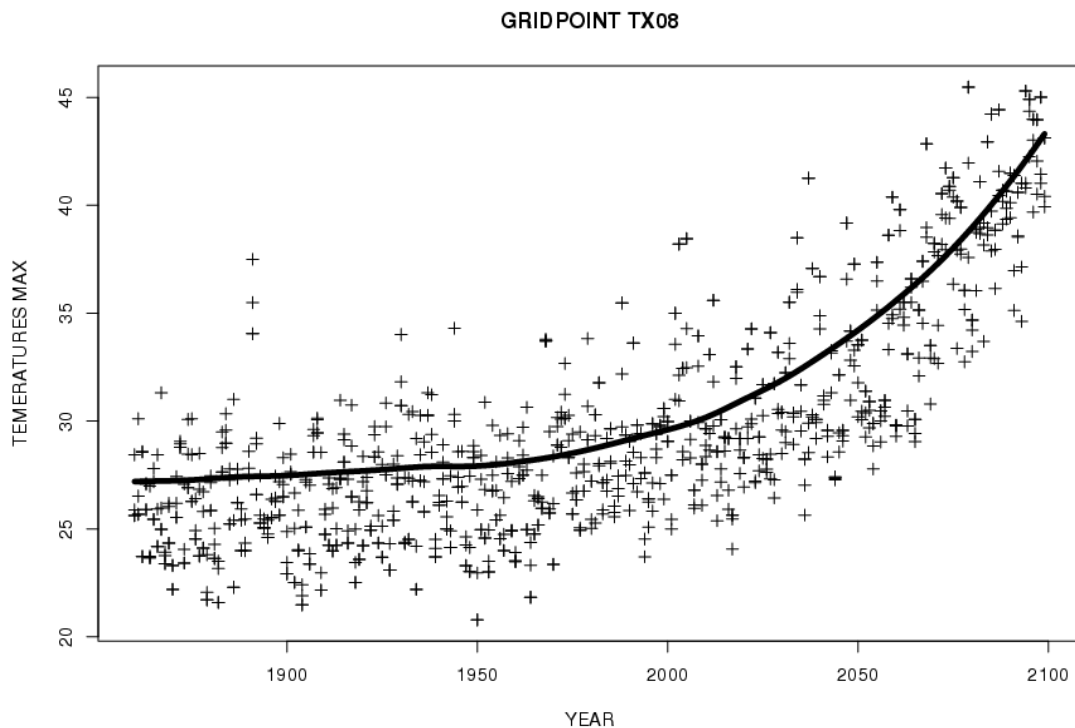


Figure 1-24. Evolution des trois plus fortes valeurs annuelles de température pour un point de grille du modèle ARPEGE, avec le scénario d'émission A2. Figure réalisée par Olivier Mestre (Météo France).

L'évolution des précipitations est plus contrastée, avec une faible évolution à l'échelle annuelle, mais une augmentation en hiver (entre 5 et 20%, Figure 1-25) et surtout une importante diminution en été (de l'ordre de 20 à 35%, Figure 1-26). Ces chiffres sont cependant très dépendants du scénario d'émission choisi, et les évolutions sont bien plus modérées avec le scénario B2. En hiver, l'augmentation du nombre de jours de précipitations supérieures à 10 mm augmente assez nettement (Figure 1-27), cette évolution n'apparaissant pas à l'échelle annuelle, sans doute à cause de la baisse de la pluviométrie en été. L'évolution des précipitations très intenses est plus incertaine, et les maxima annuels des cumuls sur 5 jours ne présentent pas de changement marqué (Figure 1-28).

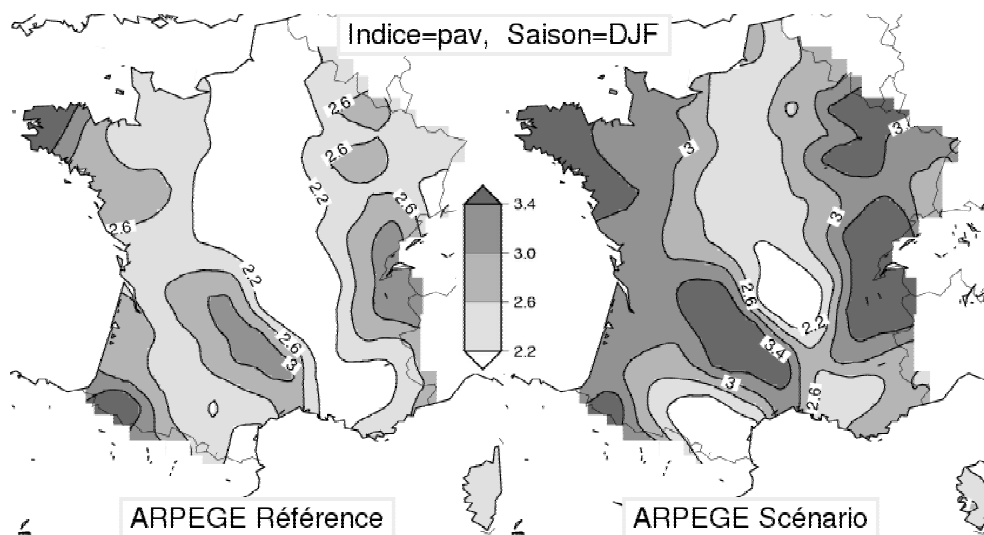


Figure 1-25. Comparaison entre les précipitations hivernales des périodes 1960-1989 et 2070-2099 issues d'un modèle climatique régional, avec le scénario d'émissions A2. Source : Projet IMFREX (<http://medias.dsi.cnrs.fr/imfrex/web/>).

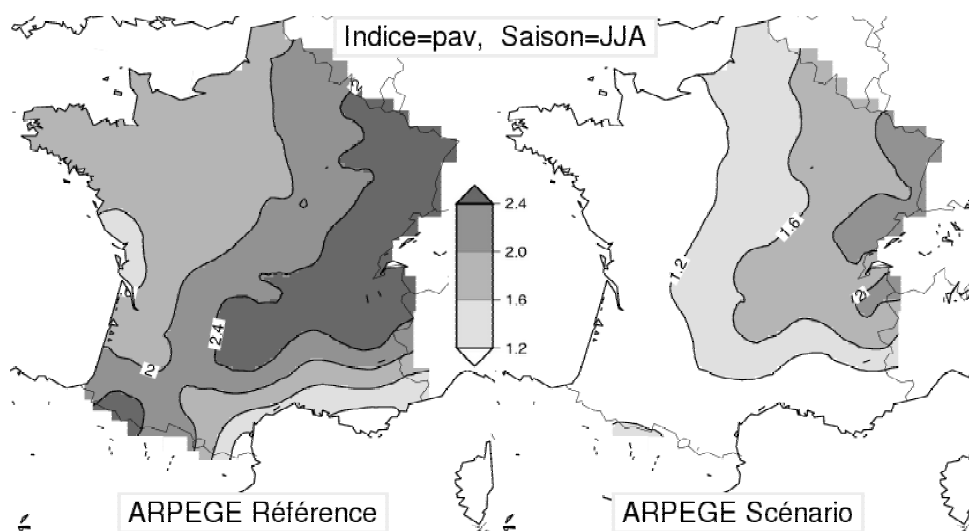


Figure 1-26. Comparaison entre les précipitations estivales des périodes 1960-1989 et 2070-2099 issues d'un modèle climatique régional, avec le scénario d'émissions A2. Source : Projet IMFREX (<http://medias.dsi.cnrs.fr/imfrex/web/>).

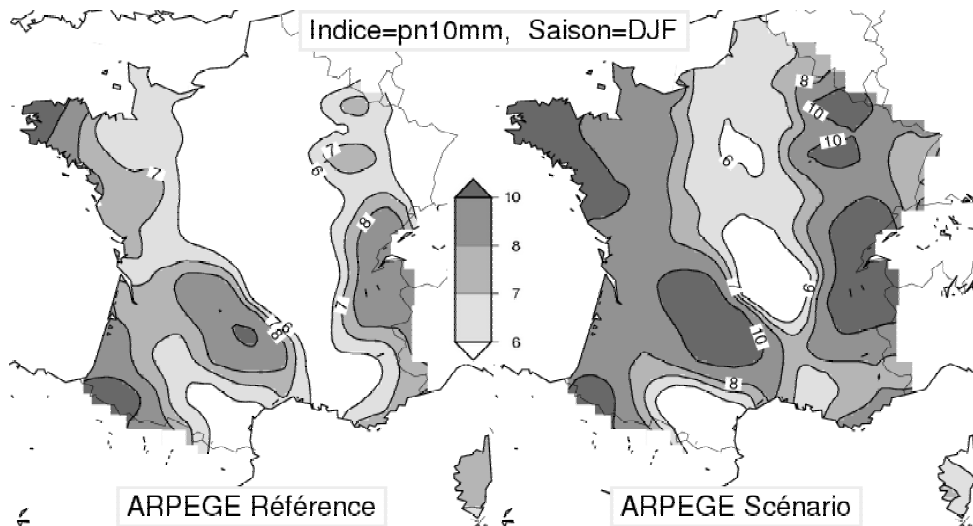


Figure 1-27. Comparaison entre le nombre de jours de fortes précipitations (>10 mm) en hiver pour les périodes 1960-1989 et 2070-2099 issus d'un modèle climatique régional, avec le scénario d'émissions A2. Source : Projet IMFREX (<http://medias.dsi.cnrs.fr/imfrex/web/>).

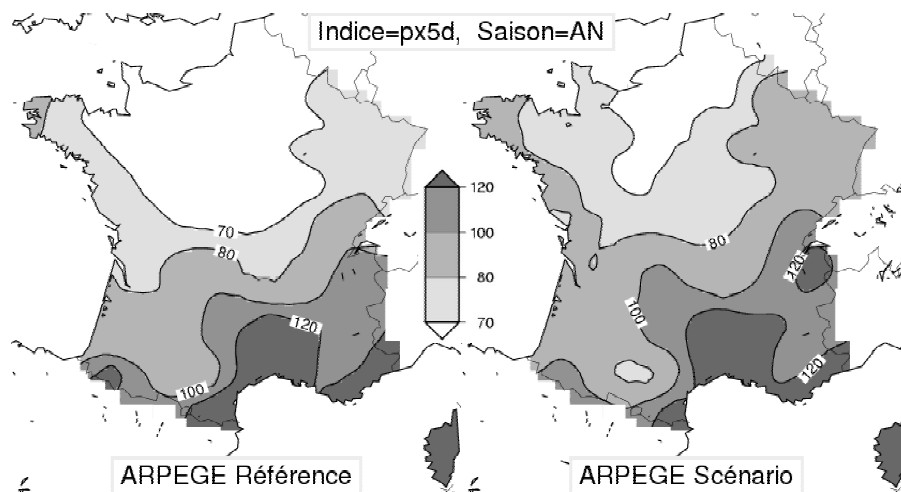


Figure 1-28. Comparaison entre les maxima annuels des cumuls sur cinq jours pour les périodes 1960-1989 et 2070-2099 issus d'un modèle climatique régional, avec le scénario d'émissions A2. Source : Projet IMFREX (<http://medias.dsi.cnrs.fr/imfrex/web/>).

Les impacts potentiels du changement climatique ont également été évalués sur trois grands bassins français, dans le cadre du programme GICC (Gestion et Impacts du Changement Climatique) du MEDD. Pour le bassin versant du Rhône (Etchevers *et al.*, 2002; Leblois, 2002), des simulations issues du modèle ARPEGE avec un scénario d'augmentation du CO₂ de 1% par an sont utilisées afin de déduire des anomalies climatiques mensuelles sous hypothèse de doublement du CO₂. Ces anomalies sont ensuite utilisées pour perturber les données climatiques observées, qui sont finalement utilisées en entrée d'un modèle hydrologique. Les résultats montrent une réponse contrastée du bassin du Rhône. Les modules semblent ainsi augmenter dans la partie nord du bassin (Saône), et diminuer dans la partie sud (Ardèche, Durance). Les affluents dont le régime possède une composante nivale marquée sont les plus sensibles au changement climatique simulé. Cependant, une étude d'incertitude basée sur l'utilisation de plusieurs modèles climatiques et hydrologiques montre que ces résultats sont à considérer avec précaution. Ainsi, les pointes de crues semblent augmenter avec le modèle ARPEGE, mais diminuer avec le modèle LMD, pour des données de forçage

identiques. L'incertitude liée au choix d'un modèle hydrologique ou à son calage semble d'un ordre de grandeur moins important. Le même type d'étude a été entrepris sur le bassin de la Seine (Ducharne *et al.*, 2003), avec le modèle ARPEGE et trois scénarios d'évolution du CO₂ (1% par an, A2 et B2). Les résultats suggèrent une intensification des contrastes saisonniers, avec des débits mensuels en baisse en été et en automne et, dans une moindre mesure, en hausse en hiver. Cependant, lorsque d'autres modèles climatiques sont utilisés, l'augmentation des débits hivernaux n'apparaît plus (certains modèles prédisant même une forte diminution), alors que les résultats semblent plus robustes pour les débits de fin d'été et d'automne. Enfin, le bassin Adour-Garonne a également été analysé (Caballero *et al.*, 2003) en utilisant sept scénarios de climat futurs, issus de quatre GCM dont deux possèdent une version haute résolution et une version basse résolution. Six de ces scénarios correspondent à un doublement de la teneur en CO₂, le septième correspondant au scénario B2 de l'IPCC. Le résultat le plus robuste concerne les étiages qui devraient s'aggraver, comme en témoigne la diminution des minima annuels et des débits mensuels estivaux. Comme cela était le cas pour le bassin du Rhône, la composante nivale du régime devrait également être très sensible, avec un glissement progressif vers un régime pluvial prépondérant. L'incertitude sur les débits hivernaux est trop importante pour obtenir des conclusions définitives.

SYNTHESE DU CHAPITRE 1

Le but de ce premier chapitre était de présenter les phénomènes physiques auxquels nous nous intéresserons au cours de cette thèse. La première partie a été consacrée à l'hydrologie. Nous avons ainsi illustré la variabilité spatio-temporelle des écoulements, et montré que le pas de temps ou d'espace considéré influait grandement sur la perception des phénomènes. La genèse des épisodes de débits extrêmes a été exposée, en mettant en avant les différences de comportement que l'on peut observer entre bassins versants. Enfin, quelques éléments relatifs à la gestion des risques hydrologiques ont été présentés. En particulier, la théorie des valeurs extrêmes est couramment utilisée pour la quantification de l'aléa, et constituera le cadre général des développements méthodologiques effectués dans cette thèse.

La seconde partie est consacrée au contexte du changement climatique. Dans un premier temps, le fonctionnement général du climat a été analysé, en insistant sur la forte variabilité naturelle des phénomènes climatiques. Le principe des modèles de circulation générale a ensuite été exposé, ainsi que les incertitudes inhérentes à ces outils numériques. Enfin, un état des lieux des connaissances actuelles sur l'ampleur observée et prédite du changement climatique a été présenté. Le signal le plus net concerne les températures, dont l'augmentation est observée de manière globale, y compris pour les valeurs extrêmes lorsque les données à pas de temps fin sont disponibles. L'évolution des précipitations est plus contrastée, certaines régions du globe allant vers une plus grande pluviométrie, d'autres enregistrant des déficits parfois importants. Enfin, bien que des évolutions aient été détectées à l'échelle régionale, il semble difficile à l'heure actuelle de mettre en avant des comportements globaux pour les écoulements dans les cours d'eau. Les prédictions issues des modèles climatiques indiquent que les évolutions observées devraient persister au cours du siècle, mais la quantification des tendances futures est limitée par de nombreuses sources d'incertitudes. Enfin, le changement climatique global ne s'est pas arrêté aux frontières de la France, et les observations montrent que le réchauffement y est aussi une réalité.

CHAPITRE 2. DONNEES ET ECHANTILLONS



Photographie par Charles Perrin.

Dans toute analyse statistique, le recueil et la critique de données constituent une étape souvent ingrate, mais qui conditionne grandement la qualité des résultats. L'analyse que nous envisageons d'effectuer portera sur un grand nombre de stations et de variables, ce qui complique évidemment la réalisation de cette première étape. Ne souhaitant cependant pas négliger cet aspect de l'analyse, nous avons mis en place une stratégie de collecte, de critique et d'échantillonnage des données de la manière la plus rigoureuse possible. L'objectif de ce chapitre est de décrire les choix effectués dans ce contexte, et de présenter les données que nous utiliserons pour l'analyse de stationnarité en France métropolitaine.

Les données de base seront des chroniques de débits calculés au pas de temps journalier. En France, la banque HYDRO regroupe une grande partie des données hydrométriques

disponibles. Plus de 3600 stations de mesure des eaux superficielles y sont recensées, auxquelles il faut rajouter quelques stations gérées notamment par EDF ou par la CNR qui ne figurent pas dans la base. Nous détaillerons dans ce chapitre l'ensemble des critères qui nous ont permis de sélectionner un échantillon de stations de mesure. Les caractéristiques générales de cet échantillon seront aussi décrites, afin d'en apprécier la représentativité par rapport aux régimes hydrologiques rencontrés en France. Nous définirons également des régions hydro-climatiques, qui nous permettront d'effectuer l'analyse de stationnarité à l'échelle régionale.

Une fois les stations sélectionnées, des variables descriptives des hautes et des basses eaux doivent être extraites des chroniques journalières. Nous verrons qu'il ne s'agit pas là d'un problème trivial : les crues et les étiages étant des phénomènes physiques complexes, il n'est pas aisé de les résumer en quelques variables. Nous exposerons donc en détail les descripteurs que nous avons décidé d'adopter. De plus, d'un point de vue plus technique, l'extraction de ces variables à partir d'une chronique journalière peut poser un certain nombre de problèmes algorithmiques. Nous décrirons donc également les procédures informatiques qui ont été implémentées.

I. Chroniques de débits journaliers

I.1. Les données hydrométriques en France

I.1.1. La mesure des débits

La caractérisation d'un phénomène physique passe nécessairement par son observation et sa mesure. Dans le cas des débits, plusieurs méthodes sont disponibles. Nous nous contenterons d'exposer celles qui sont, en pratique, le plus souvent utilisées en conditions naturelles. Une première famille de techniques consiste à estimer, sur une section bien définie, le champ de vitesses. Pour ce faire, on utilise en général des moulinets, qui permettent de mesurer une vitesse en comptant le nombre de tours effectués par une hélice, le calibrage du matériel ayant été effectué en conditions contrôlées. Sur une verticale de la section, la mesure est répétée à plusieurs profondeurs, puis plusieurs verticales sont évaluées de la même façon. On obtient ainsi une vision ponctuelle du champ de vitesse, qui, une fois intégré sur la section mouillée, donne le débit. Dans le cas de petits cours d'eau, des courantomètres remplacent parfois les moulinets. Pour les grandes rivières, un nouvel outil est utilisé depuis peu par les hydromètres, l'ADCP (Acoustic Doppler Current Profiler). Cet appareil utilise l'effet Doppler pour estimer un profil vertical de vitesses. En traversant la rivière, cette mesure est répétée tout au long de la section, ce qui permet d'obtenir le champ de vitesses, avec une résolution bien meilleure que dans le cas des mesures au moulinet. Une intégration sur la section mouillée permet encore une fois de calculer le débit. L'autre famille est constituée par les méthodes de jaugeage par dilution. Un élément traceur est injecté, à concentration connue, en un point de la rivière. Plus en aval, à une distance qui permet au traceur d'être correctement mélangé à l'eau, des prélèvements sont effectués. La concentration du traceur dans ces échantillons est alors mesurée, et la comparaison avec la concentration initiale permet d'estimer le débit. Cette technique est particulièrement adaptée aux petits cours d'eau torrentiels d'altitude, pour lesquels il s'avère difficile de trouver une section de mesure propice à l'exploration du champ de vitesse.

Ces techniques permettent d'effectuer des mesures ponctuelles, mais sont impossibles à mettre en oeuvre pour fournir une chronique continue de débits. Il est donc nécessaire de recourir à une transformation hauteur/débit, matérialisée par une courbe de tarage. Pour tracer cette courbe, on effectue quelques mesures conjointes de débit et de hauteur, et on ajuste une courbe à ce nuage de points. Cette transformation permet de passer d'une chronique de

hauteurs, mesurées quasiment en continu, à une chronique de débits. Dans cette chaîne de traitements, plusieurs sources d'incertitudes apparaissent :

- ✓ Incertitude métrologique liée à la mesure de la hauteur. Elle peut généralement être considérée comme mineure, comparativement aux autres sources d'erreurs potentielles.
- ✓ Incertitude métrologique et opératoire liée aux mesures de débits. Ce type d'erreur peut être assez important pour les débits de crues et d'étiages. Par exemple, en basses eaux, les faibles vitesses d'écoulement sont parfois difficiles à mesurer, et le développement d'algues et de végétation peut modifier temporairement la relation hauteur-débit.
- ✓ Incertitudes liées à l'ajustement d'une courbe de tarage : plusieurs types de courbes sont ajustables, et plusieurs méthodes d'ajustement peuvent être utilisées. Les lois de l'hydraulique permettent de guider cette étape.
- ✓ Incertitude liée à l'extrapolation : par nature, les événements extrêmes sont rares, il n'est donc pas possible de jauger une crue décennale tous les ans (Bois, comm. pers.). De plus, effectuer un jaugeage en crue est une opération délicate, voire périlleuse (pour le matériel et les hommes). Ainsi, les débits très faibles ou très forts peuvent ne jamais avoir été jaugés, ce qui conduit, lors de la survenue d'un tel événement, à calculer le débit par une extrapolation éloignée du domaine d'observation, d'où un important risque d'erreur.

L'extrapolation de la courbe de tarage est un des problèmes centraux de la métrologie des débits. Une étude récente a montré que sur un échantillon de 325 stations hydrométriques principalement situées dans le bassin administratif Rhône-Méditerranée-Corse, plus de la moitié des stations ne comportait pas de jaugeages effectués au-delà d'une crue de période de retour 2 ans (Lang *et al.*, 2006). De plus, les lois de l'hydraulique indiquent que la relation hauteur-débit peut ne pas être univoque en crue : à une même hauteur correspond deux débits différents en crue et en décrue, lorsqu'il y a déformation de l'onde de crue, ou lorsqu'un ouvrage hydraulique en aval est progressivement noyé. Il conviendra donc, dans les analyses ultérieures, de ne pas négliger le fait que l'incertitude associée aux débits extrêmes est particulièrement importante. De plus, dans le cadre de la recherche de changements, une rupture dans la relation hauteur/débit doit toujours être envisagée. Il est en effet difficile de garantir la validité d'une telle relation sur le long terme. Parmi les causes possibles d'évolution de cette relation, citons les changements d'extrapolation suite à un jaugeage en très hautes ou très basses eaux, une modification de la section (envasement ou creusement suite à une crue), ou des biais de mesure liés à des déplacements d'échelle ou de station. Les métrologues tentent de suivre ces modifications et d'adapter les barèmes en conséquence, mais le temps nécessaire à l'établissement d'une courbe de tarage suffisamment jaugée limite la perception de ces détarages. Enfin, certains événements plus ponctuels peuvent venir perturber la relation hauteur/débit. Parmi les stations hydrométriques que nous avons évaluées au cours de ce travail, on peut citer l'effondrement des ponts de Lyon durant la seconde guerre mondiale, ou de manière plus anecdotique, la présence de castors construisant leur barrage dans la section de mesure.

I.1.2. Quelques repères historiques

L'hydrométrie française depuis le XIX^e siècle

Les hommes ayant été de tout temps confrontés aux effets dévastateurs liés aux extrêmes hydrologiques, des descriptions d'événements remarquables datant de plusieurs siècles, et parfois très détaillées, peuvent être retrouvées. L'œuvre de Maurice Champion (1858-1864) est une référence en la matière, puisqu'elle décrit les plus grandes crues en France depuis le

VI^e siècle. C'est au milieu du XIX^e siècle que l'hydrologie et l'hydrométrie commencent à se structurer en France, en réaction aux nombreuses inondations catastrophiques qui se succèdent sur les grands fleuves (Loire en 1846, 1856 et 1866 ; Seine en 1876 ; Rhône en 1840 et 1856 ; Garonne en 1855 et 1875 ...). Dans une circulaire datant du 26 juillet 1856, le Ministre de l'Agriculture, du Commerce et des Travaux Publics écrit : « *L'immense étendue des désastres causés par les dernières inondations, le retour en quelque sorte périodique de ce fléau, qui semble s'aggraver sans cesse, imposent à l'administration des travaux publics des devoirs impérieux. (...). Rétablir les lieux dans leur état primitif, ce serait laisser le pays exposé à de nouvelles catastrophes semblables à celles que nous déplorons* » (cité par Champion (1858-1864)). Cette volonté d'évolution de l'hydrologie française va se traduire par d'importants efforts de collecte de données de débit, et la naissance des premiers services d'annonce des crues à partir de 1875, de services dédiés à l'hydrométrie aux Ponts et Chaussées et d'un service des Grandes Forces Hydrauliques, chargé de dresser l'inventaire du potentiel hydraulique français. Un premier réseau de mesure de débit est mis en place sur les grands cours d'eau français, avec un développement progressif jusqu'en 1914. La première guerre mondiale portera un important coup d'arrêt à cette activité, et l'entre-deux-guerres sera marqué par un quasi-abandon des mesures hydrométriques, à l'exception de relevés de cotes pour l'annonce des crues. Les débits relatifs à cette période dont nous disposons aujourd'hui sont donc des reconstitutions ultérieures basées sur des chroniques de hauteurs. A partir du début des années 1950, c'est l'essor de l'hydroélectricité qui permettra un nouveau développement des réseaux de mesure, sous l'égide d'Electricité De France (EDF), des Circonscriptions Electriques (CE) du Ministère de l'Industrie, et de la Compagnie Nationale du Rhône (CNR). La couverture spatiale de ces réseaux hydrométriques reste limitée aux bassins présentant un potentiel pour l'hydroélectricité. Les Services Régionaux d'Aménagement des Eaux (SRAE) voient également le jour dans les années 1960, permettant d'étendre le réseau de mesures à des cours d'eau plus petits, avec des objectifs liés à l'irrigation notamment. En 1968, une importante réorganisation des services de l'Etat liés à l'hydrologie est effectuée : la France est découpée en six grands bassins, et les réseaux existants sont centralisés dans chacun d'eux au sein d'un Service Hydrologique Centralisateur (SHC). Puis, une nouvelle réorganisation est effectuée en 1991, avec la création des Directions Régionales de l'Environnement (DIREN), qui vont regrouper, au niveau régional, les réseaux gérés par les SHC et les SRAE. Enfin, une réorganisation est aujourd'hui en cours suite à la création en 2003 du service central d'hydrométéorologie et d'appui à la prévision des inondations (SCHAPI), ainsi que de 22 services de prévision des crues (SPC), en remplacement des 52 services d'annonce des crues (SAC) préexistants. Ces derniers services cohabitent avec les DIREN, avec, à terme, l'existence de deux réseaux de mesure, le premier plutôt dédié à la prévision des crues, le second à l'hydrométrie générale. Signalons enfin que d'autres organismes peuvent également gérer des données hydrométriques, notamment les compagnies d'aménagement de bassins, les services de navigation ou les DDE, et les instituts de recherche.

Implications sur les chroniques de débits

Ce bref aperçu historique montre que l'organisation de l'hydrométrie en France a connu un certain nombre de bouleversements depuis 150 ans. Avoir conscience de cette évolution permet de mieux comprendre l'origine des données, et surtout d'envisager les conséquences sur leur validité à long terme. Tout d'abord, les objectifs des différents services gestionnaires de réseaux de mesure étaient divers. Ainsi, les stations dédiées à l'annonce des crues étaient plutôt localisées sur des sites à enjeux économiques ou humains importants, et non sur des sections permettant d'assurer la qualité de la transformation hauteur/débit. Inversement, les gestionnaires d'électricité s'intéressaient, pour la gestion de leurs ouvrages, à la mesure en

elle-même du débit, et s'attachaient donc à utiliser des sections propices à la construction de courbes de tarage précises. De plus, ces réorganisations successives posent le problème de la perte d'expertise des personnels chargés des mesures et de la construction des courbes de tarage, et dans ce domaine, l'expérience et la connaissance du terrain sont certainement les outils les plus importants. Enfin, l'éclatement géographique des gestionnaires est également un frein à une critique efficace des données. La création de la banque centralisatrice HYDRO permet d'accéder aux valeurs de débit de manière très simple, mais l'accessibilité aux méta-données (historique des courbes de tarages, déplacements, changements de matériel, faits marquants, etc.) reste problématique. Ce sont pourtant ces informations qui permettent de se faire une opinion sur la fiabilité des débits utilisés.

I.1.3. Critères de sélection des stations

Il n'est pas nécessaire, ni même souhaitable d'utiliser la totalité des stations hydrométriques françaises pour notre étude : la masse de données poserait inévitablement des problèmes de gestion, et empêcherait de réaliser un réel travail de critique de la qualité des données. Il a donc fallu sélectionner un sous-échantillon de stations, si possible représentatif de la variabilité hydrologique observée en France, et répondant à un certain nombre de critères que nous allons détailler dans les paragraphes suivants.

Ancienneté

Les phénomènes extrêmes présentent par nature une variabilité très importante. Parvenir à détecter une évolution sur de telles variables réclame donc soit un changement marqué, soit une durée d'observation suffisante. Plusieurs auteurs (Chiew et McMahon, 1993; Wetherald et Manabe, 1999; Radziejewski et Kundzewicz, 2004) ont ainsi étudié la « détectabilité » de changements dans des séries hydrométéorologiques. Pour donner un ordre de grandeur, environ 50 années d'observations sont nécessaires pour détecter un changement d'amplitude 40% par rapport à la moyenne. Même si ces performances peuvent être améliorées en utilisant des tests plus puissants ou une approche régionale, ce résultat montre qu'il est illusoire d'espérer détecter un changement modéré dans des séries trop courtes. Dans la plupart des études de stationnarité, des séries de 30 à 60 années de données au minimum sont utilisées. En France, les réseaux de mesure se sont considérablement développés au début des années 60. Nous avons donc imposé un critère d'ancienneté de 40 années au minimum, qui permettra de travailler avec un nombre suffisant de stations. En contrepartie, le pas de temps utilisé sera journalier, car les chroniques à pas de temps variable sont généralement bien moins anciennes. Pour les bassins versants les plus rapides, la description des événements de crue sera donc assez imprécise, notamment pour les variables décrivant des hydrogrammes.

Caractère non influencé

Le climat n'est pas le seul facteur influant sur le régime des eaux. Tout d'abord, de nombreuses rivières sont aménagées et volontairement influencées, afin de satisfaire à différents besoins : production d'électricité, soutien d'étiages, écrêtement de crues, prélèvement pour l'irrigation, etc. Dans le meilleur des cas, ces influences sont quantifiables, et des débits naturels reconstitués (c'est à dire les débits que l'on observerait en l'absence de l'aménagement) sont calculés. La qualité de ces reconstitutions est parfois douteuse, surtout en basses eaux. Par exemple, si le calcul prend en compte des variations de réserve dans le cas d'un barrage à grande capacité, une faible variation de hauteur peut se traduire par une forte variation de débit, rendant la mesure très sensible à des biais météorologiques comme la présence de vaguelettes créées par le vent, par exemple. D'autres influences sont plus diffuses à l'intérieur du bassin versant, et sont difficilement quantifiables. On peut citer notamment les changements d'occupation du sol ou les prélèvements sauvages.

Parvenir à discriminer ce qui relève d'une influence climatique de ce qui relève d'influences anthropiques risque fort d'être impossible, étant donné les difficultés à ne serait-ce que détecter un changement, comme expliqué précédemment. Nous avons donc choisi d'utiliser uniquement des séries réputées pas ou faiblement influencées (suivant la terminologie de la banque HYDRO), ou des débits naturels reconstitués réputés de bonne qualité. Néanmoins, une difficulté supplémentaire est liée au fait qu'un cours d'eau peut n'être influencé que pour une partie de son régime. Par exemple, un ouvrage dédié à l'écrêtement des crues aura une influence sur la distribution des hautes eaux, mais quasiment aucune sur le module, puisque à l'échelle de l'année hydrologique, toute l'eau stockée est rendue à la rivière. De même, des prélèvements peuvent influencer les étiages, mais être totalement négligeables pour les crues. En conséquence, les stations sélectionnées ne seront pas systématiquement étudiées en hautes et en basses eaux, mais uniquement pour les variables que l'on pourra considérer comme exemptes de toute influence significative directe.

Qualité

Nous avons souhaité écarter les stations pour lesquelles la qualité de la mesure nous semblait incertaine. Un tel jugement est forcément subjectif en l'absence d'un indice de qualité hydrométrique issu d'un calcul d'incertitude. Cependant, un certain nombre d'informations peuvent conduire à douter de la validité de certains sites de mesure. Les principales sont listées ci-après :

- ✓ Nombreux détarages ;
- ✓ Peu de jaugeages, ou absence de jaugeages sur les extrêmes de la courbe de tarage ;
- ✓ Grande dispersion des débits jaugés ;
- ✓ Section de mesure changeante (envasement ou creusement) ;
- ✓ Absence de contrôle hydraulique ;
- ✓ Nombreux changements de gestionnaire ;
- ✓ Nombreux changements d'appareils de mesure ;
- ✓ Présence de végétation.

Afin d'écarter les stations « à problème », un tour de France des gestionnaires a été effectué. Une discussion avec les personnels chargés de la mesure et de l'élaboration des courbes de tarage, ainsi qu'une inspection de l'historique de la station de mesure, dans lequel est notamment répertorié l'ensemble des courbes de tarage, permet d'estimer la qualité des mesures. Notons qu'un tel jugement peut encore une fois varier sur les différentes parties du régime de la rivière : une section de mesure peut être fiable en hautes eaux et inutilisable en étiages, et inversement. Dans la pratique, Philippe Bois (LTHE) a rendu visite à une grande partie des DIREN pour cette étape de critique des données, Eric Sauquet (Cemagref) s'est occupé des stations gérées par la CNR, et les stations EDF ont été évaluées conjointement par Anne Dupeyrat (EDF), Philippe Bois, Christelle Prudhomme (CEH Wallingford) et moi-même.

I.2. Description des données utilisées

I.2.1. Description générale

Chroniques

A l'issue de l'étape de critique des données, 195 chroniques de débits journaliers ont été retenues. Parmi celles-ci, 179 sont utilisables en hautes eaux et 128 en étiages. Nous avons également volontairement introduit quelques stations influencées, afin de vérifier que les tests statistiques que nous utiliserons seront bien capables de détecter des changements avérés. 106

stations sont gérées par les DIREN, 75 par EDF et 14 par la CNR. La répartition des stations sur le territoire est représentée sur la Figure 2-1. Cette répartition n'est évidemment pas homogène : la densité de points est plus importante en zone montagneuse qu'en zone de plaine.

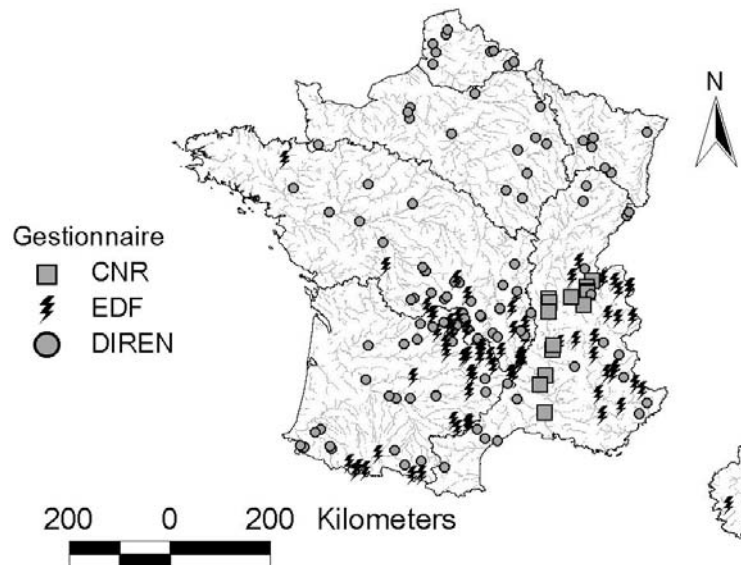


Figure 2-1. Localisation des 195 stations étudiées

Rares sont les stations qui possèdent une chronique continue de débits journaliers depuis leur mise en service : la présence de lacunes est ainsi susceptible de diminuer le nombre d'années réellement utilisables pour l'analyse. L'histogramme présenté en Figure 2-2a montre néanmoins que dans la plupart des cas, la taille des chroniques reste supérieure à 40 ans. Trois stations présentent un grand nombre de lacunes, conduisant à des séries de durées réelles inférieures à trente ans. Elles sont situées sur des cours d'eau dont le régime est glaciaire, pour lesquels seules les données estivales sont disponibles. Ceci n'est pas une limitation, puisque les crues glaciaires se produisent exclusivement durant l'été. Treize stations auront une durée réelle comprise entre 30 et 40 ans.

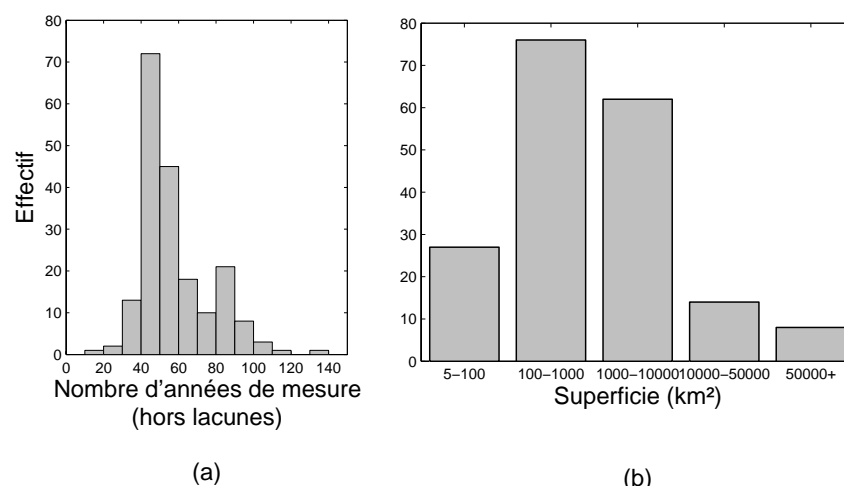


Figure 2-2. Histogrammes du nombre d'années de mesures réellement disponibles (a) et des superficies des bassins versants (b).

Diversité des bassins versants

Tout d'abord, des bassins versants de toutes tailles ont été utilisés, comme le montre la Figure 2-2b. Les bassins versants de taille modérée (moins de 1000 km²) sont sur-représentés, d'une part car ils sont naturellement plus nombreux, d'autre part car il sont plus souvent exempts d'influences significatives. Cependant, certains très grands bassins pourront être utilisés, en crue notamment, malgré des influences avérées sur les débits courants. C'est le cas par exemple du Rhône à Beaucaire (96500 km²), qui est situé en aval de nombreux aménagements, dont l'influence demeure négligeable pour les fortes crues.

Parmi ces bassins versants, on retrouve la plupart des régimes présents sur le territoire français. La Figure 2-3 présente quelques diagrammes de débits mensuels inter-annuels : on y retrouve des représentants des régimes pluvial et nival, ainsi que des différents sous-régimes (cévenol) et régimes mixtes.

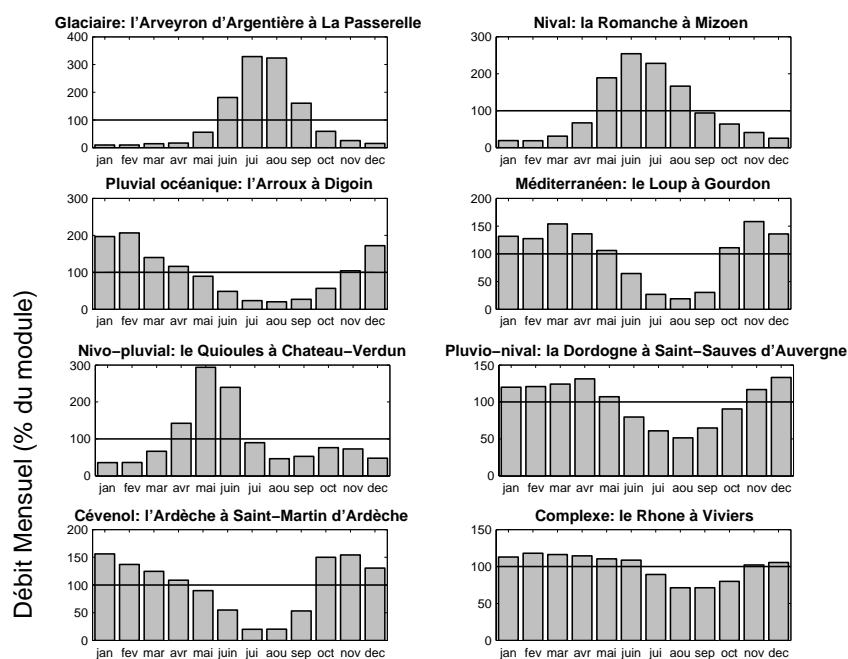


Figure 2-3. Exemples de stations représentatives des différents régimes hydrologiques français.

I.2.2. Régions hydro-climatiques

De nombreuses classifications climatiques existent afin de décrire les principales caractéristiques des grands climats observés dans le monde. Elles sont par exemple basées sur les diagrammes de températures et de précipitations mensuelles, ce qui permet de différencier assez simplement un climat océanique tempéré d'un climat Méditerranéen, par exemple. Dans le contexte de cette thèse, nous travaillons à une échelle relativement petite (la France Métropolitaine), et notre objet d'étude, les écoulements, dépend certes de variables climatiques, mais de manière fortement non linéaire. De plus, il est important de garder en mémoire la raison pour laquelle nous souhaitons construire des régions homogènes : il s'agira de vérifier la cohérence d'éventuels changements à l'intérieur d'une même zone. En conséquence, la classification devra être effectuée en répondant à certains critères, que nous allons détailler.

Un premier impératif est qu'à l'intérieur d'une même région, les conditions climatiques conduisant aux événements hydrologiques extrêmes soient comparables. Par exemple, les crues ne sont pas liées aux mêmes phénomènes météorologiques sur le pourtour méditerranéen ou en Bretagne. De plus, en cas de changement climatique avéré, rien ne

prouve que ces deux phénomènes distincts évolueront de la même façon. Il n'y a donc pas de raison de rechercher une cohérence dans ce cas. Dans un second temps, il conviendra de distinguer les différents régimes rencontrés à l'intérieur d'une même zone climatique. Encore une fois, la genèse des événements de crues peut être complètement différente sur deux rivières méditerranéennes, l'une à régime pluvial et l'autre à dominante nivale par exemple. Les effets d'un changement climatique pouvant être très variables suivant les saisons, il n'y a encore une fois aucune raison de penser que des événements liés à des précipitations intenses automnales et d'autres liés à un cumul de précipitations hivernales réagissent de manière cohérente. On voit donc qu'il sera nécessaire de croiser deux classifications pour la définition de nos régions hydro-climatiques : l'une basée sur des variables climatiques (les précipitations), l'autre utilisant la saisonnalité des événements hydrologiques.

Classification basée sur les pluies

Nous allons reprendre une classification réalisée par Météo France (Champeaux et Tamburini, 1995). Elle est basée sur les précipitations journalières entre 1971 et 1990 de 1976 postes du Réseau Climatologique d'Etat (RCE), et a été effectuée avec une classification hiérarchique ascendante utilisant la distance euclidienne et le critère de Ward pour l'agrégation des classes. Les auteurs fournissent deux zonages différents : le premier comprend 51 classes, ce qui constitue un degré de finesse trop important pour notre étude, puisque de nombreuses régions risquent alors de ne comporter qu'une ou deux stations de débits. Le second zonage comprend 12 classes, et est représenté dans la Figure 2-4.

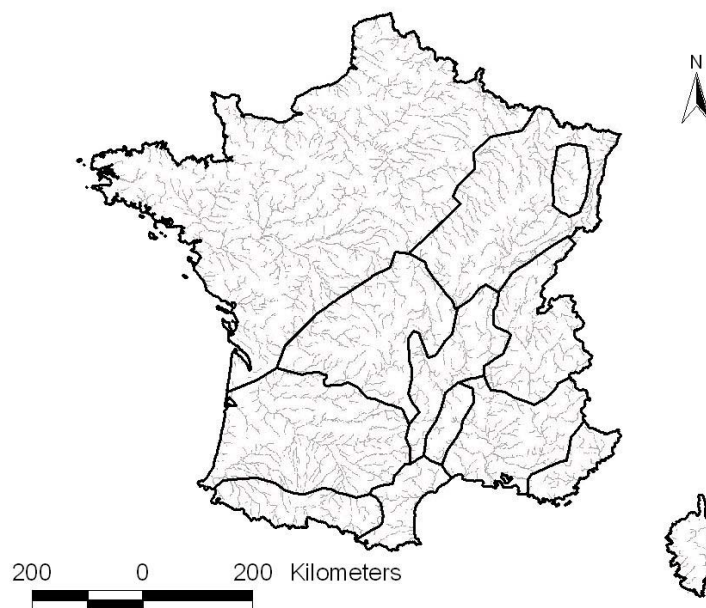


Figure 2-4. Zonage de la France en 12 classes à partir des précipitations quotidiennes, d'après Champeaux et Tamburini (1995). La Corse et le Roussillon appartiennent à la même classe.

On retrouve dans ce zonage les grands ensembles climatiques français. La partie Nord-Ouest de la France, au-dessus d'une ligne Champagne-Gironde, est caractérisée par des pluies océaniques, en général peu intenses mais fréquentes. La zone Aquitaine-Midi Pyrénées est également à influence océanique, mais est soumise à l'influence des Pyrénées, qui ont tendance à bloquer les perturbations venant du nord-ouest, augmentant ainsi les cumuls pluviométriques, dans le Pays Basque notamment. La chaîne des Pyrénées constitue également une zone pluviométrique distincte, de par son climat de type montagnard. Au nord-

est de la France, on distingue une zone d'influence continentale, au milieu de laquelle le massif Vosgien, aux précipitations plus importantes, est isolé. La partie ouest du massif central est caractérisée par des cumuls de pluie importants, notamment en automne et en hiver. La région des Cévennes ressort comme une zone bien délimitée, ce qui traduit la particularité de cette zone du point de vue pluviométrique (précipitations convectives très intenses). Entre les deux, on retrouve une zone plus abritée, avec un nombre de jours de pluie annuel assez faible. Le nord des Alpes et le Jura sont regroupés pour former une vaste zone montagnarde, aux précipitations abondantes. Enfin, le pourtour méditerranéen est découpé en trois zones, ce qui démontre l'importante variabilité spatiale des précipitations. Une première zone s'étend de la frontière espagnole jusqu'à Montpellier, et comprend étonnamment la Corse. Une seconde zone englobe le littoral Varois et la côte d'Azur. Enfin, la Provence et les Alpes du Sud sont regroupées pour former la dernière région. Ces trois régions méditerranéennes sont caractérisées par un nombre de jour de pluie annuel faible, et des précipitations automnales intenses.

Notons que le fait que la classification soit établie à partir de cumuls journaliers n'est pas sans influence sur l'interprétation de ce zonage. En effet, deux stations seront « proches » au sens pluviométrique si elles sont soumises au même type de pluie de manière relativement concomitante. C'est la raison pour laquelle les Alpes du Nord et les Pyrénées, par exemple, sont classées comme deux zones distinctes, bien qu'il s'agisse dans les deux cas d'un climat de type montagnard. C'est également ce qui explique la connexité des zones obtenues, qui est liée à une certaine continuité spatiale des perturbations génératrices de pluies. Il ne s'agit donc pas, à proprement parler, d'un zonage climatique, mais plutôt d'un zonage reflétant, en un point du territoire, à la fois le type de circulation atmosphérique et le relief. Or, un des effets majeurs du changement climatique devrait précisément être de perturber la circulation atmosphérique. Nous pensons donc que ce zonage est adéquat dans le cadre de notre étude : à titre d'illustration, il n'y a pas de raison de rechercher une cohérence entre des changements affectant des rivières nivales des Alpes et des Pyrénées, puisque ces régions ne sont pas forcément influencées par le même type de conditions atmosphériques.

Classification basée sur les régimes de crue

A partir des séries de débits journaliers, des événements de crue ont été extraits par la méthode SUP-SEUIL, en choisissant en moyenne un événement par an. Cette méthode d'échantillonnage sera décrite en détail dans la section II.1. de ce chapitre. Les dates des pics de ces événements ont ensuite été calculées. Enfin, pour chaque mois de l'année, nous avons calculé le pourcentage d'événements de crue. Nous disposons ainsi d'un tableau de 195 lignes (les stations) et 12 colonnes (les pourcentages mensuels), grâce auquel nous pouvons calculer une matrice de distance entre les stations, en calculant la distance euclidienne dans l'espace de dimension 12 défini par les pourcentages mensuels d'événements. Une classification hiérarchique ascendante utilisant le critère de Ward est ainsi effectuée sur la base de cette matrice de distances. Le dendrogramme issu de cette classification est présenté en Figure 2-5. Nous avons décidé de retenir un découpage en six classes, sur la base de la pertinence hydrologique des groupes constitués. On observe que la classification sépare très nettement les régimes à influence pluviale de ceux à influence nivale. A l'intérieur de chacune de ces deux catégories, trois classes apparaissent, et sont décrites dans la Figure 2-6. Le groupe « Pluvial pur » est constitué de rivières assez homogènes où les crues se produisent majoritairement en hiver, et pratiquement jamais en été. Le groupe « pluvial complexe », plus hétérogène, comprend des cours d'eau où la saisonnalité des crues est peu marquée. Enfin, le groupe « Méditerranéen » est caractérisé par une saison des crues automnale, avec un maximum net pour les mois d'octobre et novembre. Le groupe « Nival pur » est constitué de rivières où les crues se produisent majoritairement durant les mois de mai, juin, et juillet, avec

éventuellement quelques événements durant l'automne. Le groupe « Pluvio-Nival » présente quant à lui deux saisons de crues, l'une située au printemps (crues nivales), l'autre en automne (crues pluviales). Enfin, le groupe « Glaciaire » est constitué de trois stations où les crues liées à la fonte du glacier se produisent exclusivement durant l'été, et principalement durant juillet et août.

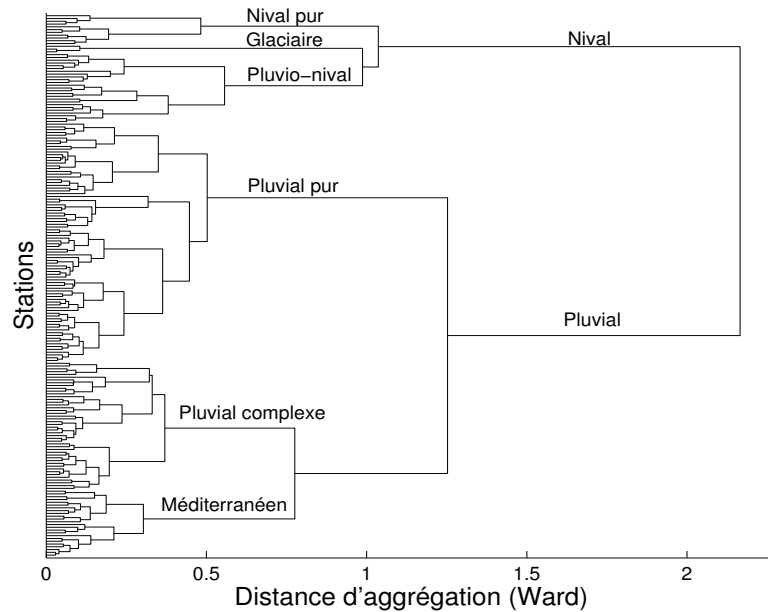


Figure 2-5. Dendrogramme issu d'une classification hiérarchique ascendante sur les dates de crue.

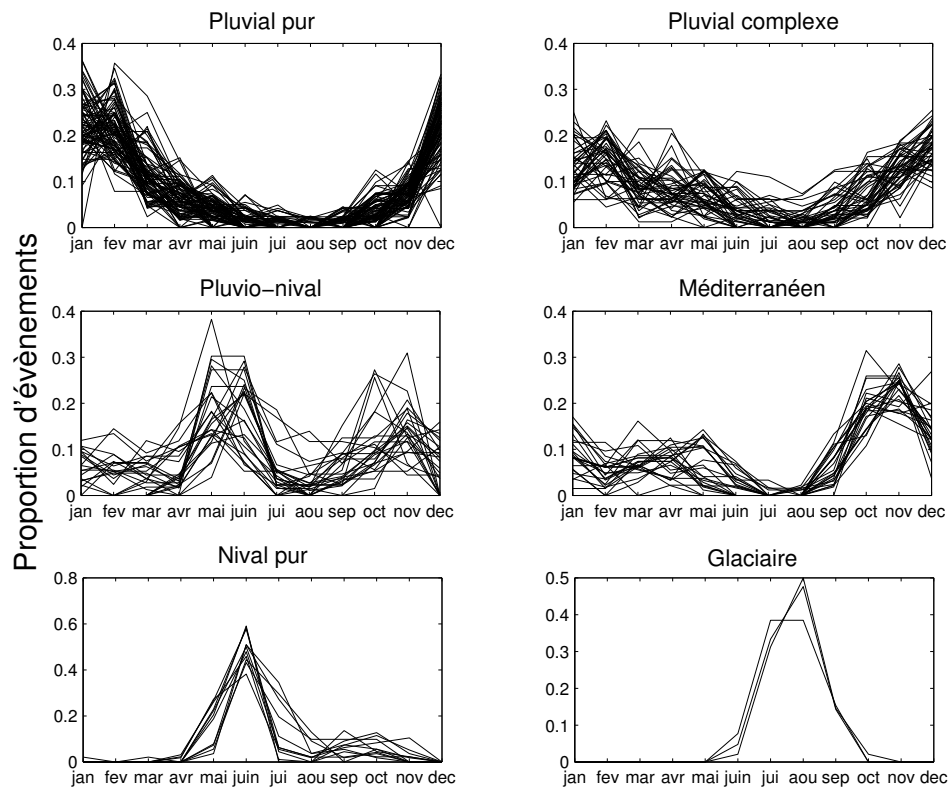


Figure 2-6. Saisonnalité des événements de crue pour les différents groupes constitués.

La répartition des différentes classes constituées sur le territoire français est décrite dans la Figure 2-7, et confirme la cohérence de la classification effectuée.

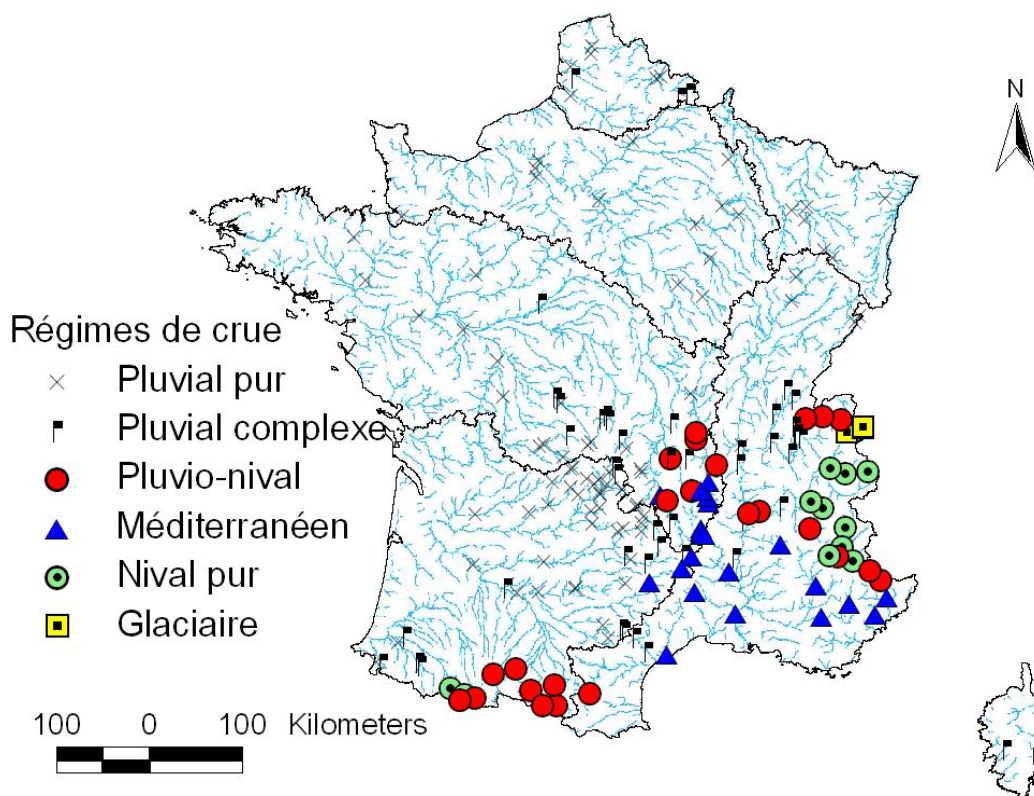


Figure 2-7. Cartographie des différents groupes constitués pour les régimes de crue.

Classification basée sur les régimes d'étiage

La méthode de classification est en tout point identique à celle utilisée pour les crues. La variable utilisée est la date du minimum annuel des débits moyens sur 7 jours. Seul les 128 stations réputées utilisables en étiages ont été prises en considération, car des aménagements de soutien d'étiage, par exemple, sont capables de modifier artificiellement la saisonnalité des basses eaux. On remarque encore une fois sur le dendrogramme de la Figure 2-8 une séparation des influences nivales et pluviales. Nous avons cette fois-ci choisi de distinguer trois classes de comportements en étiages, décrits dans la Figure 2-9. Le groupe « nival » présente une variabilité importante, et est constitué de bassins situés à des altitudes élevées, où les étiages sont dus au stockage des précipitations sous forme de neige. L'altitude du bassin contrôle la date du minimum annuel, ce qui explique que la période des basses eaux s'étende pour ce groupe d'octobre à mars. Le groupe « pluvial pur » est assez homogène, et comprend des bassins où les étiages ont lieu exclusivement entre la fin de l'été et le début de l'automne (juillet-octobre). Enfin, un dernier groupe, baptisé « complexe », possède une variabilité importante, et une saisonnalité peu marquée pour les étiages. Un tel groupe témoigne certainement d'un comportement original des bassins qui le composent, plutôt que d'une réelle unité de comportement hydrologique. Les résultats concernant ce groupe devront donc être interprétés avec prudence.

Nous avons également évalué des classifications plus fines, mais les segmentations supplémentaires consistaient à différencier des stations précoces et tardives dans chacun des groupes « nival » et « pluvial pur ». Etant donné que nous avons introduit une discrétisation

en mois pour la définition des variables classificatrices, il nous a semblé que cette séparation était artificielle, et ne reflétait pas un certain continuum entre les périodes de basses eaux à l'intérieur d'un même groupe. Le fait de ne pouvoir distinguer clairement que trois classes peut de plus s'interpréter comme une conséquence de l'importante portée spatiale du phénomène « absence de pluie ». Une classification plus fine nécessiterait d'utiliser d'autres variables, géologiques notamment, dont nous ne disposons pas dans le cadre de ce travail.

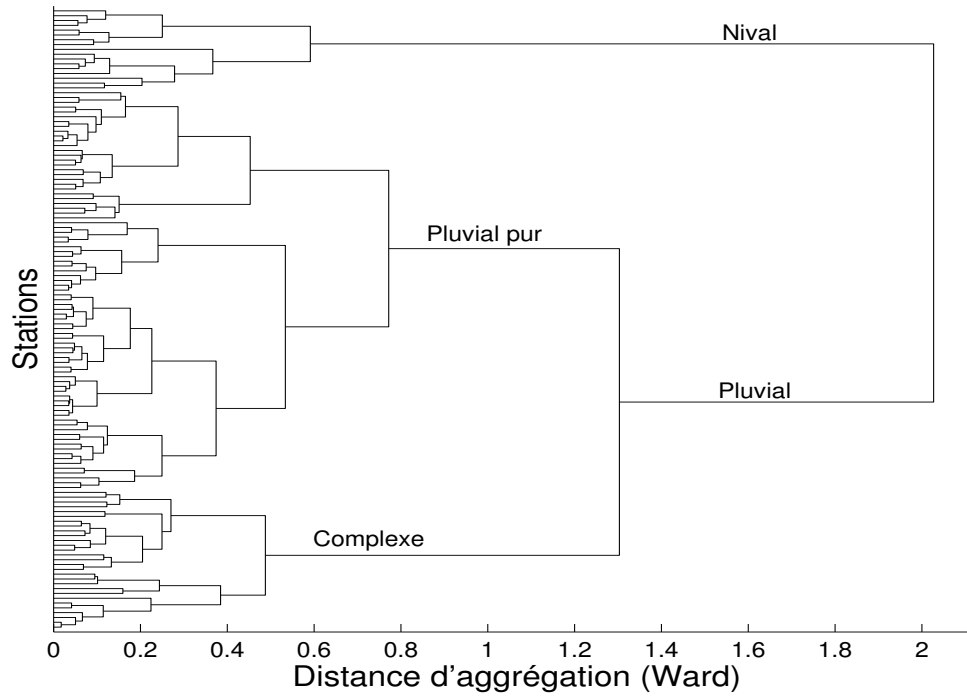


Figure 2-8. Dendrogramme issu d'une classification hiérarchique ascendante sur les dates d'étéage.

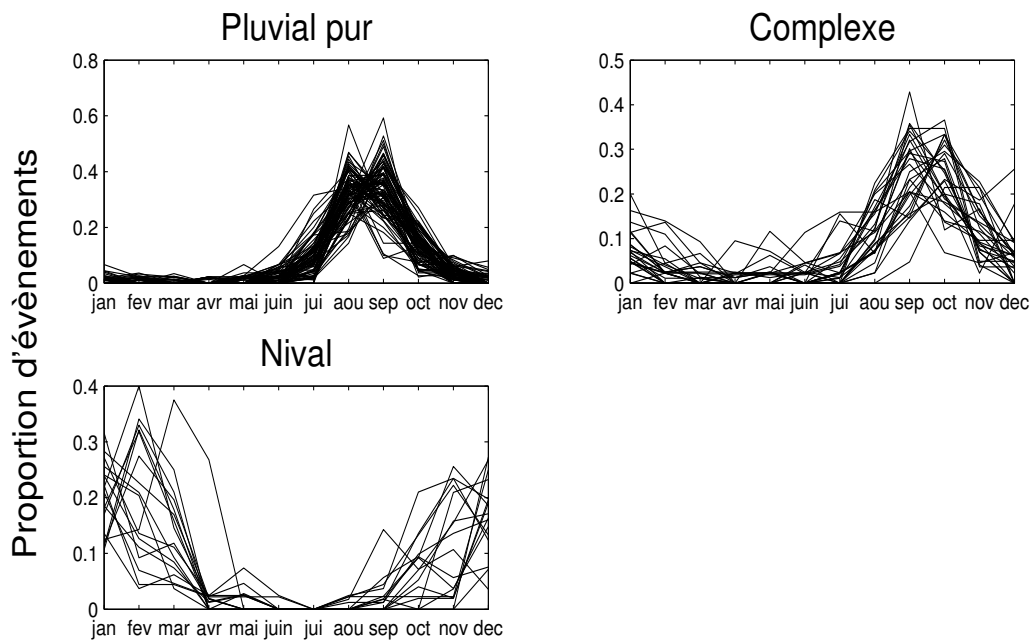


Figure 2-9. Saisonnalité des étiages pour les différents groupes constitués.

La Figure 2-10 montre la répartition géographique des différentes classes. On peut observer une certaine cohérence, même si la classification semble moins informative que dans le cas des crues. Le groupe « complexe » notamment semble se répartir sans grande homogénéité sur une large partie du territoire, ce qui incite à respect prudent sur la pertinence hydrologique d'une telle classe.

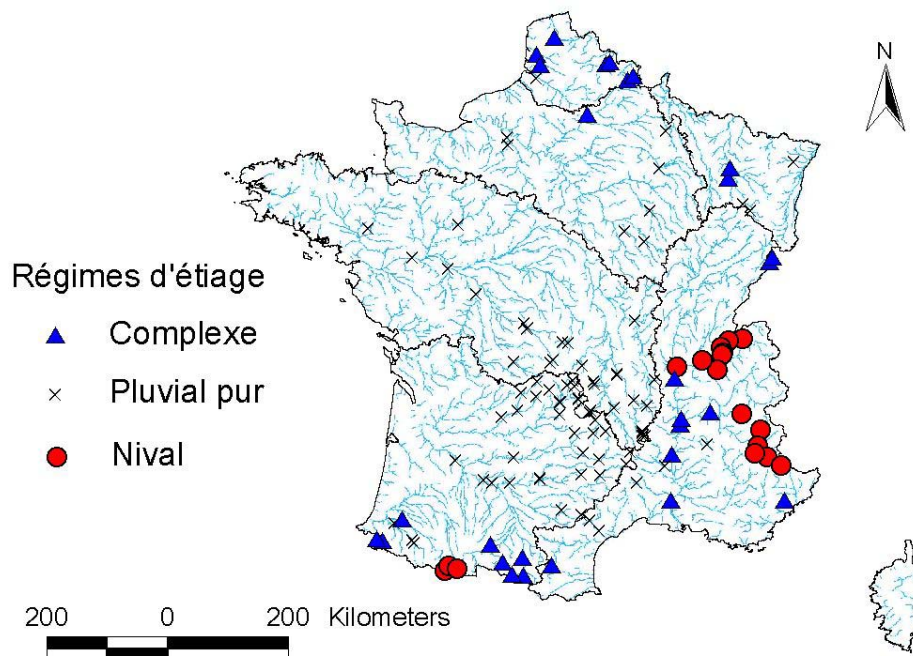


Figure 2-10. Cartographie des différents groupes constitués pour les régimes d'étiage.

Conclusion : les régions hydro-climatiques considérées

Les deux classifications obtenues doivent à présent être croisées afin de créer les différentes régions hydro-climatiques. Avec 12 régions climatiques et 6 régimes de crues, il existe théoriquement $12 \times 6 = 72$ combinaisons possibles. Cependant, réaliser un tel croisement de manière brutale n'a que peu d'intérêt. Tout d'abord, certaines combinaisons ne sont pas observées (par exemple, il n'y a pas de régimes méditerranéens en zone océanique). Ensuite, les limites des zones climatiques ne sont pas absolues, puisque le phénomène pluvieux est continu. Enfin, la localisation d'une station conditionne beaucoup moins son comportement que la localisation du bassin versant qu'elle contrôle. Evidemment, certains bassins sont répartis sur plusieurs zones climatiques, ce qui empêche de réaliser le croisement de manière automatique et déterministe. En conséquence, nous avons créé les régions hydro-climatiques « à la main », en nous basant sur les deux classifications préexistantes, sans pour autant les considérer comme des contraintes absolues. De plus, certaines grandes régions ont été découpées en sous-régions, afin de préserver un maximum d'homogénéité hydro-climatique.

Les classes ainsi constituées sont présentées dans le Tableau 2-1 et la Figure 2-11. A titre d'exemple, toutes les stations présentant des crues méditerranéennes ont été regroupées dans une même classe, considérant ainsi que le mécanisme climatique à l'origine des crues était identique (précipitations automnales liées à des flux de sud). Nous avons ainsi considéré que la partition climatique de la zone méditerranéenne fournie par Champeaux et Tamburini (1995) reflétait d'avantage la variabilité spatiale des précipitations lors de chaque événement, plutôt qu'une différence fondamentale de mécanisme climatique. Cependant, deux sous-groupes ont également été créés, afin de prendre en compte les différences hydrologiques liées à la présence de reliefs marqués (les Cévennes et les Alpes).

Régions Hydro-Climatique en crue	Sous-Régions
AlpN : Stations des Alpes du Nord à influence nivale (11)	AlpN-N : Nival pur (5) AlpN-PN : Pluvio-Nival (6)
AlpS : Stations des Alpes du Sud à influence nivale (6)	AlpS-N : Nival pur (3) AlpS-PN : Pluvio-Nival (3)
Basque : Stations pluviales du Pays Basque (6)	/
Centre : Stations pluviales de l'ouest du Massif Central (46)	Centre-P : Pluvial (34) Centre-C : Complexe (12)
Cont : Stations pluviales en zone de pluies continentales (14)	Cont-Ouest : Stations en limite de zone océanique (6) Cont-Vosges : Stations sous les versants ouest des Vosges (5)
Gla : Régimes glaciaires (3)	/
Lozère : Stations à régime complexe de la Lozère (6)	/
Med : Stations aux crues méditerranéennes (24)	Med-Cev : Rivières Cévenoles (9) Med-Alp : Rivières Alpines (6)
NivCent : Stations à influence nivale du Massif Central (6)	/
Ocean : Stations pluviales en zone océanique (19)	Ocean-Ouest : Stations proches de la façade Atlantique (8) Ocean-BP : Bassin Parisien (5) Ocean-Nord : Stations du nord (6)
Pyr : Stations à influence nivale des Pyrénées (12)	Pyr-N : Nival pur (2) Pyr-PN : Pluvio-Nival (10)
Rhône : Stations à régime complexe du bassin du Rhône (12)	/
Roussillon : Stations à régime complexe du Roussillon (4)	/
SudOuest : Stations pluviales du Sud-Ouest (11)	/
Stations non classées (15)	/

Tableau 2-1. Définition des régions hydro-climatiques en crue. Le nombre entre parenthèses indique l'effectif de chaque région ou sous-région.

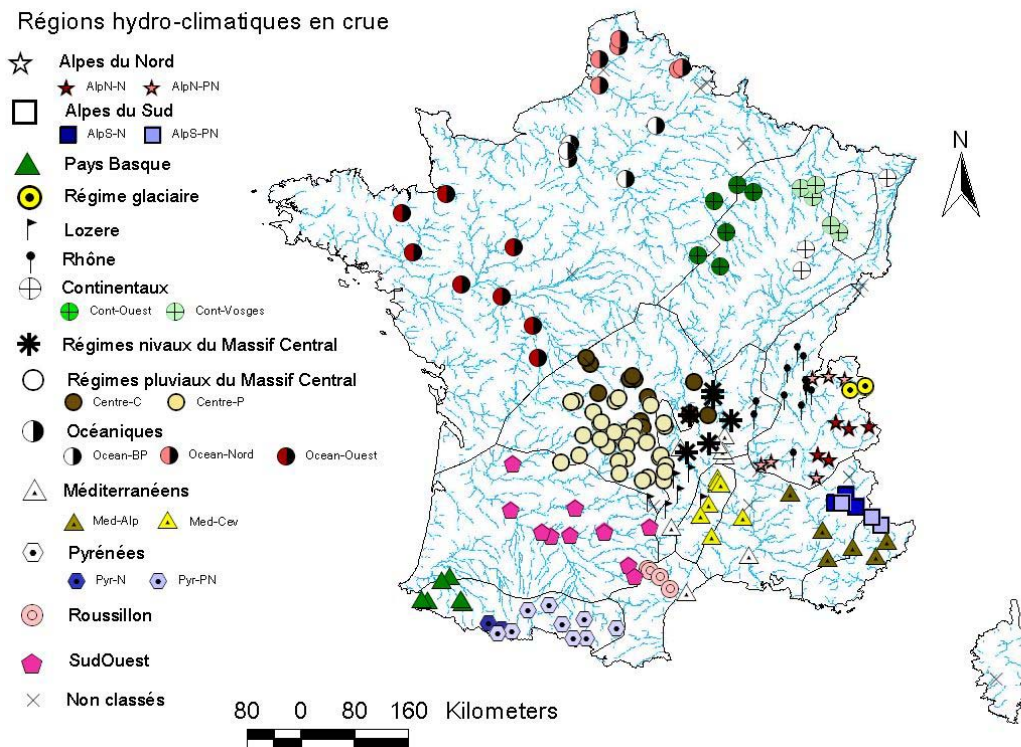


Figure 2-11. Localisation des stations étudiées suivant leur région hydro-climatique en crue.

La création de groupes hydro-climatiques en étiage a suivi la même logique que pour les crues. Le nombre de régimes d'étiage étant plus limité, les régions ont été principalement constituées sur la base du zonage pluviométrique. Quelques sous-régions ont été définies afin de différencier si nécessaire les régimes complexes des régimes purs. La description des groupes est fournie dans le Tableau 2-2, et la Figure 2-12 montre leur répartition géographique.

Régions Hydro-Climatique en étiage	Sous-Régions
AlpN : Stations des Alpes du Nord (10)	AlpN-N : Nival (8) AlpN-C : Complexe (2)
AlpS : Stations des Alpes du Sud à influence nivale (5)	/
Basque : Stations du Pays Basque (6)	Basque-P : Pluvial pur (3) Basque-C : Complexe (3)
Centre1 : Stations pluviales du Massif Central, à l'Ouest des Cévennes (14)	/
Centre2 : Stations pluviales de l'Ouest du Massif Central (23)	/
Cont : Stations en zone de pluies continentales (9)	Cont-Ouest : Stations en limite de zone océanique (4) Cont-C : Stations complexes (2)
Med : Stations méditerranéennes (11)	/
Ocean : Stations en zone océanique (16)	Ocean-Ouest : Stations proches de la façade Atlantique (5) Ocean-Nord : Stations du nord (8)
Pyr : Stations des Pyrénées (9)	Pyr-N : Nival (3) Pyr-C : Complexe (6)
Rhône : Stations à régime complexe du Rhône (4)	/
SudOuest : Stations pluviales du Sud-Ouest (10)	/
Stations non classées (11)	/

Tableau 2-2. Définition des régions hydro-climatiques en étiage. Le nombre entre parenthèses indique l'effectif de chaque région ou sous-région.

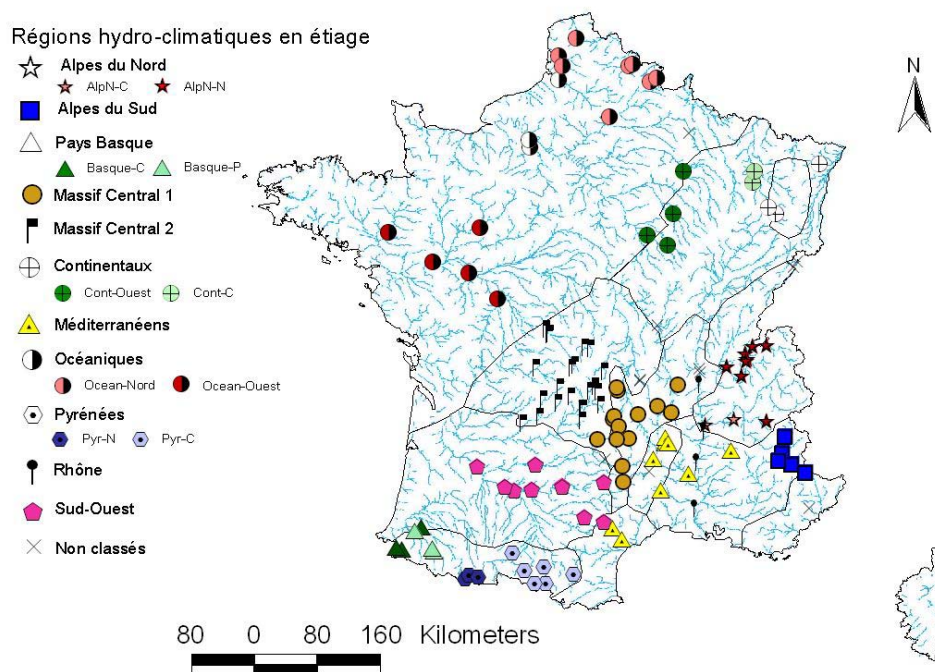


Figure 2-12. Localisation des stations étudiées suivant leur région hydro-climatique en étiage.

II. Choix de variables descriptives

Comment définir une crue ou un étiage ? Cette question, d'apparence simple, est en fait quasiment insoluble, tant les événements hydrologiques extrêmes sont des objets physiques complexes. En guise d'illustration, voici quelques définitions trouvées sur des sites Internet français :

- ✓ Météo France : « Toute élévation du niveau d'un cours d'eau — fleuve, rivière, torrent, ru... — , d'un canal ou d'un plan d'eau tel qu'un lac est appelée une crue lorsqu'elle a pour cause un apport important en eau liquide consécutif à une période de précipitations ou bien de fonte de neige ou de glace. »
(http://www.meteofrance.com/FR/glossaire/designation/863_curieux_view.jsp).
- ✓ BRGM : « Une crue correspond à l'augmentation du débit d'un cours d'eau dépassant plusieurs fois le débit moyen. »
(<http://www.brgm.fr/Risques/Antilles/guad/ino.htm>)
- ✓ Glossaire international d'hydrologie : « Montée, en général brève, du niveau d'un cours d'eau jusqu'à un maximum dont il redescend en général plus lentement. »
(<http://www.cig.ensmp.fr/~hubert/glu/aglo.htm>).

On observe à la lecture de ces définitions qu'il existe plusieurs caractérisations possibles du phénomène de crue : on peut invoquer les causes de l'événement (première définition), le caractère anormalement haut des valeurs de débit (seconde définition), ou encore la forme d'un hydrogramme (troisième définition). Dans tous les cas, cependant, ces définitions restent qualitatives.

Le cas des étiages est peut-être plus complexe encore, puisqu'il peut exister une confusion entre les termes « étiage », « sécheresse », « déficit », qui ne sont pourtant pas synonymes.

Cette complexité sémantique traduit bien la complexité physique des régimes de hautes eaux et de basses eaux. Etant donné qu'il n'existe pas de caractérisation unique des phénomènes hydrologiques extrêmes, nous avons choisi de les décrire par un ensemble de variables, dont la description constitue le corps de cette section.

II.1. Variables de crues

II.1.1. Définition des variables

Echantillonnages MAXAN et SUP-SEUIL

Deux méthodes d'échantillonnage sont principalement utilisées par les hydrologues afin d'extraire d'une chronique des descripteurs des hautes eaux. L'échantillonnage par valeurs maximales annuelles (MAXAN) consiste à sélectionner chaque année le débit observé le plus fort. L'échantillonnage par valeurs supérieures à un seuil (SUP-SEUIL) consiste quant à lui à choisir un seuil, puis à sélectionner les pointes des événements dépassant ce seuil. Dans la pratique, plutôt qu'un seuil, on se donne en général un nombre d'événements à sélectionner par an (en moyenne), et par itérations successives, on calcule le seuil conduisant à cet objectif. Ces deux techniques sont illustrées dans la Figure 2-13 (MAXAN) et la Figure 2-14 (SUP-SEUIL).

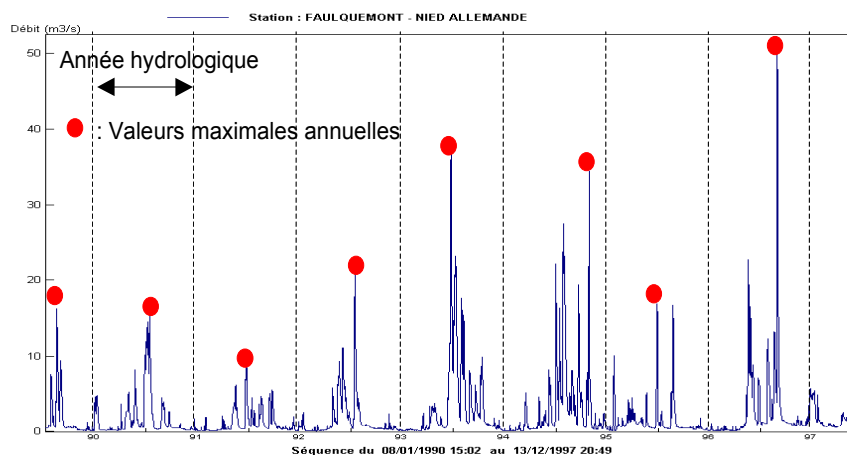


Figure 2-13. Echantillonnage de type MAXAN.

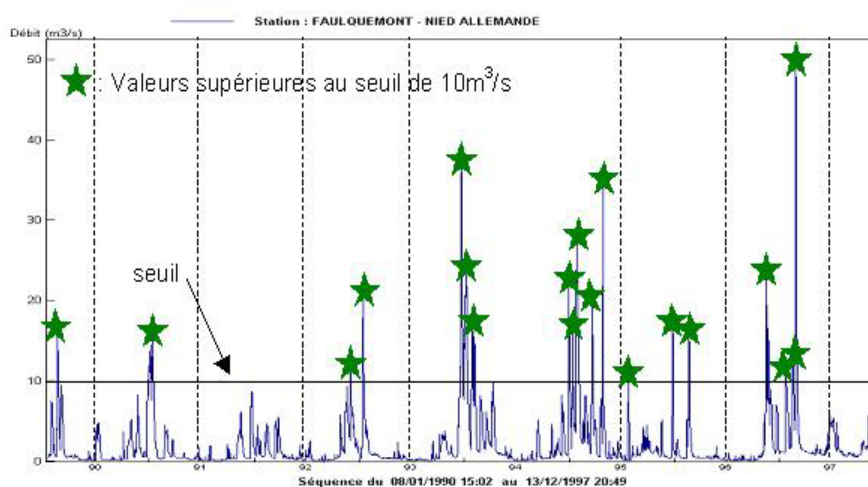


Figure 2-14. Echantillonnage de type SUP-SEUIL.

Chacune de ces méthodes présente des avantages et des inconvénients, notamment dans l'optique d'une analyse statistique où nous souhaiterions disposer d'un échantillon *iid*. La mise en œuvre de la méthode MAXAN est très simple. En ne sélectionnant qu'un unique événement par an, on s'assure de l'indépendance entre les valeurs, sauf éventuellement si une crue importante se produit à cheval sur deux années successives, conduisant, par exemple, à échantillonner un débit le 31 décembre 2000 et un autre le 1^{er} janvier 2001. Pour cette raison, on échantillonne souvent par année hydrologique, chaque année étant séparée par une période d'étiage. L'inconvénient est d'ignorer un certain nombre d'événements lors des années où beaucoup de crues se sont produites, et inversement de prendre en compte des événements peu importants lors des années peu actives. L'homogénéité de l'échantillon n'est donc pas optimale.

La technique du SUP-SEUIL est plus difficile à mettre en œuvre (Lang, 1995; Lang *et al.*, 1997; Lang *et al.*, 1999; Lang, 2000; Rosbjerg et Madsen, 2004). En effet, il faut ajouter des contraintes d'indépendance afin de ne pas échantillonner plusieurs fois le même événement hydrologique, comme dans l'exemple présenté en Figure 2-15. Cette pratique, dénommée « déclusterisation » dans le domaine des statistiques, consiste généralement à imposer une contrainte d'espacement temporel minimal entre deux pointes sélectionnées, ainsi qu'une contrainte de redescente vers un débit de base. Bien choisies, ces contraintes permettent de garantir l'indépendance de l'échantillon. Cet échantillon sera également plus homogène que

celui fournit par la méthode MAXAN. Autre avantage, il est possible d'étoffer l'échantillon en choisissant, en moyenne, plus d'un événement par an.

Ces deux méthodes permettent d'obtenir des échantillons décrivant les pointes de débit que connaît une rivière. En calculant la date des événements sélectionnés, on peut également étudier leur saisonnalité. Enfin, la méthode SUP-SEUIL permet également de s'intéresser au processus d'occurrence des crues, ce qui donne une information sur la fréquence des événements.

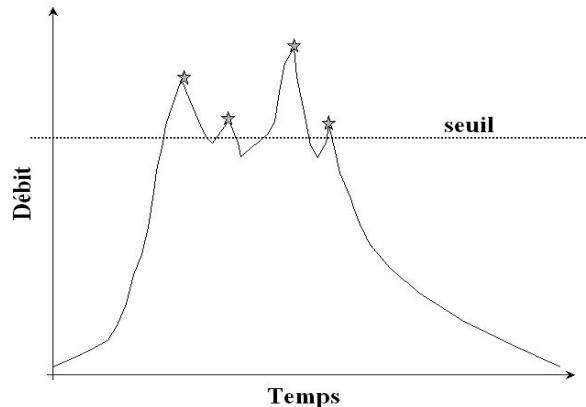


Figure 2-15. Illustration de la nécessité des contraintes d'indépendance.

Evaluation multi-durées

Dans le cas d'un cours d'eau soumis à des crues très volumiques, le débit journalier maximal n'est pas forcément la variable la plus informative : ce pic de débit peut en effet s'expliquer par une reprise des pluies, même modérée, qui vient s'ajouter sur une onde de crue pouvant s'étaler sur plusieurs semaines. Dans ce type de situation, il sera plus intéressant de s'intéresser aux volumes calculés sur une certaine durée caractéristique d , représentative de la réponse du bassin versant aux pluies. Ce raisonnement a conduit à caractériser le régime des hautes eaux non pas sur un unique pas de temps, mais sur une gamme de durée s'étalant du pas temps le plus fin (la journée calendaire dans notre cas) jusqu'à un multiple de la durée caractéristique. Cette approche est connue sous le nom de caractérisation débit-durée-fréquence (QdF) (Prudhomme, 1995; Galéa et Prudhomme, 1997; Javelle *et al.*, 1999; Javelle, 2001; Meunier, 2001; Javelle *et al.*, 2002b). Elle généralise l'approche intensité-durée-fréquence (IdF), utilisée couramment pour la description des précipitations.

Dans la pratique, les variables descriptives sont extraites d'une chronique en appliquant les échantillonnages MAXAN ou SUP-SEUIL non pas sur les chroniques à pas de temps fin, mais sur les séries modifiées suivantes (Lang, 1995) :

- ✓ Débits moyens sur la durée d (Vd) : la série journalière est remplacée par la série des moyennes mobiles calculées sur la durée d . Cette variable permet de décrire les volumes générés par les événements extrêmes.
- ✓ Débits seuils sur la durée d (Qd) : la série journalière est remplacée par la série des minimums mobiles calculées sur la durée d . Cette variable s'interprète comme étant la valeur continûment dépassée sur la durée d . Son interprétation physique est moins évidente, mais elle présente l'intérêt d'être adaptée à des problèmes concrets liés à des débordements.

Description d'hydrogrammes

Nous avons également souhaité décrire la forme des hydrogrammes de crue. D'un point de vue pratique, ceci pose la question de l'extraction d'hydrogrammes à partir de la série journalière. Garçon *et al.* (2002) et Le Clerc (2004) fournissent de nombreux conseils pratiques dans cette optique. Nous avons décidé de définir un hydrogramme de crue de la manière suivante : tout d'abord, des pointes de débits sont sélectionnées par la méthode SUP-SEUIL, à partir de la chronique journalière. De part et d'autre de la pointe, l'ensemble des débits continûment supérieurs à la moitié de la pointe constituent l'hydrogramme à décrire, que nous nommerons également de façon quelque peu abusive l'« événement de crue ». A partir de cet événement, des indices de durée et de forme seront calculés. La Figure 2-16 illustre la définition de ces variables.

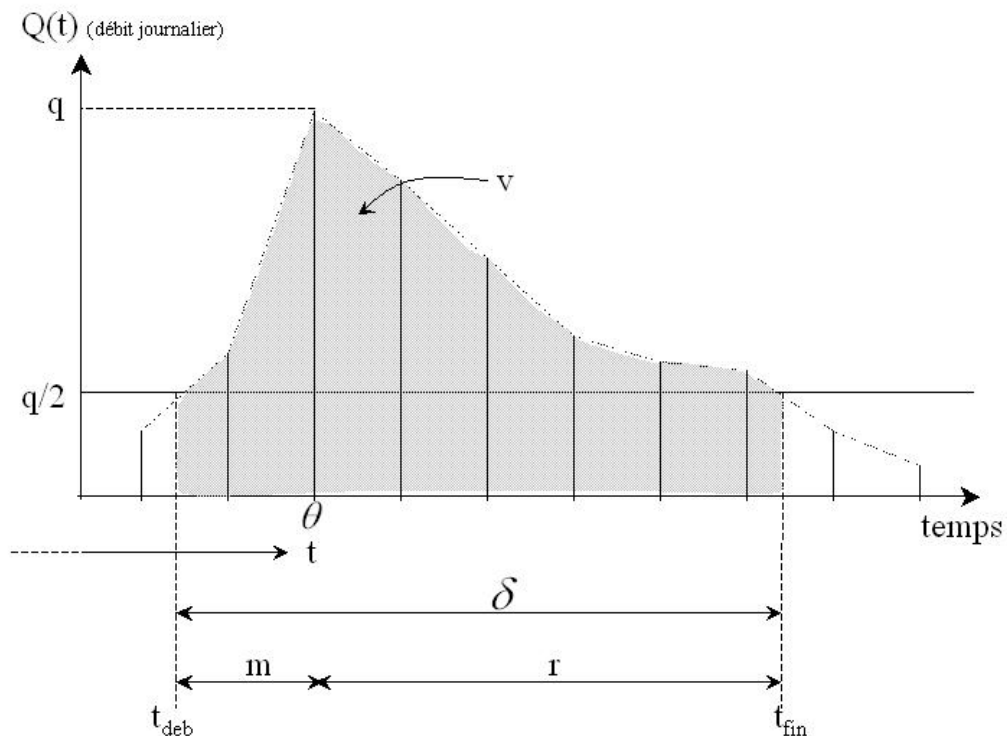


Figure 2-16. Schématisation des variables descriptives des hydrogrammes de crue.

Le caractère informatif de ces indices dépend en fait d'une hypothèse sous-jacente d'invariance de forme, qui consiste à supposer que tous les hydrogrammes de crue ont sensiblement la même forme. En d'autres termes, les hydrogrammes normalisés (*i.e.* divisés par le débit de pointe) doivent être relativement homogènes. En conséquence, les indices calculés entre le pic et la moitié du pic sont alors indépendants de l'intensité de la crue, et devraient pouvoir être considérés comme issus d'une même population. Evidemment, cette invariance de forme est loin d'être une hypothèse universelle. Elle ne s'applique pas, par exemple, à certains bassins à régime mixte, ou aux crues multiples. Elle peut également être totalement fautive sur des cours d'eau à régime complexe, pour lesquels il existe de nombreux mécanismes de genèse des crues.

Notons enfin que des méthodes beaucoup plus sophistiquées existent pour l'extraction d'hydrogrammes de crue, notamment pour décider quand commence la crue et quand finit la décrue. L'ajustement de courbes de récession pour décrire la décrue peut notamment être envisagée. Malheureusement, ces techniques se sont avérées assez peu robustes, et donc peu adaptées à une automatisation complète des procédures d'extraction.

Conclusion : définition des variables de crue.

L'ensemble des variables descriptives des hautes eaux est finalement résumé dans le Tableau 2-3.

	Variable	Description
MAXAN ⁽¹⁾	$q_1=(VCXd)_A$	Maximum annuel du débit moyen calculé sur la durée d
	θ_1	Date angulaire du débit q_1 ⁽²⁾
	$q_2=(QCXd)_A$	Maximum annuel du débit continuellement dépassé sur la durée d
	θ_2	Date angulaire du débit q_2
SUP-SEUIL	$q_3=(VCXd)_S$	Débit moyen (sur la durée d) maximal supérieur au seuil S
	θ_3	Date angulaire de q_3
	t_3	Temps cumulé depuis le début de la chronique ⁽³⁾
	$N_3(i)$	Nombre de débits q_3 sélectionnés pour l'année i
	$q_4=(QCXd)_S$	Débit maximal continuellement dépassé sur la durée d , supérieur au seuil S
	θ_4	Date angulaire de q_4
	t_4	Temps cumulé depuis le début de la chronique
Sélection d'hydrogrammes ⁽⁴⁾	δ	Durée de l'événement ⁽⁵⁾
	m	Temps de montée ⁽⁶⁾
	r	Temps de redescente ⁽⁷⁾
	σ	Ecart type de l'événement ⁽⁸⁾
	γ	Asymétrie de l'événement ⁽⁹⁾
	v	Volume écoulé durant l'événement ⁽¹⁰⁾

Tableau 2-3. Variables de crues utilisées.

- (1) L'année hydrologique commence au premier jour du mois ayant le plus faible débit mensuel inter-annuel.
- (2) $\theta=2\pi j/365.25$, où j est le nombre de jours écoulés depuis le début de l'année hydrologique.
- (3) Il s'agit de temps reconstitué (on ne compte pas les périodes de lacunes).
- (4) Un hydrogramme est défini comme la série des débits journaliers sur $[t_{deb}, t_{fin}]$ telle

$$\text{que : } \begin{cases} q(t_{deb}) = q(t_p) / 2 \\ q(t) > q(t_p) / 2, \forall t \in]t_{deb}; t_{fin}[, \text{ où } t_p \text{ est la date du pic.} \\ q(t_{fin}) = q(t_p) / 2 \end{cases}$$

Le pas de temps étant journalier, nous effectuons une interpolation linéaire au début et à la fin de la période de dépassement pour calculer t_{deb} et t_{fin} .

$$(5) \delta = t_{fin} - t_{deb}, (6) m = t_p - t_{deb}, (7) r = t_{fin} - t_p, (8) \sigma = \left(\frac{\sum_i (t_i - t_p)^2 \frac{q(t_i)}{\sum_j q(t_j)}}{\sum_j q(t_j)} \right)^{1/2},$$

$$(9) \gamma = \left(\frac{\sum_i (t_i - t_p)^3 \frac{q(t_i)}{\sum_j q(t_j)}}{\sum_j q(t_j)} \right) / \sigma^3, (10) v = \sum_i (t_{i+1} - t_i) \left(\frac{q(t_i) + q(t_{i+1})}{2} \right).$$

Etant donné le grand nombre de variables à extraire, sur plusieurs durées et avec plusieurs seuils, il est totalement indispensable d'automatiser les procédures d'extraction. De plus, plusieurs choix restent à effectuer afin de calculer les variables, notamment celles relatives aux méthodes SUP-SEUIL ou de sélection d'hydrogrammes : quelles contraintes d'indépendance utiliser ? Comment calculer une durée caractéristique ? Comment traiter les régimes mixtes ? Les parties suivantes ont pour but de lever ces interrogations, et de décrire de manière détaillée les algorithmes utilisés pour l'échantillonnage SUP-SEUIL.

II.1.2. Algorithme d'extraction de variables

Procédure 1 : Extraction sans contraintes, seuil connu

Soit X_t la chronique de départ. L'algorithme est le suivant :

- ✓ Recherche du maximum de X_t
- ✓ De part et d'autre de ce maximum : on efface tous les points tant que le débit n'est pas repassé sous le seuil S
- ✓ Réexécuter les deux étapes précédentes, à partir de la série tronquée
- ✓ Arrêt de la procédure lorsque le maximum de la série tronquée est inférieur au seuil

Procédure 2 : Application des contraintes

A partir de l'échantillon fourni par la procédure 1, on élimine les valeurs qui ne respectent pas les contraintes d'indépendance, qui seront discutées plus tard.

Procédure 3 : Calcul d'un seuil à partir d'un objectif μ_0 (nombre de valeurs par an)

Le nombre μ est égal à l'effectif de l'échantillon obtenu par la procédure 2, divisé par le temps d'observation hors lacunes, exprimé en années. L'algorithme de recherche du seuil S associé au nombre μ est le suivant (Javelle, 2001):

- ✓ $S_1 = \min(X_t)$; $S_2 = \max(X_t)$;
- ✓ Calcul de μ pour 10 valeurs de S équiréparties entre S_1 et S_2 .
- ✓ Choix de deux nouvelles valeurs S_1 et S_2 conduisant à μ_1 et μ_2 tels que $\mu_1 > \mu_0 > \mu_2$.
- ✓ Les deux étapes précédentes sont itérées 3 fois.

On choisira finalement comme valeur de seuil $S = 0.5 * (S_1 + S_2)$. Dans la pratique, nous choisirons $\mu_0 = 0.5, 1$ et 2 . Notons qu'il n'est pas forcément possible de trouver un seuil correspondant à un objectif μ_0 trop élevé. Par exemple, les régimes nivaux ou glaciaires ne connaissent souvent qu'une unique crue chaque année. Dans ce cas, μ_0 sera égale à la plus grande valeur possible.

Traitement des lacunes

Un des avantages de la procédure d'échantillonnage SUP-SEUIL est d'être moins sujette aux problèmes de données manquantes. Il peut cependant arriver qu'une lacune soit présente pendant un événement de crue. Un « code lacune » permet de détecter de tels cas. Les événements problématiques seront alors retirés de l'échantillon.

La procédure d'extraction de pointes supérieures à un seuil correspondant à μ_0 valeurs par an en moyenne est finalement schématisée dans la Figure 2-17.

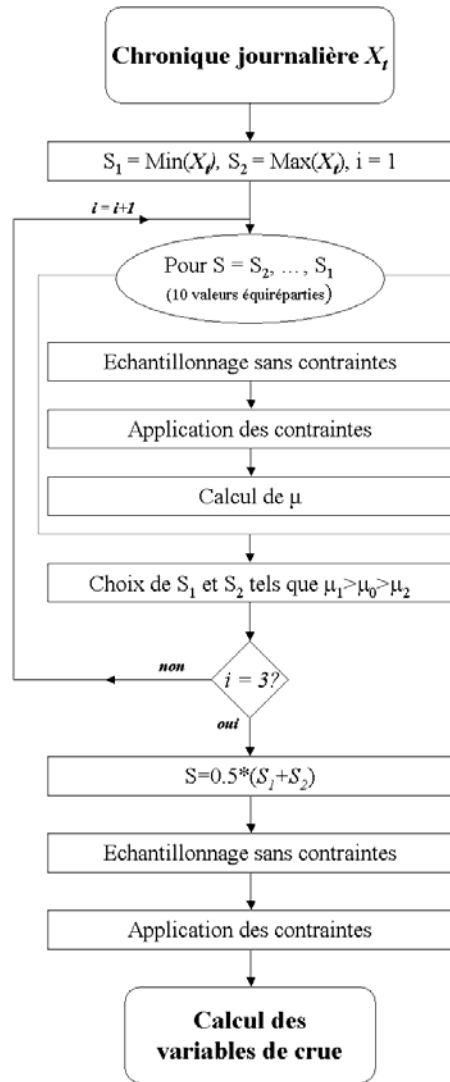


Figure 2-17. Fonctionnement de la procédure d'extraction de variables.

II.1.3. Critères d'indépendance

Contrainte C1 : redescente

Cette contrainte peut être décrite de la manière suivante : pour chaque événement, caractérisé par une pointe de débit $q(t_i)$, le débit minimum entre t_i et t_{i+1} , la date de l'événement suivant, doit être inférieur à $C_1 q(t_i)$. En cas de non respect de cette contrainte entre deux pics, seul le plus grand des deux est conservé. Il faut alors vérifier que la contrainte est également respectée entre t_i et t_{i+2} , puis éventuellement t_{i+3} , etc. L'algorithme d'application de cette contrainte est résumé dans la Figure 2-18. Etant donné la définition des hydrogrammes de crue, qui sont constitués des débits supérieurs à la moitié du pic, il semble cohérent d'imposer une contrainte de redescente $C_1=0.5$.

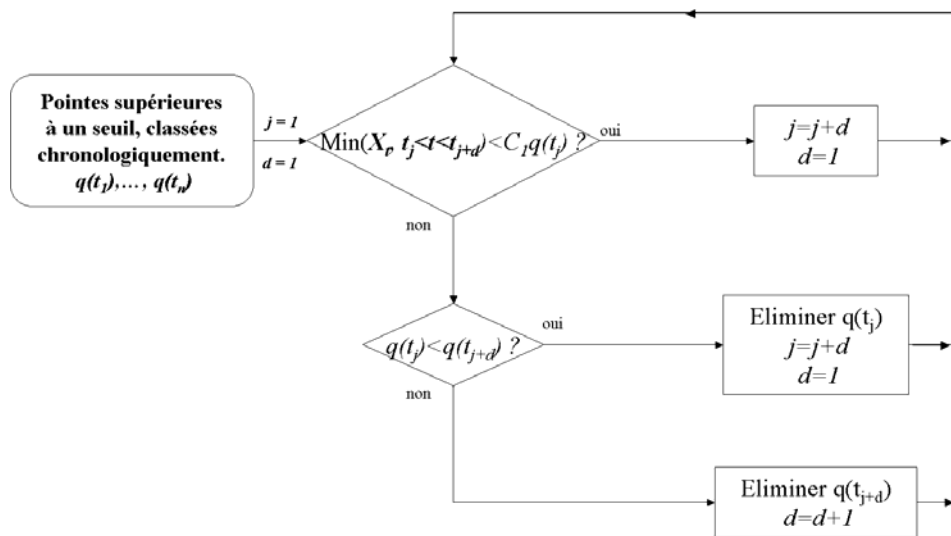


Figure 2-18. Algorithme d'application de la contrainte C1.

Contrainte C2 : espacement

Il s'agit ici de vérifier que toutes les pointes sélectionnées sont espacées d'une durée supérieure ou égale à C_2 . Cette durée sera caractéristique du régime des crues. Le premier problème vient du choix de cette durée caractéristique (D), qui sera discuté plus loin. Quelle que soit la définition adoptée, un second problème demeure : un échantillon est nécessaire pour estimer D ... mais quelle contrainte d'espacement adopter pour cet échantillon ? Trois solutions sont possibles :

- ✓ Estimer D par des formules empiriques utilisant des variables descriptives du bassin versant (méthode Socose).
- ✓ Estimer D à partir de l'échantillon des maximums annuels, puis effectuer l'échantillonnage sup-seuil.
- ✓ Échantillonner sans contrainte d'espacement, estimer D , puis appliquer une contrainte d'espacement multiple de D .

Nous opterons pour la dernière solution, qui permet de disposer d'un échantillon plus conséquent en choisissant $\mu=2$.

Contrainte C3 : homogénéité sélection de pics / sélection d'hydrogrammes

Une des difficultés provient du fait que les variables que nous souhaitons extraire ne sont pas relatives au même objet : certaines sont relatives à une pointe de crue, d'autres à un épisode de débits supérieurs à la moitié du pic. Or, les contraintes C_1 et C_2 ne concernent que les pointes de crues. Il arrive ainsi que soient sélectionnées deux pointes indépendantes au sens des deux contraintes exposées précédemment, mais qui conduisent à la sélection d'épisodes recouvrants, où le pic n'est pas forcément la valeur la plus forte de l'épisode ! La Figure 2-19 présente un tel cas de figure.

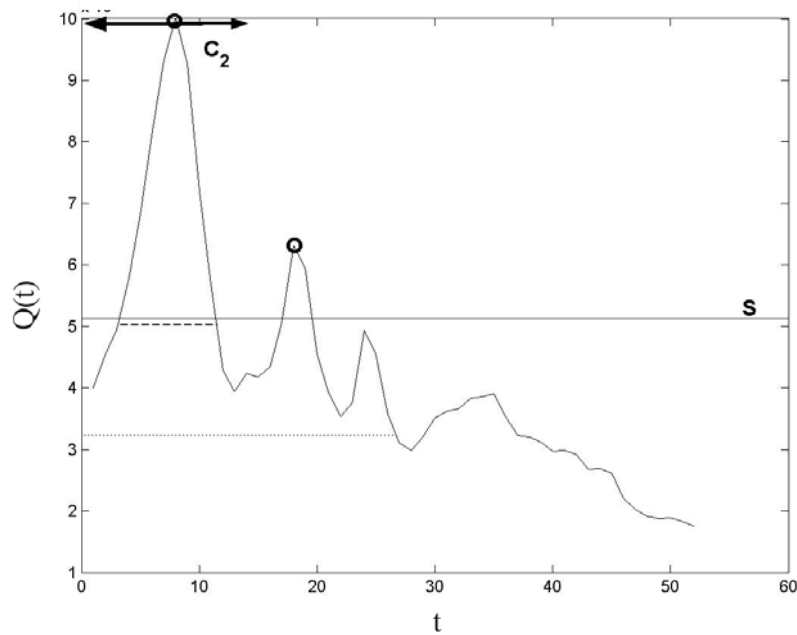


Figure 2-19. Illustration de l'insuffisance des contraintes C_1 et C_2 .

Les deux pics entourés sont ici sélectionnés : espacement supérieur à C_2 et redescente sous la moitié du premier pic (ligne pointillée épaisse). Cependant, l'épisode relatif au second pic comprend tous les débits supérieurs à la moitié de ce pic, figurée en pointillés fins : il comprend donc en particulier le premier pic ! Pour éviter ce désagrément, la contrainte C_3 consiste à vérifier que le pic est bien le maximum sur la durée de l'épisode. Dans le cas où la contrainte C_1 est cohérente avec la définition d'un événement (i.e. $C_1=0.5$), ceci est en fait équivalent à imposer une contrainte de redescente de la forme $C_1 \times \min(q(t_1), q(t_2))$, où t_1 et t_2 sont les instants des deux pics successifs. Pour simplifier, nous adoptons cette nouvelle définition comme contrainte de redescente.

II.1.4. Calcul d'une durée caractéristique de crue

La durée caractéristique de crue, que nous noterons D , sera utilisée pour définir une contrainte d'espacement C_2 , et pour savoir sur quelle durée calculer les débits moyens et seuils. Dans un premier temps, nous supposons que l'échantillon de crues utilisé pour l'estimation de la durée D est homogène. Nous discuterons par la suite du cas des mélanges de populations (régime pluvio-nival, par exemple).

La durée Δ issue du modèle QdF convergent

Cette durée est un paramètre du modèle QdF convergent. Ce modèle, développé par Javelle (2001), permet de relier le comportement fréquentiel des débits moyens sur différentes durées de la manière suivante :

$$V(d, T) = \frac{V(0, T) - P}{1 + d/\Delta} + P, \quad (2-1)$$

où $V(d, T)$ est le débit moyen sur la durée d de période de retour T ($d=0$ correspondant aux débits instantanés), et P et Δ sont deux paramètres à estimer. Le paramètre P , qui a la dimension d'un débit, correspond à l'abscisse du point d'intersection des distributions de différentes durées, tandis que le paramètre Δ , qui a la dimension d'un temps, règle la courbure de la forme hyperbolique des courbes QdF en fonction de la durée d .

Grâce à une méthode itérative proposée par Meunier (2001), Δ peut être ajusté sans faire d'hypothèses sur la distribution de l'échantillon. Cependant, la validité de ce paramètre reste dépendante des hypothèses d'affinité des hydrogrammes de crue, de forme hyperbolique des courbes QdF en fonction de d , et de convergence des courbes en un point, dont l'abscisse est elle aussi à déterminer plus ou moins empiriquement (valeur nulle ou égale au module pour P ?). De plus, certains problèmes numériques demeurent, comme le montre la Figure 2-20: pour certaines stations, l'algorithme d'estimation de la durée Δ ne converge pas.

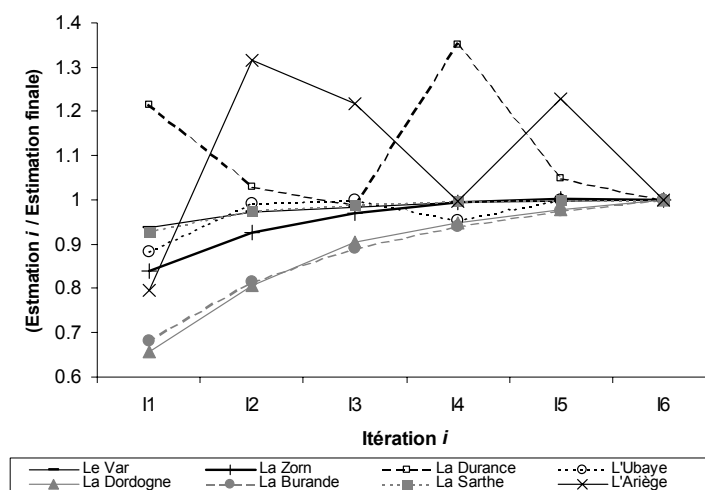


Figure 2-20. Convergence de la méthode de calcul de la durée Δ pour quelques stations.

Une autre difficulté est liée au fait que nous ne disposerons que de chroniques au pas de temps journalier. La précision d'estimation du paramètre Δ risque d'être moindre sans l'information instantanée. Ce biais est illustré dans la Figure 2-21, qui découle de données obtenues par Chetail (2004): il s'agit de la comparaison des valeurs de Δ estimées à partir de l'information instantanée et à partir de l'information journalière. On voit clairement la surestimation systématique par cette dernière méthode, qui ne semble pas se stabiliser avant une durée de l'ordre de sept jours. Quelle confiance alors accorder à une durée Δ estimée égale à cinq jours? Ces problèmes sont renforcés par l'impossibilité d'accéder à des indicateurs de variabilité pour le paramètre Δ , à moins d'avoir recours à des techniques de rééchantillonnage type Bootstrap. L'utilisation de ce paramètre ne semble donc pas approprié au cadre de cette étude.

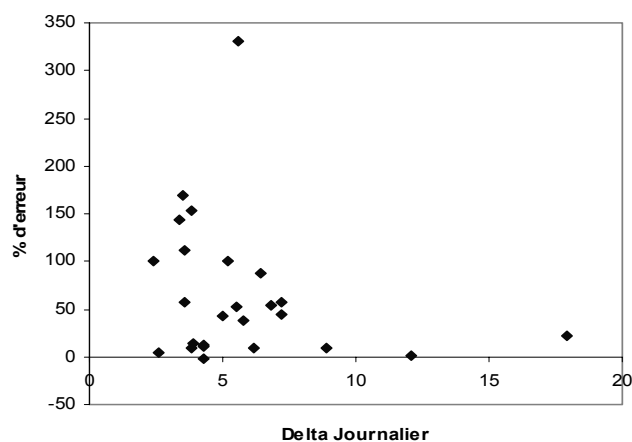


Figure 2-21. Erreur d'estimation de Δ à partir du pas de temps journalier.

Les durées de type Socose

Il s'agit ici d'estimer une durée caractéristique à partir de la durée de dépassement de la moitié du pic. Cette échantillon pourra provenir soit des maximums annuels, soit d'un premier échantillonnage SUP-SEUIL sans contrainte d'espacement. La durée de Socose à proprement parler (CTGREF *et al.*, 1980-1982) présente le désavantage de nécessiter une étape de modélisation probabiliste pour le calcul du débit décennal, et reste en partie empirique (découpage de l'échantillon en plusieurs classes). La méthode la plus simple consisterait à calculer la médiane de l'échantillon total, moins sensible aux valeurs extrêmes que la moyenne. Cependant, cette valeur reste influencée par un certain nombre de crues multiples, dont la durée ne peut plus être considérée comme représentative d'une réponse du bassin versant à la pluie. Pour éviter ce biais il est possible de calculer pour chaque événement, la complexité de la crue, définie comme le nombre d'extremums locaux de l'hydrogramme. En ne basant le calcul de la médiane que sur les crues simples (ie. un unique extremum local, le pic), on s'affranchit de l'influence des crues multiples. La Figure 2-22 présente les résultats de cette technique appliquée à la Sarthe : la densité, estimée par un noyau Gaussien, est moins attirée par les fortes valeurs de durées pour l'échantillon restreint aux crues simples. Ceci se traduit par une durée médiane qui passe de 7.9 jours (échantillon total) à 5.5 jours (série restreinte).

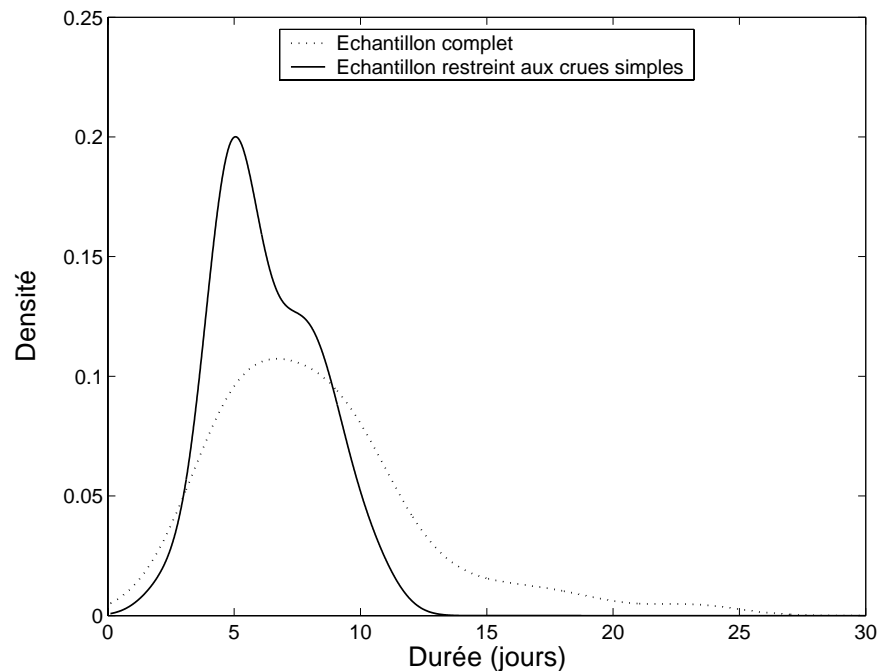


Figure 2-22. Densité des durées de dépassement de la moitié du pic, estimée par un noyau Gaussien, pour la Sarthe à Spay.

Après quelques essais, cette méthode semble s'appliquer efficacement sur les crues d'origine pluviale. Le pas de temps journalier permet à notre critère de ne pas être trop contraignant : avec un pas de temps plus fin, la plupart des hydrogrammes présenterait de légers extremums locaux secondaires, il faudrait donc dans ce cas réfléchir à un critère de complexité plus souple. Par contre, pour les régimes nivaux, les hydrogrammes ont généralement une complexité supérieure à 1. Il conviendrait donc de n'appliquer cette technique que si le nombre d'hydrogrammes rejetés ne dépasse pas un certain pourcentage, que nous fixerons à 50%. Dans le cas contraire, la durée caractéristique sera définie comme la médiane de l'échantillon complet.

II.1.5. Cas des mélanges de population (régime pluvio-nival)

Sur certains bassins versants, les crues peuvent résulter de plusieurs phénomènes physiques bien différenciés, comme par exemple la fusion nivale et les précipitations intenses. Dans ce cas, l'homogénéité de l'échantillon ne pourra pas être garantie. Il est donc impératif de prendre en compte cette multimodalité pour le calcul d'une durée caractéristique, qui peut varier de l'ordre du jour pour les phénomènes pluviaux à l'ordre du mois pour les phénomènes nivaux. La Figure 2-23 illustre ce cas de figure sur les données de l'Ubaye à Barcelonnette.

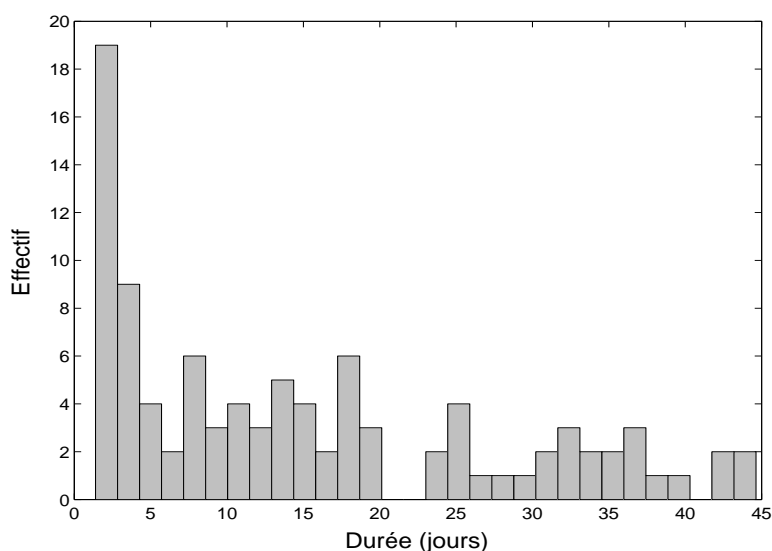


Figure 2-23. Histogramme des durées de dépassement de la moitié du pic pour l'Ubaye à Barcelonnette

Détecter un mélange de populations

Le problème se rattache au domaine de la classification, bien connu en statistiques. Il s'agit de segmenter en plusieurs groupes (nous supposons qu'il y a au plus deux groupes), puis de décider si cette représentation en sous-échantillons est plus pertinente (au sens d'un critère à définir) qu'une représentation avec une unique population. Cette segmentation peut se faire sur plusieurs variables, descriptives de la forme des hydrogrammes : durées (dépassement, montée et redescente), forme (écart-type et asymétrie). Cependant, sur les bassins français auxquels nous nous intéressons, il semble que la variable la plus pertinente soit la date d'occurrence des crues, puisque les phénomènes physiques à l'origine des crues sont soumis à une certaine saisonnalité.

La démarche adoptée est la suivante : dans un premier temps, il s'agit de définir une distance entre deux événements, sur la base de leur date angulaire d'occurrence. Cette distance ne peut pas être euclidienne, à cause de la topologie circulaire de cette variable : les points 0 et 2π sont ainsi situés à une distance nulle l'un de l'autre (et non égale à 2π). La distance entre deux dates d'occurrence sera donc définie comme suit (Figure 2-24):

$$d(\theta_1, \theta_2) = \min\left(|\theta_1 - \theta_2|, 2\pi - |\theta_1 - \theta_2|\right), \quad (2-2)$$

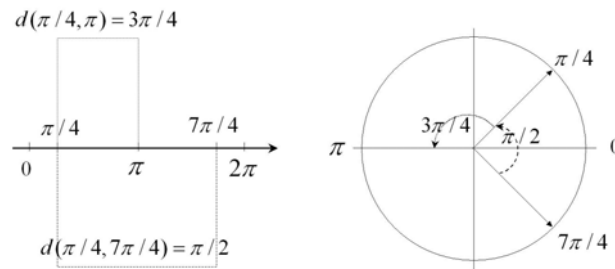


Figure 2-24. Illustration de la distance angulaire définie entre deux dates d'occurrence.

Sur la base de la matrice de distance ainsi définie, une classification hiérarchique ascendante (critère du saut moyen) permet d'effectuer la segmentation.

Le résultat de cette technique est présenté dans la Figure 2-25 pour l'Ubaye à Barcelonnette : les crues d'origine nivale (croix) et pluviale (cercles) sont ainsi parfaitement discriminées. La distribution des durées, en ordonnées, semble ainsi plus homogène à l'intérieur de chaque groupe.

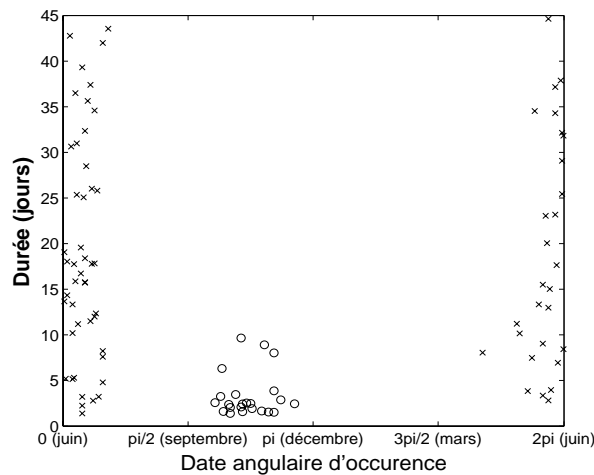


Figure 2-25. Classification des événements de l'Ubaye à Barcelonnette en deux groupes.

Il reste maintenant à définir un critère pour évaluer la pertinence de cette classification en deux groupes. Rappelons que la priorité est de définir une durée caractéristique de crue, sur des populations suffisamment homogènes. Nous allons donc effectuer un test de comparaison des médianes (Wilcoxon) des durées de chaque groupe. Si la partition reflète réellement une différence physique dans la genèse des événements, alors les médianes devraient être significativement différentes. Le Tableau 2-4 donne les statistiques de test, ainsi que la p -valeur correspondante. Les régimes nivo-pluviaux sont représentés en grisés.

Rivière	code	Wilcoxon	p
La Burande	P0115010	0	1
La Zorn	A3472010	0.08	0.936
La Sarthe	M0500610	0.16	0.873
Le Var	Y6432010	2.17	0.030
La Dordogne	P0010010	2.51	0.012
La Durance	X0130010	3.86	<0.001
L'Ariège	O1252510	5.63	<0.001
L'Ubaye	X0434010	6.25	<0.001

Tableau 2-4. Résultat du test de Wilcoxon sur 8 stations.

On voit ainsi que le test est hautement significatif pour les régimes mixtes, la partition semble donc adaptée. Cependant, le test est également significatif au risque 5% pour les bassins du Var et de la Dordogne, bien que la différence de distribution des durées ne semble pas suffisante pour parler de deux régimes distincts, comme le montre la Figure 2-26. Il semble donc nécessaire de s'assurer que la différence soit très hautement significative entre les deux groupes, afin de ne pas segmenter inutilement des séries. Nous choisirons donc un risque de 10^{-3} , ce qui correspond à un seuil de rejet de 3.29 sur la statistique de test de Wilcoxon.

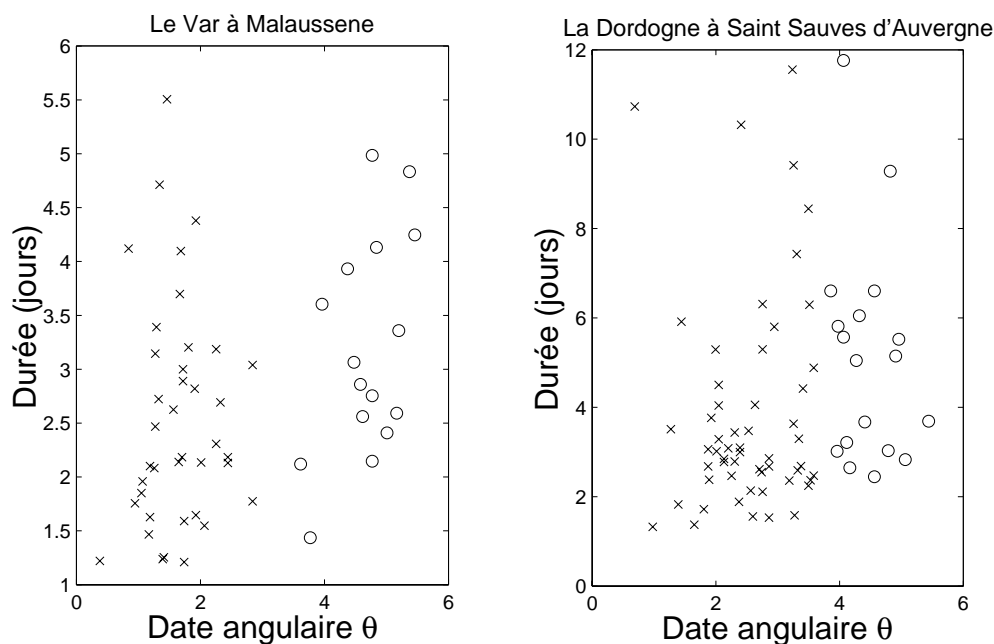


Figure 2-26. Date d'occurrence et durée des événements pour des bassins du Var et de la Dordogne.

Comment traiter les mélanges de populations ?

La procédure précédemment décrite permet donc de définir, si nécessaire, deux sous-populations homogènes. Dans ce cas, la procédure d'échantillonnage doit être saisonnalisée afin de fournir deux échantillons distincts. De même, le calcul des débits moyens devra être effectué sur les deux durées caractéristiques des deux populations. La limite des saisons peut être calculée comme la moyenne entre l'événement le plus tardif du groupe 1 et l'événement le plus précoce du groupe 2 (avec les précautions nécessaires résultant de la topologie circulaire...). Des méthodes plus sophistiquées pourraient être envisagées pour calculer ces limites (cf. les problèmes de classement en statistiques, avec minimisation des probabilités d'erreur). Cependant, leur application pourrait s'avérer délicate, pour un gain qui ne semble pas évident, puisqu'il faudrait définir une loi de probabilité sur un espace circulaire. Enfin, au niveau régional, la composante nivale des régimes mixtes pourra être traitée avec celle des régimes purement nivaux.

II.1.6. Procédure complète d'extraction

Les quelques résultats présentés ici permettent de définir une procédure d'extraction, schématisée dans la Figure 2-27.

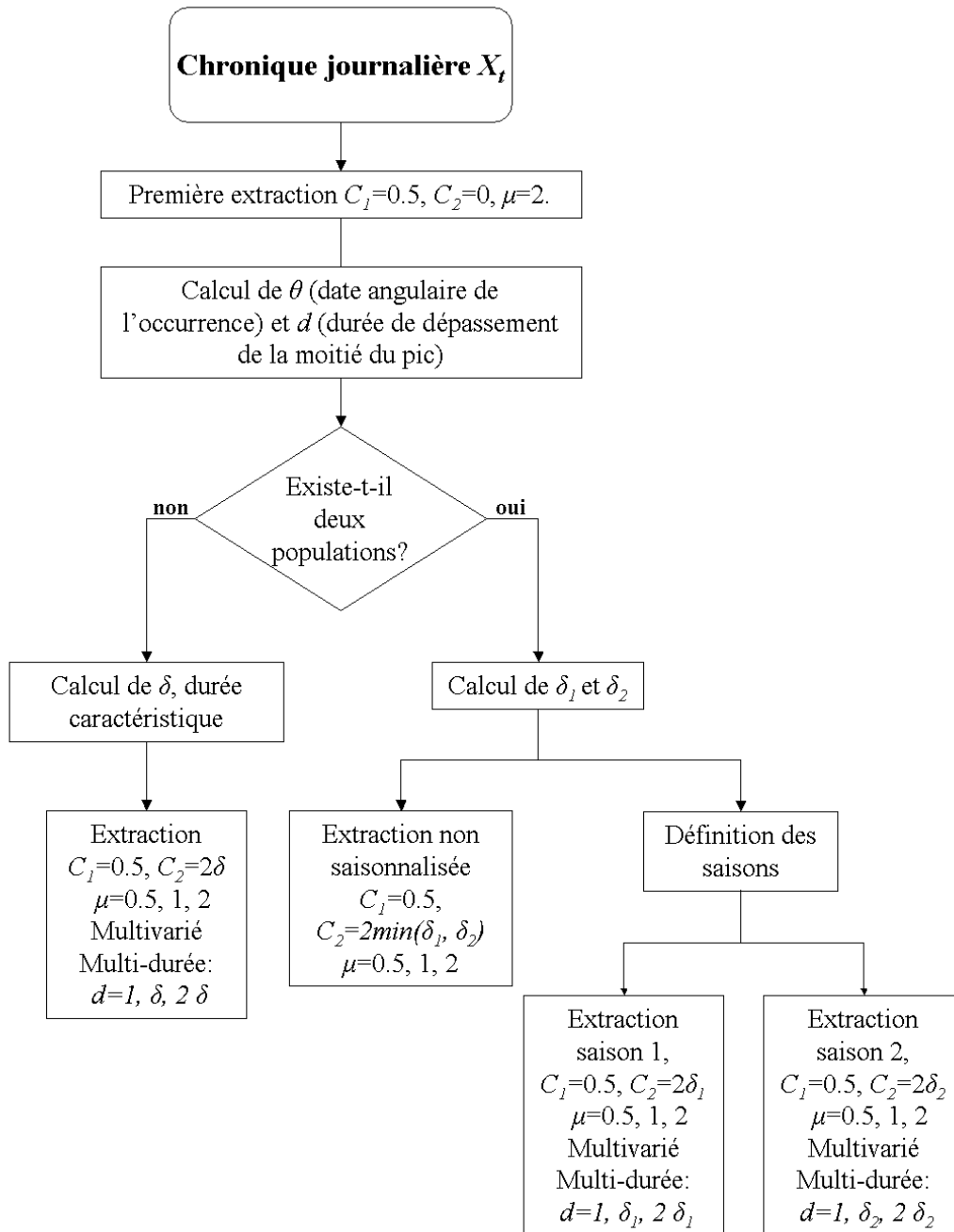


Figure 2-27. Procédure d'extraction de variables.

II.1.7. Traitement particulier des crues nivales

Les crues nivales, liées à la fonte des neiges stockées durant l'hiver, possèdent la particularité d'être très longues et volumiques. La Figure 2-28 donne un exemple de rivière soumise à ce type de crue. On observe que des brusques variations de débits, liées à des épisodes pluvieux, se superposent à une onde de fonte beaucoup plus lente. Ainsi, la pointe de ce type de crue caractérisera plutôt l'importance de l'épisode de pluie, et non celle de l'onde nivale. La date du maximum est peut être encore plus sensible à ce problème. Or, les gestionnaires de ce type de rivière, dans les Alpes notamment, ressentent une précocité accrue de la fonte des neiges depuis quelques années. Il serait donc particulièrement intéressant de quantifier ce phénomène afin de pouvoir le confirmer ou l'infirmer.

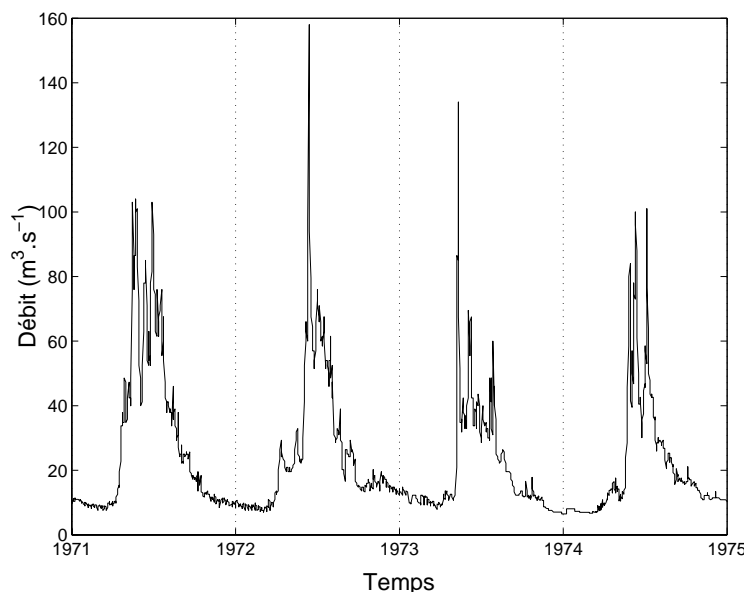


Figure 2-28. Extrait de la chronique de débits journaliers de la Durance à l'Argentière La Bessée.

Définition des variables

La première possibilité consiste à lisser la chronique des débits journaliers en calculant des moyennes (Vd) ou des minimums (Qd) mobiles. Après quelques essais, il s'avère que les moyennes mobiles restent assez sensibles à de forts épisodes pluvieux, contrairement aux minimums mobiles, qui permettent de gommer plus efficacement ces brusques variations de débit. Le calcul des minimums mobiles s'effectuera sur une durée de 10 jours, qui permet « d'effacer » la plupart des pluies, tout en décrivant l'onde nivale de manière suffisamment précise. Nous nous intéresserons donc à la valeur q_1 et la date θ_1 de la pointe de cette série des débits seuils.

Une seconde possibilité consiste à considérer que la fonte des neiges contribue principalement au débit de base, alors que les pluies vont contribuer au ruissellement. En conséquence, une procédure de séparation du débit de base devrait permettre de limiter l'analyse à la partie du débit qui est due à la neige. Nous avons appliqué l'algorithme proposé par Tallaksen et Van Lanen (2004). A titre d'exemple, la Figure 2-29 permet de comparer le débit journalier et le débit de base estimé par cette procédure. Le résultat semble satisfaisant, puisque les événements pluvieux, même importants, n'influencent pas la courbe du débit de base. Comme précédemment, la valeur q_2 et la date θ_2 de la pointe de débit seront extraites de cette série transformée.

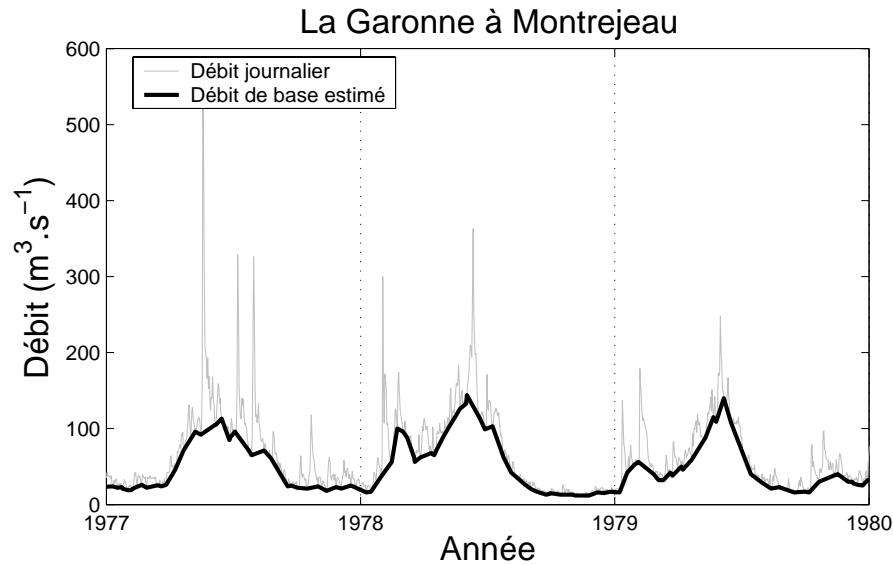


Figure 2-29. Comparaison entre le débit journalier et le débit de base estimé par la procédure de Tallaksen et Van Lanen (2004).

Un dernier indice caractérisant la saisonnalité des crues nivales a été proposé par Stewart *et al.* (2005) : pour une année donnée, le centre de masse (χ) est défini comme la date à laquelle la moitié du module a été écoulé. Plutôt que de considérer l'année hydrologique complète, nous nous limiterons à la saison des crues nivales. Cette caractérisation permet de minimiser l'influence des épisodes pluvieux, puisqu'à l'échelle de la saison, le volume qu'ils représentent est négligeable par rapport au volume total de l'onde de fonte. Enfin, nous calculerons également le module de la saison nivale q_3 .

Les variables spécifiques aux crues nivales sont finalement résumées dans le Tableau 2-5.

Variable	Description
$q_1=(QCX_{10j})$	Maximum saisonnier ⁽¹⁾ du débit seuil calculé sur 10 jours
θ_1	Date angulaire du débit q_1
q_2	Maximum saisonnier du débit de base
θ_2	Date angulaire du débit q_2
χ	Centre de masse évalué sur la saison
$q_3=QS$	Module saisonnier

Tableau 2-5. Variables spécifiques aux crues nivales.

(1) La saison des crues nivales s'étend de mars à septembre pour les régimes nival et pluvio-nival, de juin à octobre pour le régime glaciaire.

II.2. Variables descriptives des basses eaux

Dans cette section, nous nous intéressons aux faibles valeurs de débits. On parle en général de basses eaux ou d'étiages, ce dernier terme intégrant le caractère exceptionnellement bas des débits. Nous n'étudierons pas par contre le phénomène de sécheresse, dont le terme contient une connotation temporelle et hydrique et est plutôt expliqué par les déficits pluviométriques, ni les déficits, qui désignent des valeurs faibles par rapport à la normale, sans pour autant que ces valeurs soient faibles dans l'absolu.

II.2.1. Définition des variables

Dans un premier temps, il semble naturel d'utiliser la dualité existant entre les crues et les étiages pour la définition des variables : un débit faible n'est rien d'autre qu'une forte valeur de l'opposé du débit ! Malheureusement, le phénomène physique est plus compliqué que cette simple observation arithmétique. Si la méthode d'échantillonnage par valeurs maximales annuelles se transpose sans difficulté aux bas débits (méthode MINAN), la sélection de valeurs inférieures à un seuil est plus problématique : il est en effet particulièrement difficile de déterminer des événements indépendants au sein d'une même année hydrologique. La simple application de contraintes d'espacement temporel ou de remontée ne suffit pas toujours à garantir l'indépendance des valeurs sélectionnées, à cause de l'inertie du mécanisme remplissage/vidange de la nappe d'accompagnement. La position la plus sage consiste donc à ne considérer qu'un unique étiage (éventuellement fragmenté) par année hydrologique (Figure 2-30) : à partir d'un seuil de bas débit, il est possible de calculer le nombre de jours δ passés sous le seuil, le déficit de volume v par rapport au seuil (cumulé sur l'année), ainsi que les dates de début et de fin d'étiage θ_1 et θ_2 .

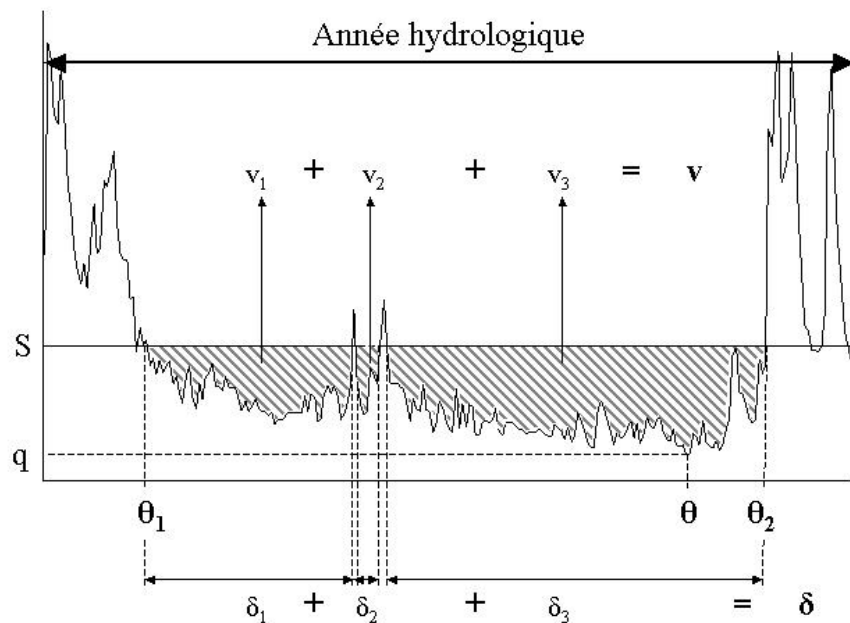


Figure 2-30. Schématisation des variables d'étiage.

II.2.2. Choix d'un seuil de bas débit

Le seuil de bas débit sera égal à un quantile de la courbe des débits classés. La valeur de ce seuil déterminera le caractère exceptionnel des événements sélectionnés. Cependant, un seuil trop bas peut être problématique, car de nombreuses années peuvent alors ne connaître aucun étiage, conduisant à des événements de durée nulle. Du point de vue statistique, l'application d'un test de stationnarité à une série comportant de nombreux zéros peut s'avérer problématique. Inversement, un seuil trop haut ne correspondra plus forcément à une caractérisation des extrêmes, conduisant à des durées d'étiages trop importantes. La Figure 2-31 décrit l'évolution comparée sur huit stations du nombre de données nulles et supérieures à 90 jours, en fonction du seuil choisi (exprimé comme un quantile de la courbe des débits classés). Un seuil égal au quantile à 15% de la courbe des débits classés semble être un bon compromis.

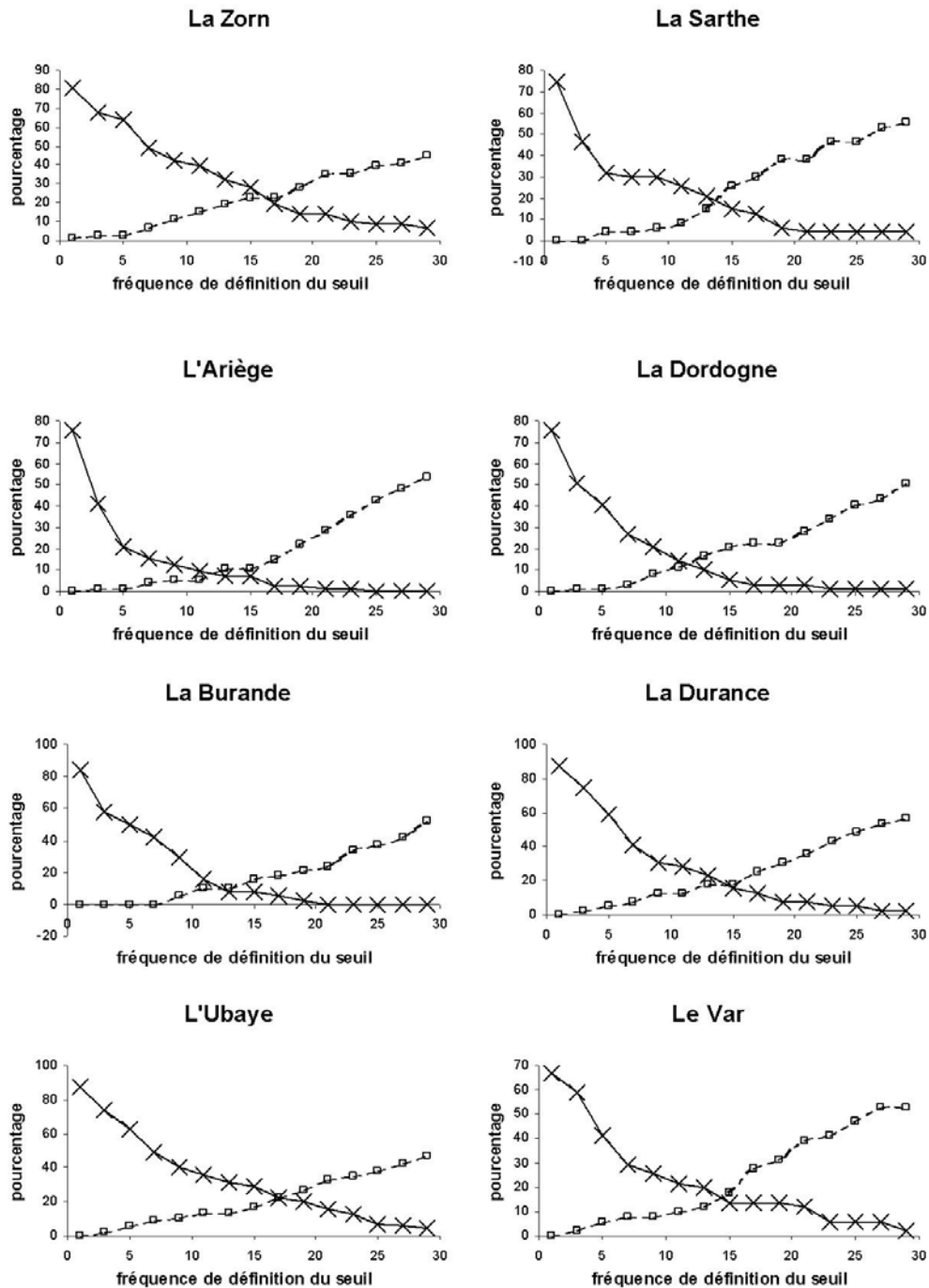


Figure 2-31. Comparaison du pourcentage d'étiages de durée nulle (croix) et supérieure à 90 jours (carrés) en fonction du seuil choisi.

II.2.3. Analyse multi-durées

Comme dans le cas des crues, le pas de temps journalier n'est pas forcément le plus adapté pour la description des étiages. Les variables relatives à l'échantillonnage par valeurs minimales annuelles seront donc également extraites à partir des séries de débits moyens et seuils calculés sur une durée de d jours. Le débit seuil est ici défini comme le débit continuellement non dépassé pendant d jours, et est obtenu en calculant des maximums mobiles sur d jours.

Il est par contre plus difficile de calculer une durée caractéristique d'étiage. En effet, la durée d'un étiage est principalement contrôlée par la durée de la période de déficit pluviométrique, plutôt que par les caractéristiques intrinsèques au bassin. Ainsi, la Figure 2-32 ne montre pas de différence fondamentale entre les distributions des durées d'étiages pour huit bassins versants aux comportements bien différenciés. Par contre, la vitesse à laquelle le bassin réagit à l'absence de pluie est une caractéristique intrinsèque, et dépend principalement de sa géologie et de l'importance de la nappe. Il est malheureusement difficile de quantifier ce temps de réaction sans les données pluviométriques correspondantes. L'ajustement de courbes de récession est envisageable, mais ne s'automatise pas simplement sur un grand nombre de bassins versants hétérogènes. Nous avons donc choisi d'utiliser des durées calendaires pour le calcul des débits moyens et seuils, à savoir 7 et 30 jours.

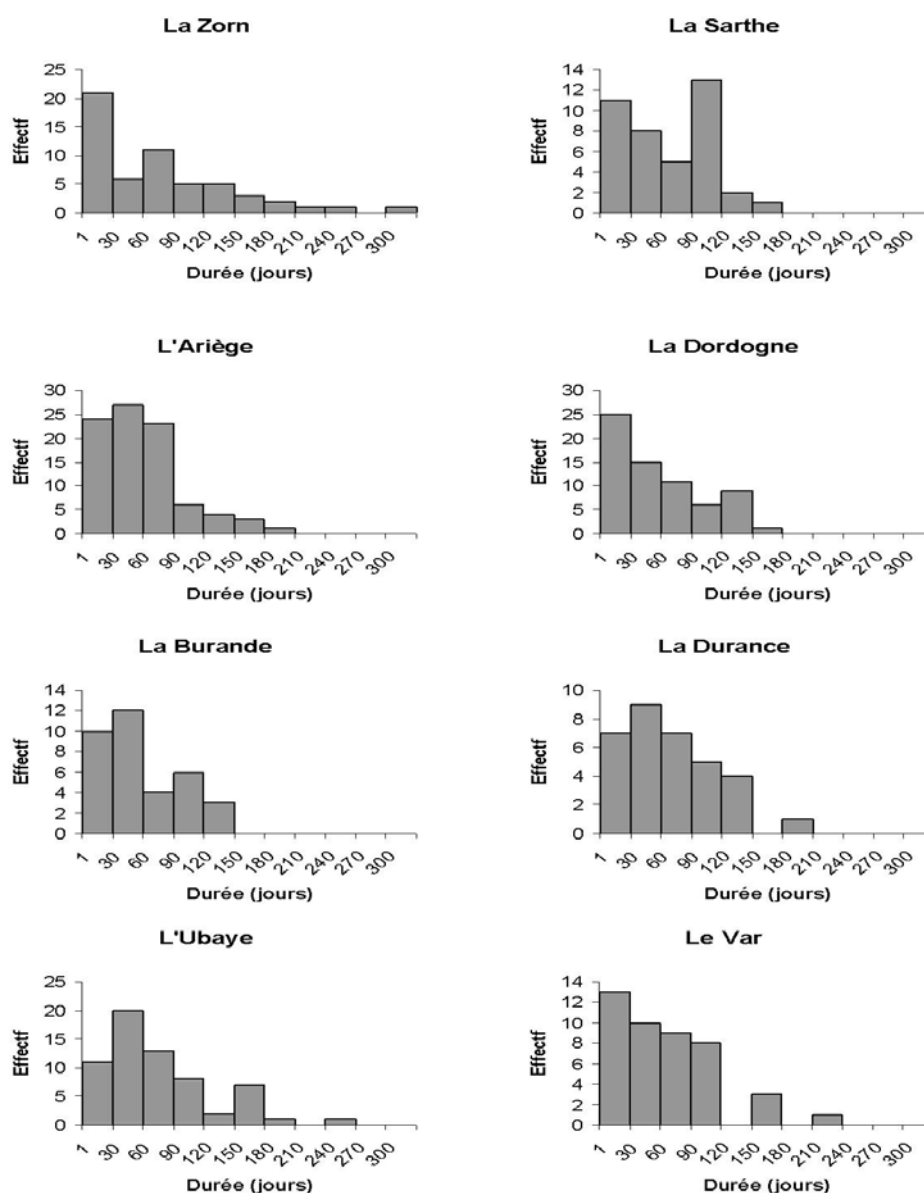


Figure 2-32. Histogrammes du nombre de jours annuel passé sous un seuil de bas débit pour 8 stations hydrométriques.

L'ensemble des variables utilisées pour la description des basses eaux est finalement décrit dans le Tableau 2-6.

Variable	Description
$q_1=(VCNd)_A$	Minimum annuel ⁽¹⁾ du débit moyen calculé sur d jours ⁽²⁾
θ_1	Date angulaire du débit q_1
$q_2=(QCNd)_A$	Minimum annuel du débit seuil calculé sur d jours
θ_2	Date angulaire du débit q_2
δ	Durée cumulée sur l'année de non-dépassement du seuil ⁽³⁾
θ_{deb}	Date angulaire du premier passage sous le seuil de l'année
θ_{fin}	Date angulaire du dernier jour sous le seuil de l'année
v	Déficit de volume par rapport au seuil, cumulé sur l'année

Tableau 2-6. Variables descriptives des basses eaux.

- (1) L'année hydrologique commence au premier jour du mois ayant le plus fort débit mensuel inter-annuel.
- (2) $d=1, 7$ et 30 jours
- (3) Le seuil est égal au quantile à 15% de la courbe des débits classés.

II.3. Variables descriptives du régime des eaux

Bien que le travail de thèse soit dédié aux basses et aux hautes eaux, nous avons également décidé d'extraire et d'analyser deux variables non relatives aux extrêmes. Tout d'abord, nous étudierons le module annuel $m(i)$, égal au débit moyen au cours de l'année i . Nous nous intéresserons également aux variations de régime en calculant, pour chaque année hydrologique, la distance au régime moyen de la manière suivante :

$$d(i) = \sqrt{\sum_{m=1}^{12} (\dot{p}_m - p_m(i))^2}, \quad (2-3)$$

où \dot{p}_m désigne le débit moyen inter-annuel du mois m (exprimé en pourcentage du module inter-annuel), et $p_m(i)$ le débit moyen du mois m pour l'année i (exprimé en pourcentage du module de l'année i).

SYNTHESE DU CHAPITRE 2

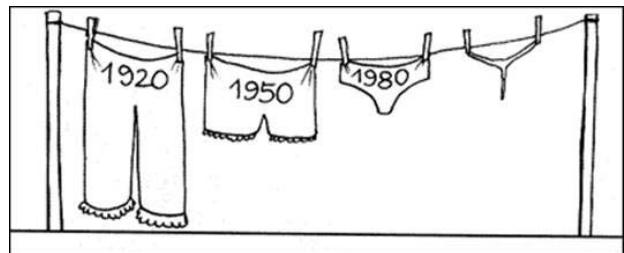
L'objet de ce chapitre était de présenter les données qui seront utilisées dans cette thèse, ainsi que les traitements préliminaires à l'analyse que nous avons effectués.

Dans un premier temps, des chroniques de débits journaliers ont été sélectionnées sur la base de critères de qualité, d'ancienneté et d'absence d'influences significatives. Les données provenant de 195 stations hydrométriques ont ainsi été récupérées auprès des DIREN, d'EDF et de la CNR. Les bassins versants contrôlés par ces stations présentent une grande gamme de taille et de fonctionnement, couvrant ainsi la totalité des régimes observés en France. Des classifications basées sur la saisonnalité des crues et des étiages ont ensuite été effectuées. Les classes obtenues ont été croisées avec un zonage pluviométrique préexistant afin de définir des régions hydro-climatiques de crue et d'étiage, qui seront utiles pour l'analyse à l'échelle régionale.

Dans un second temps, une réflexion a été menée quant à la définition de variables descriptives des événements hydrologiques extrêmes. L'extraction des variables réclame d'effectuer un certain nombre de choix, que nous avons étudiés afin de limiter au maximum les biais d'échantillonnage. Pour les crues, les méthodes permettant d'extraire des pointes de débits supérieurs à un seuil ou des hydrogrammes ont été évaluées. De plus, une procédure de détection de mélanges de populations (pluvial / nival) a été implémentée, afin d'effectuer le cas échéant un échantillonnage saisonnalisé. Pour les étiages, le manque d'indépendance des différentes périodes de bas débits au cours d'une même année a conduit à la description des étiages à l'échelle annuelle.

A l'issue de ce travail de pré-traitement des données, la prochaine étape consiste à définir le cadre méthodologique qui permettra d'étudier leur stationnarité.

CHAPITRE 3. COMMENT DETECTER DES CHANGEMENTS ?



“The indisputable evidence of the global climate warming”

La recherche de changements brusques ou graduels dans des séries statistiques est une préoccupation assez ancienne, qui a conduit à une grande abondance de méthodes disponibles dans la littérature. Cependant, chacune de ces méthodes possède son propre champ d’application et ses propres hypothèses. La question du choix de la « meilleure » technique à utiliser ne possède donc pas de réponse unique, mais dépend du contexte de l’étude. Les principaux facteurs influençant le choix d’une méthode concernent la structure probabiliste des variables (distribution, présence d’autocorrélation, données univariées ou multivariées), la taille des échantillons et le type de changement recherché (tendance, rupture unique, rupture multiple, changement en moyenne, en variance).

On peut classer les différentes méthodes disponibles en quatre grandes familles :

- ✓ Les approches descriptives : il s’agit de représenter visuellement une série de données afin de mettre en avant son comportement vis-à-vis de l’hypothèse que nous souhaitons évaluer (en l’occurrence, sa stationnarité). La première technique consiste à

créer le graphique des valeurs chronologiques contre le temps, mais d'autres méthodes plus sophistiquées peuvent également être utilisées (lissages ou approches descriptives multivariées). Nous utiliserons cette approche descriptive à de nombreuses reprises. Cependant, étant donné la taille des échantillons que nous aurons à traiter, il n'est pas envisageable de faire l'économie d'une étape supplémentaire d'inférence statistique, c'est à dire de l'évaluation de la significativité des résultats.

- ✓ L'approche Bayésienne : bien que très intéressante puisque adaptable à de nombreuses situations, la mise en place d'une telle analyse sur un grand nombre de stations n'est pas sans poser un certain nombre de problèmes, liés notamment à la spécification de lois *a priori* ou à la complexité informatique des méthodes MCMC. Nous n'utiliserons donc pas cette approche pour la détection de changements en France métropolitaine.
- ✓ Les procédures de segmentation et d'homogénéisation (Hubert *et al.*, 1989; Peterson *et al.*, 1998; Mestre, 2000; Kehagias, 2004): elles consistent à segmenter une série en plusieurs sous-séries stationnaires. Elles sont donc plutôt dédiées à la recherche de ruptures multiples, d'origine météorologique notamment, bien qu'elles aient aussi été appliquées à des études climatiques, dans le cas de phénomènes présentant une variabilité à l'échelle de la dizaine d'années (Hubert *et al.*, 1989). Bien que ce type de variabilité existe également en Europe de l'Ouest (oscillation Nord Atlantique), son influence sur les débits (*a fortiori* sur les débits extrêmes) n'est pas évidente à discerner, en France tout du moins. De plus, la significativité d'une segmentation n'est pas aisément accessible, sauf dans le cas d'hypothèses trop restrictives pour l'étude des extrêmes (normalité des variables). Nous n'utiliserons donc pas cette approche dans notre étude.
- ✓ Les tests statistiques : ce sont les outils que nous avons décidé d'adopter, car d'une part leur objectif est précisément d'évaluer la significativité d'une hypothèse, et d'autre part, ils sont en général assez simples à mettre en place, et sont donc particulièrement adaptés à une étude sur un grand nombre de stations. La littérature fournit un très grand nombre de tests de détection de changements, qui couvrent la quasi-totalité des situations que nous rencontrerons. La difficulté consiste à choisir le test adapté à une situation donnée. Un certain nombre d'études ont déjà été effectuées en ce sens (Lubès-Niel *et al.*, 1998; Kundzewicz et Robson, 2000; Lemaitre, 2002; Dindar, 2003; Kundzewicz, 2004; Kundzewicz et Robson, 2004; Zhang *et al.*, 2004), mais leurs conclusions ne sont pas toujours cohérentes. Par exemple, Kundzewicz et Robson (2004) conseillent d'utiliser des tests non paramétriques car ils sont presque aussi puissants que les tests paramétriques, tout en étant très souples d'utilisation. A l'inverse, Zhang *et al.* (2004) recommandent d'utiliser des tests paramétriques adaptés aux lois des valeurs extrêmes, en démontrant un net gain de puissance. Nous souhaitons donc reprendre et généraliser ces études afin de définir un schéma de sélection du « bon » test statistique.

Une fois la démarche d'analyse locale définie, un certain nombre de problèmes devront être résolus avant de pouvoir conclure à un effet du changement climatique. Le premier est lié à la répétition de tests dont nous maîtrisons le risque local : à partir de quel pourcentage de tests significatifs pouvons-nous être sûrs que tous ces changements ne sont pas dus au hasard ? Le second est lié à l'attribution des changements à une cause climatique, qui devrait s'exprimer de manière cohérente spatialement. Le but de ce chapitre est d'étudier l'ensemble de ces phénomènes, et de définir une démarche d'analyse garantissant un maximum de puissance au niveau local, et permettant d'évaluer la significativité et la cohérence régionale des résultats.

I. Echelle locale : quel test pour quelle variable ?

I.1. Présentation des principaux tests

I.1.1. Rappel : le principe des tests statistiques

Un test est un mécanisme qui permet de trancher entre deux hypothèses, dont une et une seule est vraie, au vu des résultats d'un échantillon (Saporta, 1990). Soient H_0 et H_1 ces deux hypothèses. Il existe alors quatre possibilités, de probabilités différentes, résumées dans le Tableau 3-1. α et β sont appelés les risques de première et de seconde espèce. Dans le cadre de la détection de tendances, α est la probabilité de rejeter à tort l'hypothèse de stationnarité, β est la probabilité de ne pas détecter une tendance qui existe. $1-\beta$ est la puissance du test.

Décision \ Vérité	H_0	H_1
H_0	$1-\alpha$	β
H_1	α	$1-\beta$

Tableau 3-1. Probabilités de bonnes et de mauvaises décisions.

L'application d'un test statistique suit généralement le canevas suivant :

1. Choix des hypothèses H_0 et H_1
2. Choix du risque de première espèce α
3. Choix de la statistique de test
4. Détermination de la loi de cette statistique sous H_0
5. Calcul de la zone de rejet
6. Calcul de la valeur expérimentale de la variable de test et décision

Dans la pratique, l'étape la plus difficile est généralement la détermination de la loi de la statistique de test sous l'hypothèse nulle. Cette distribution peut être exacte, asymptotique, ou tabulée sur la base de simulations. Pour y accéder, des hypothèses sur la distribution des observations sont généralement formulées : on parle alors de test paramétrique. Le test est dit non-paramétrique si la loi de la statistique ne dépend pas de la distribution des données.

Plusieurs tests sont généralement disponibles pour évaluer une même hypothèse. Le choix du test le plus approprié pourra se faire sur la base d'un certain nombre de critères de qualité. La première qualité d'un test est de respecter le risque de première espèce α : on parlera de test non biaisé. Si l'hypothèse nulle est trop souvent rejetée, on parlera de test libéral, et dans le cas contraire, de test conservatif. Dans un deuxième temps, on pourra rechercher le test le plus puissant. Enfin, la robustesse pourra être évaluée : il s'agit de la faculté du test à rester non biaisé lorsque les observations ne respectent plus les conditions d'application.

I.1.2. Quatre tests classiques

Nous allons présenter en détail quatre tests statistiques qui comptent parmi les plus utilisés dans les études de stationnarité. Il s'agit des tests de Mann-Kendall (Mann, 1945; Kendall, 1975), de Pettitt (1979), de Buishand (1982) et de régression linéaire. Ils illustreront la manière dont les statistiques de test ont été construites, dans l'optique d'une détection de tendance ou de rupture, dans les cas paramétrique et non paramétrique.

Le test de Mann-Kendall

Soit (x_1, \dots, x_n) un échantillon de valeurs indépendantes issu d'une variable aléatoire X , dont on cherche à évaluer la stationnarité. La statistique de Mann-Kendall est définie comme suit :

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{signe}(X_j - X_i) \tag{3-1}$$

En d'autres termes, la statistique est calculée en dénombrant, pour tous les couples $(x_i, x_j)_{i < j}$, le nombre de cas où la seconde valeur est supérieure à la première, et le nombre de cas où la seconde est inférieure, puis en faisant la différence entre ces deux quantités. Sous l'hypothèse H_0 de stationnarité de la série, la statistique devrait être proche de zéro. A l'inverse, si les valeurs de la série ont tendance à croître (resp. décroître), S devrait prendre des valeurs positives (resp. négatives). Ce test est donc plutôt dédié à la détection de tendances monotones, même s'il devrait également être capable de détecter des ruptures, mais sans donner d'estimation de la date de changement. Par contre, il n'est pas adapté à la détection de tendances non monotones. Sous H_0 , Mann (1945) et Kendall (1975) ont démontré que :

$$\begin{aligned} E(S) &= 0 \\ \text{Var}(S) &= n(n-1)(2n+5)/18 \end{aligned} \tag{3-2}$$

De plus, dès que l'échantillon compte une dizaine d'observations, alors la loi de la statistique de test Z ci-dessous peut être approchée par une Gaussienne centrée-réduite :

$$Z = \begin{cases} \frac{S-1}{\sqrt{\text{Var}(S)}} & \text{si } S > 0 \\ 0 & \text{si } S = 0 \\ \frac{S+1}{\sqrt{\text{Var}(S)}} & \text{si } S < 0 \end{cases} \tag{3-3}$$

Il s'agit donc d'un test non paramétrique, puisque aucune hypothèse n'a été émise quant à la distribution des données. Notons qu'une correction doit être apportée à la variance de S en présence d'ex aequo dans la série, de la manière suivante :

$$\text{Var}(S) = \left(n(n-1)(2n+5) - \sum_{k=1}^n t_k k(k-1)(2k+5) \right) / 18, \tag{3-4}$$

où t_k désigne le nombre d'égalités impliquant k valeurs. Pour autant, le test reste peu utilisable dans le cas d'une variable à valeurs entières créant un trop grand nombre d'ex aequo.

Le test de Pettitt

Ce test est très semblable dans sa construction au test de Mann-Kendall, mais il sera plutôt dédié à la détection d'une rupture. Soit k un entier compris entre 1 et $n-1$. On commence par calculer la statistique :

$$U(k) = \sum_{i=1}^k \sum_{j=k+1}^n \text{signe}(X_j - X_i) \tag{3-5}$$

On cherche ici encore à évaluer si les couples de valeurs ont tendance à être rangés préférentiellement dans un sens, mais cette fois-ci les couples sont formés par sélection d'une valeur de chaque coté d'un point de rupture k . La statistique de test vaut alors :

$$Z = \max_{k=1, \dots, n-1} |U(k)|, \quad (3-6)$$

et une estimation du point de rupture est naturellement :

$$\hat{k} = \text{Arg max}_{k=1, \dots, n-1} |U(k)| \quad (3-7)$$

La fonction de répartition de Z peut être approchée asymptotiquement par (Pettitt, 1979):

$$F(z) = P(Z \leq z) = 1 - 2 \exp\left(-\frac{6z^2}{n^3 + n^2}\right) \quad (3-8)$$

Il s'agit donc encore une fois d'un test non paramétrique. Il permet d'illustrer la façon dont sont construits la quasi-totalité des tests de rupture : on choisit un point de rupture k , puis on calcule une statistique permettant de quantifier les différences de part et d'autre de la rupture. La statistique de test sera finalement égale au maximum de ces différentes quantités obtenues pour toutes les valeurs possibles de k . La difficulté vient évidemment du calcul de la loi d'une telle statistique, qui est le maximum de variables aléatoires fortement dépendantes.

Le test de régression linéaire

Ce test est basé sur le modèle paramétrique suivant :

$$X_t = \alpha + \beta t + \varepsilon_t, \quad (3-9)$$

où les erreurs ε_t suivent une loi normale centrée de variance σ^2 . Par la méthode des moindres carrés, les paramètres sont estimés de la manière suivante :

$$\hat{\beta} = \frac{\sum_{t=1}^n (t - \bar{t}) X_t}{\sum_{t=1}^n (t - \bar{t})^2}, \quad \hat{\alpha} = \bar{X} - \hat{\beta} \bar{t} \quad (3-10)$$

La variance de l'estimateur de tendance peut être estimée par :

$$\text{Var}(\hat{\beta}) = \frac{12 \sum_{t=1}^n (X_t - \hat{\alpha} - \hat{\beta} t)^2}{(n-2)n(n^2-1)} \quad (3-11)$$

Le test de la régression linéaire consiste alors à vérifier que l'estimateur du coefficient directeur β est proche de zéro. Pour cela, on compare la statistique de test suivante aux quantiles d'une loi de Student à $n-2$ degrés de liberté :

$$Z = \frac{\hat{\beta}}{\sqrt{\text{Var}(\hat{\beta})}} \quad (3-12)$$

Il s'agit donc ici d'un test paramétrique, supposant la normalité des données, et spécialement conçu pour détecter une tendance linéaire. Ce test est cependant réputé robuste à une distribution qui serait éventuellement non Gaussienne.

Le test de Buishand

On suppose encore une fois que les données sont indépendantes et Gaussiennes. Le test est construit à partir des écarts cumulés à la moyenne jusqu'à un certain rang k :

$$U(k) = \sum_{i=1}^k (X_i - \bar{X}) \quad (3-13)$$

Sous l'hypothèse H_0 , l'écart cumulé devrait osciller autour de zéro, sans atteindre de trop grandes valeurs. A l'inverse, une valeur trop importante en un point k peut être le signe d'une rupture en cet endroit. La statistique de test est donc finalement :

$$Z = \max_{k=1, \dots, n} \left(\frac{|U(k)|}{\sqrt{n}\sigma_X} \right) \quad (3-14)$$

L'instant de rupture peut encore une fois être estimé par :

$$\hat{k} = \text{Arg max}_{k=1, \dots, n} \left(\frac{|U(k)|}{\sqrt{n}\sigma_X} \right) \quad (3-15)$$

La distribution de cette statistique a été tabulée grâce à des simulations Monte Carlo par Buishand (1982) (Tableau 3-2)

n	α		
	0.1	0.05	0.01
10	1.05	1.14	1.29
20	1.1	1.22	1.42
30	1.12	1.24	1.46
40	1.13	1.26	1.5
50	1.14	1.27	1.52
100	1.17	1.29	1.55
∞	1.22	1.36	1.63

Tableau 3-2. Valeurs critiques de la statistique de Buishand

I.1.3. Autres tests

On peut trouver en parcourant la bibliographie un très grand nombre de méthodes dédiées à la détection de changements. Le but n'étant pas ici de réaliser un recensement exhaustif de ces techniques, nous nous contentons de signaler ci-dessous quelques méthodes assez fréquemment utilisées, que nous évaluerons ultérieurement dans les études de robustesse et de puissance.

Tests non paramétriques :

- ✓ Les tests du changement de signe et des points de rebroussement (Kendall et Stuart, 1943), dédiés à la détection de tendances.
- ✓ Le test de Mc Gilchrist et Woodyer (1975), basé sur les écarts cumulés à la médiane.
- ✓ Le test de Spearman, basé sur la corrélation entre le temps et les rangs des observations (Kundzewicz et Robson, 2000).
- ✓ Le test de Polanski et Check (2002), qui décrit un cadre général de détection de changements en moyenne basé sur des procédures de rééchantillonnage par permutations.

Tests paramétriques:

- ✓ Le test de Cox et Lewis (1966), utilisé pour tester la stationnarité d'un processus Poissonien (occurrence d'événements rares indépendants). Le test de Lang *et al.* (2004) a le même champ d'application.
- ✓ Les tests de Gardner (1969), construits dans un cadre Bayésien Gaussien, sont dédiés à la détection d'une rupture à date inconnue. Deux versions existent, suivant que l'*a priori* de l'instant de rupture soit une distribution uniforme ou une distribution en cloche privilégiant les ruptures centrales. Les valeurs critiques sont fournies par Buishand (1982).
- ✓ Le test de Hawkins (1977), basé sur l'hypothèse de normalité. Les valeurs critiques pour de petits échantillons sont fournies par Worsley (1979).
- ✓ Le test de Keim et Cruise (1998), basé sur une régression logistique, peut être utilisé pour tester la stationnarité du temps d'attente entre deux événements rares.
- ✓ Les tests de déviance, ou du rapport des vraisemblances maximales, qui constituent un cadre très général et que nous étudierons en détail ultérieurement (Coles, 2001; Mestre, 2003).

En plus de ces tests construits de manière assez classique, il existe d'autres techniques pour évaluer d'autres types de non-stationnarité. Par exemple, Radziejewski *et al.* (2000) proposent d'évaluer la stationnarité d'une chronique entière de débits journaliers en utilisant des techniques dérivées de l'analyse de Fourier (randomisation de phase). D'autres auteurs se sont attachés à évaluer la stationnarité de relations pluie-débit (Maronna et Yohai, 1978; Andréassian *et al.*, 2003), ce qui peut s'avérer très intéressant afin de vérifier le déterminisme climatique de changements observés sur les débits.

I.1.4. Le test du rapport des vraisemblances maximales

La théorie de la vraisemblance est centrale en statistiques, car elle offre un cadre très général pouvant s'appliquer à une très grande variété de situations. Dans le contexte de la détection de changements, il est également possible de tirer profit du comportement probabiliste des échantillons que nous souhaitons tester pour essayer d'améliorer l'efficacité des méthodes. Le test du rapport des vraisemblances maximales (noté LR pour Likelihood Ratio), parfois appelé test de déviance, permet d'utiliser une méthode paramétrique adaptable à tout type de distribution. Cependant, ce test a été étonnamment peu utilisé pour détecter des évolutions liées au changement climatique, jusqu'à récemment, où plusieurs auteurs rappellent son existence et ses avantages (Jaruskova, 1997; Coles, 2001; Mestre, 2003; Zhang *et al.*, 2004; Parey *et al.*, 2006). Nous allons ici présenter cette méthode, ainsi que quelques adaptations que nous avons effectuées dans le cas de la détection de rupture.

Principe du test

On suppose deux modèles alternatifs M_0 et M_1 , de dimensions d_0 et d_1 , M_0 étant emboîté dans M_1 (c'est à dire que l'on obtient M_0 en prenant une valeur particulière des paramètres de M_1). On note $L(\mathbf{X}; \boldsymbol{\theta})$ la log-vraisemblance d'un échantillon \mathbf{X} . Il est alors possible de démontrer que, sous H_0 :

$$-2 \left(L_0(\mathbf{X}; \hat{\boldsymbol{\theta}}_0) - L_1(\mathbf{X}; \hat{\boldsymbol{\theta}}_1) \right) \xrightarrow{n \rightarrow \infty} \chi^2_{(d_1 - d_0)}, \quad (3-16)$$

les vecteurs de paramètres étant estimés par maximum de vraisemblance (MV). Ce test est dès lors intéressant pour évaluer des changements sur les paramètres des distributions.

Test LR-tendance

Le modèle M_I consiste à appliquer une tendance linéaire sur tout ou partie des paramètres (Coles, 2001). Les modèles sont bien emboîtés puisque on obtient M_0 en prenant les coefficients directeurs des tendances égaux à 0. Pour la plupart des distributions que nous utiliserons, il est impossible d'obtenir une expression analytique des estimateurs MV. Le recours à une méthode d'optimisation numérique est alors indispensable. Programmée sous Matlab à l'aide de la fonction *fminsearch*, une telle procédure parvient dans la plupart des cas à trouver un maximum -au moins local- de la vraisemblance, en choisissant comme points de départ pour les paramètres les estimateurs des moments et des valeurs nulles pour les coefficients directeurs des tendances. Dans le cas de changements plus complexes (polynomiaux par exemple), les points de départ peuvent être obtenus par des estimations non paramétriques de type spline par exemple (Parey *et al.*, 2006). Mentionnons également que des problèmes d'optimisation peuvent survenir avec des petits échantillons (de l'ordre d'une vingtaine d'observations), dont l'utilisation est de toutes façons à proscrire car loi de la statistique de test est asymptotique. Précisons enfin que dans le cas des lois des valeurs extrêmes GEV et GPD, nous supposons que le paramètre de forme reste constant, étant donné la difficulté à estimer ce paramètre, même en l'absence de toute tendance. Tous les autres paramètres (moyennes, variances, paramètres de position ou d'échelle) seront en revanche susceptibles d'être affectés par une tendance.

Test LR-rupture

Il est également possible de supposer une rupture à une date connue k . Si la rupture concerne l'ensemble des paramètres, l'estimation est ici plus simple, puisque :

$$\sup_{\theta_I, \theta'_I} L_1(\mathbf{X}; \theta_I, \theta'_I) = L_0(x_1, \dots, x_k; \hat{\theta}_I) + L_0(x_{k+1}, \dots, x_n; \hat{\theta}'_I) \quad (3-17)$$

Pour une loi dont on sait estimer les paramètres par maximum de vraisemblance (analytiquement ou numériquement), estimer les paramètres du modèle M_I ne pose aucune difficulté (estimateurs MV sur des sous-échantillons). Dans notre cas, ce test n'est cependant pas utilisable car nous ne connaissons pas *a priori* la date de rupture.

Dans le test que nous nommerons LR-rupture, on suppose une rupture à une date inconnue. Supposons pour commencer que la rupture affecte tous les paramètres. Dans ce cas :

$$\sup_{\theta_I, \theta'_I, k} L_1(\mathbf{X}; \theta_I, \theta'_I, k) = \max_k \left(L_0(x_1, \dots, x_k; \hat{\theta}_I) + L_0(x_{k+1}, \dots, x_n; \hat{\theta}'_I) \right) \quad (3-18)$$

Encore une fois, l'estimation du modèle M_I semble plus simple, puisqu'il suffit d'estimer un modèle stationnaire sur des sous-échantillons, puis de choisir la date ayant conduit à la vraisemblance maximale. Cependant, les modèles ne sont plus emboîtés, car on obtient le modèle M_0 en prenant $\theta_I = \theta'_I$ et k quelconque. L'approximation par une loi du χ^2 ne sera donc plus valable. Nous allons nous intéresser dans les parties suivantes à la loi de cette statistique de test, que nous noterons Z_n .

I.1.5. Approximation de la loi du rapport des vraisemblances maximales Z_n dans le cas du test LR-rupture

De nombreux auteurs se sont intéressés à cette question, dans le cas de distributions particulières (Sen et Srivastava, 1975; Hawkins, 1977; Worsley, 1983; Horvath, 1989; James *et al.*, 1992; Horvath, 1993). Gombay et Horvath (1994; 1996a; 1996b; 1997) ont obtenu récemment les résultats très généraux suivants :

Formule Asymptotique

$$\lim_{n \rightarrow \infty} P\left(a(\log n)\sqrt{Z_n} \leq t + b_d(\log n)\right) = \exp(-2e^{-t}), \text{ où :}$$

$$d = d_1 - d_0 \text{ la différence de dimension entre les deux modèles,} \quad (3-19)$$

$$a(x) = \sqrt{2 \log x}, \text{ et } b_d(x) = 2 \log x + \frac{d}{2} \log(\log x) - \log \Gamma(d/2)$$

ce qui permet de calculer une valeur critique.

On reconnaît une loi des valeurs extrêmes pour le comportement de la statistique, ce qui est cohérent avec le fait que la variable Z_n s'écrit comme le maximum de n variables aléatoires (équation (3-18)). Cependant, la convergence vers ce type de loi est connue pour être très lente, d'autant plus que les différentes variables aléatoires dont on prend le maximum sont ici loin d'être indépendantes. L'approximation proposée ci-dessus risque de n'être valable que pour des valeurs de n très grandes (plusieurs milliers d'observations...).

Pour remédier à ce dernier inconvénient, les auteurs proposent d'approcher, pour un risque α donné, la valeur critique z de $\sqrt{Z_n}$ de la manière suivante :

Formule Brownienne

$$z \text{ solution de } \frac{z^d \exp(-z^2/2)}{2^{d/2} \Gamma(d/2)} \left\{ T - \frac{d}{z^2} T + \frac{4}{z^2} \right\} = \alpha, \text{ avec:} \quad (3-20)$$

$$d = d_1 - d_0, \quad h = \frac{(\log n)^{3/2}}{n}, \text{ et } T = \log\left(\frac{(1-h)^2}{h^2}\right)$$

Le problème est à présent de savoir laquelle de ces deux formules utiliser. Pour répondre à cette question, nous avons procédé de la manière suivante : à partir de 1000 simulations stationnaires suivant des distributions variées, nous avons calculé 1000 statistiques Z_n , puis le 0.9-quantile de ces 1000 statistiques. Nous avons ainsi pu vérifier la proximité de ces quantiles avec les valeurs critiques proposées par chacune des deux formules décrites précédemment ($\alpha=0.1$). Les résultats présentés en Figure 3-1 montrent que la formule asymptotique semble indiquée pour les lois à deux (Gauss, Gumbel) ou trois (GEV) paramètres. En revanche, la formule Brownienne donne de bien meilleurs résultats pour les lois à un paramètre (Poisson, exponentielle).

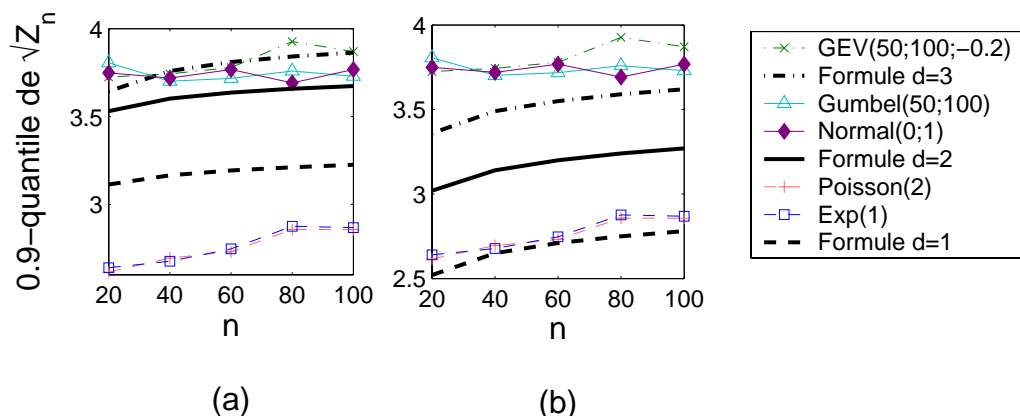


Figure 3-1. Comparaison des quantiles simulés de $\sqrt{Z_n}$ avec les valeurs critiques calculées (a) par la formule asymptotique, (b) par la formule Brownienne. Les résultats pour $d=1, 2$ et 3 sont respectivement présentés en tirets, traits pleins et tirets-pointillés.

Sur la base de ces observations, l'abaque représenté en Figure 3-2 peut être utilisé pour le calcul des valeurs critiques du test avec $\alpha=0.1$, en utilisant la formule Brownienne pour $d=1$ et la formule asymptotique pour $d=2$ et $d=3$.

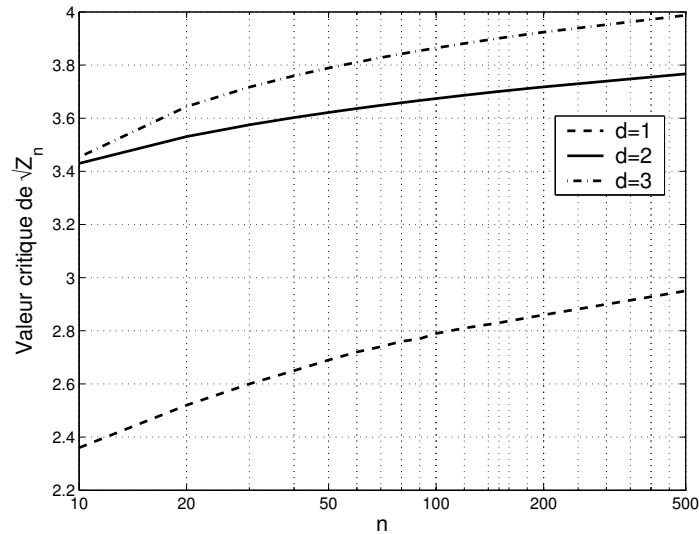


Figure 3-2. Valeurs critiques ($\alpha=0.1$) de la statistique de test $\sqrt{Z_n}$.

I.1.6. Quelques modifications du test LR-rupture

Nous avons souhaité apporter quelques modifications au test afin de gagner en puissance de détection. Tout d'abord, comme nous l'avons fait pour le test LR-tendance, nous supposons le paramètre de forme constant pour les lois des valeurs extrêmes. Ceci implique qu'il faut reprogrammer les procédures de maximisation, puisque les estimateurs du modèle M_l ne peuvent plus être obtenus par les algorithmes classiques utilisés sur des lois stationnaires. Encore une fois, ces procédures sont efficaces sous Matlab avec les lois que nous utilisons. De plus, nous avons remarqué que sur des séries stationnaires, l'instant de rupture estimé n'était pas réparti uniformément mais se concentrait aux extrémités de la chronique, comme l'illustre la Figure 3-3, réalisée à l'aide de 1000 simulations de séries exponentielles stationnaires de taille 80. Ceci s'explique par le fait que, pour une rupture précoce par exemple, les paramètres avant rupture sont estimés avec très peu de données, et donc avec beaucoup d'incertitude. La contribution de ces ruptures non centrales aux valeurs fortes de la statistique de test est donc importante, ce qui augmente les valeurs critiques et laisse présager une perte de puissance.

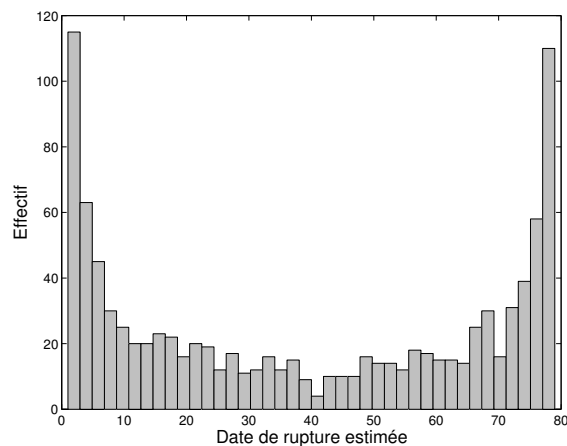


Figure 3-3. Distribution des dates de rupture estimées sur des séries stationnaires

Une alternative consiste à changer les hypothèses du test, en supposant que le changement ne peut pas se situer sur les m premiers ou sur les m derniers points de la chronique. Il faut cependant recalculer le test, puisque la statistique de test sera alors calculée en prenant le maximum des rapports de vraisemblances sur un nombre plus restreint de valeurs. Ce recalage est possible car la loi de la statistique de test ne dépend que du nombre d de paramètres soumis à la rupture et du nombre de données. Nous avons donc réalisé un nouvel abaque des valeurs critiques pour $m=10$, à l'aide de 5000 simulations de données stationnaires suivant une loi exponentielle ($d=1$) et une loi Gaussienne ($d=2$) (Figure 3-4). La valeur choisie pour m est en partie arbitraire, le but étant d'obtenir un compromis entre augmentation de puissance et perte de généralité de l'hypothèse que nous testons. Un autre choix reste toujours possible, mais il faudra dans ce cas recalculer par simulations de nouvelles valeurs critiques. Notons que l'effet de bord décrit ci-dessus a été étudié par Jaruskova (1997). Une alternative au retrait des points situés aux extrémités de l'échantillon est de pondérer le rapport de vraisemblance calculé pour une date de rupture k par une fonction donnant plus de poids aux valeurs de k proches de $n/2$. Le choix le plus classique est $w(k) = \sqrt{\frac{k}{n(1-k/n)}}$.

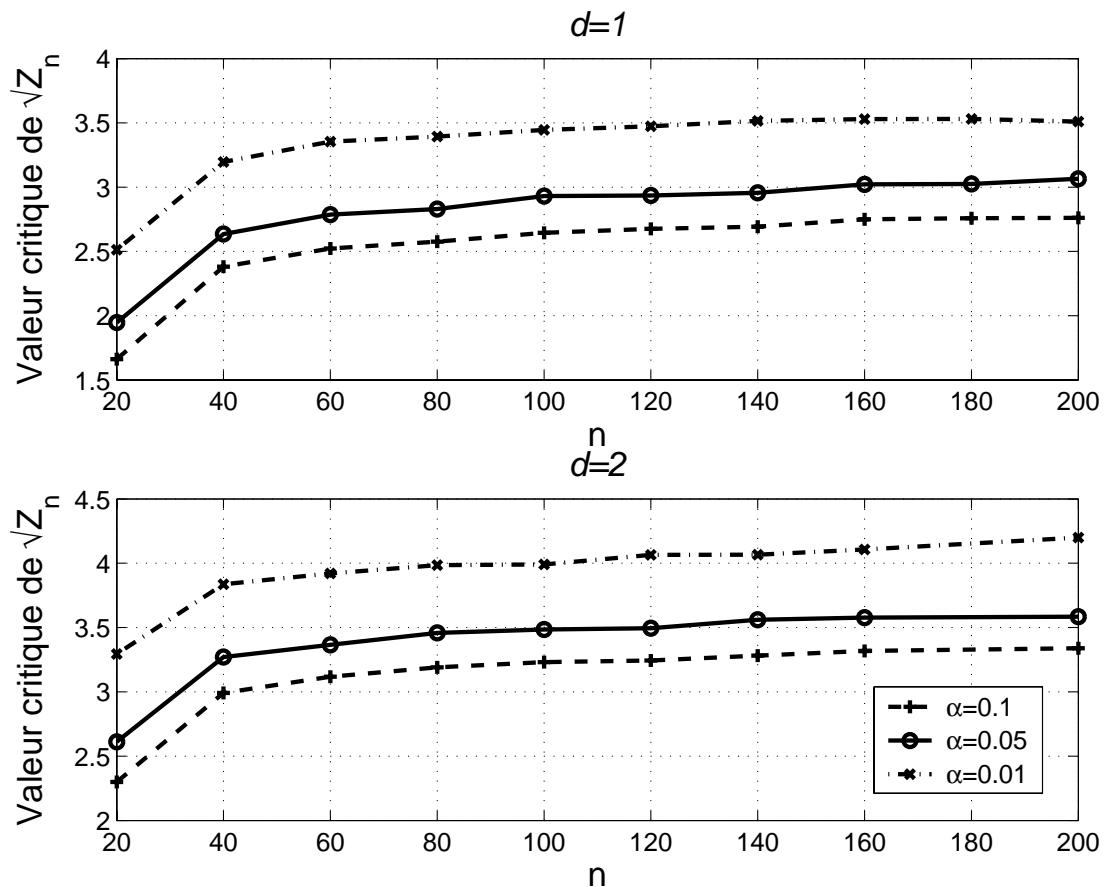


Figure 3-4. Valeurs critiques de la statistique de test modifiée en supposant que le changement n'est situé ni sur les 10 premières ni sur les 10 dernières valeurs (n est la taille de l'échantillon total).

Comme attendu, on gagne effectivement en puissance, comme le montre la Figure 3-5, réalisée à partir de 1000 simulations Gaussiennes de taille 80 (rupture centrale, $\alpha=0.1$). Ce gain de puissance est au détriment de la généralité de l'hypothèse que nous testons. Cependant, il semble totalement utopique de vouloir détecter des ruptures très précoces ou très tardives, surtout en utilisant des lois des valeurs extrêmes, dont les paramètres sont connus pour être très délicats à estimer. Ceci étant particulièrement vrai pour le paramètre de forme, il nous semble donc également illusoire de rechercher un changement sur ce paramètre.

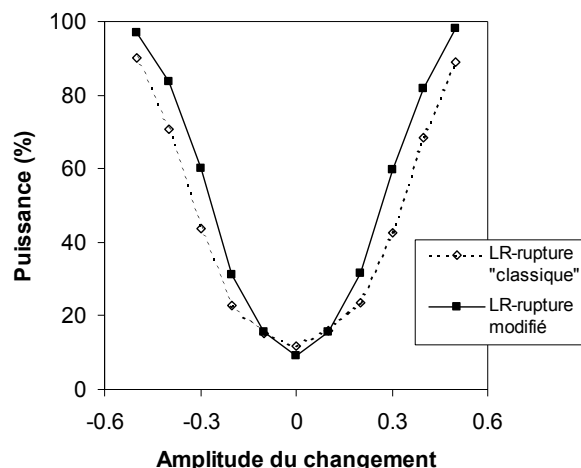


Figure 3-5. Puissance des tests LR-rupture classique et modifié

I.2. Etude de robustesse

La première qualité d'un test est de respecter le risque de première espèce spécifié. Or, les valeurs critiques ne sont en général valables que sous certaines conditions d'application (nombre de données, échantillons *iid*, distribution, ...). Si ces contraintes ne sont pas respectées, le test pourra avoir tendance à rejeter l'hypothèse H_0 trop souvent (test libéral) ou au contraire trop rarement (test conservatif). On aura tendance à préférer un test conservatif à un test libéral, car le risque de première espèce sera en réalité plus faible que celui spécifié (au détriment de la puissance). Si le risque reste stable même en dehors des conditions d'application requises, le test sera dit robuste. Cette partie va évaluer la robustesse de différents tests afin de recenser les hypothèses qui doivent absolument être respectées, et celles qui au contraire peuvent être assouplies.

I.2.1. Robustesse aux petits échantillons

Dans tout test statistique, la règle de décision est basée sur la comparaison entre la statistique de test calculée sur les données et les quantiles de la loi théorique sous H_0 de cette statistique. Cette loi peut être calculée de manière exacte, être obtenue par des simulations, ou être asymptotique. Dans ce dernier cas, il existe donc un risque à utiliser de trop petits échantillons. Ce danger est illustré par la partie gauche de la Figure 3-6, obtenue après 1000 simulations de données Gaussiennes. On observe notamment que les tests non paramétriques de Pettitt et de McGilchrist et Woodyer sont trop conservatifs, alors que le test LR-tendance est trop libéral.

On observe le même phénomène lorsque les données sont issues d'une loi des valeurs extrêmes (Figure 3-6, droite), signe que l'approximation par un χ^2 de la statistique des tests LR-tendance n'est valable qu'à partir d'une cinquantaine de données. Précisons cependant que dans le cadre de ces simulations, les procédures de maximisation de la vraisemblance entièrement automatisées ne parviennent pas toujours à trouver un maximum local avec de si

petits échantillons, ce qui peut légèrement biaiser les résultats. Dans la pratique, il sera possible de tester différents points de départ pour les algorithmes afin de contourner ce problème.

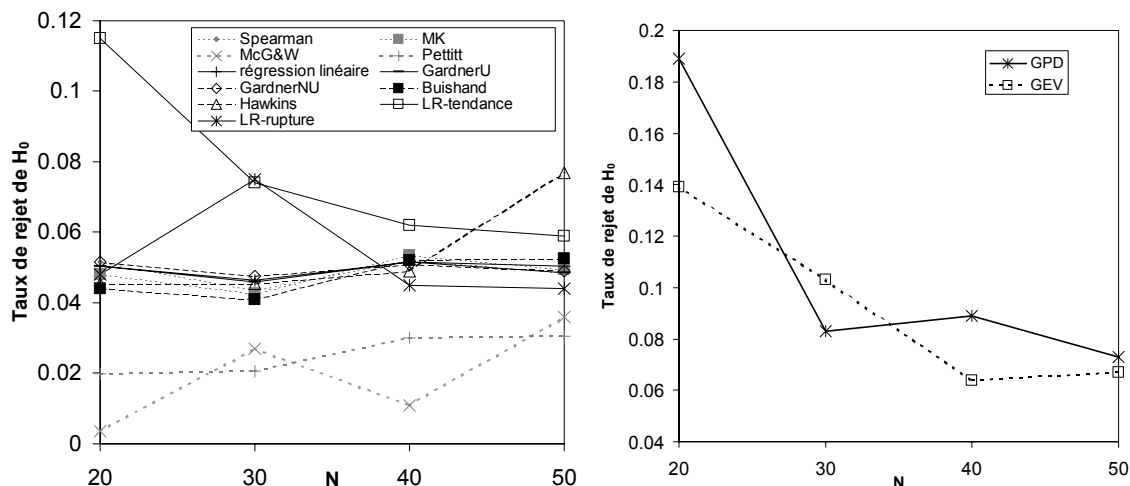


Figure 3-6. Taux de rejet de H_0 pour différentes tailles d'échantillon ($\alpha=0.05$). Gauche : données Gaussiennes. Les tests non paramétriques sont représentés en grisé. Droite : données issues de lois des valeurs extrêmes GEV et GPD.

I.2.2. Erreur sur la distribution

La plupart des tests statistiques utilisés pour la détection de changements sont réputés assez robustes. Plus spécifiquement, l'hypothèse la plus courante étant que les données suivent une loi normale, une certaine robustesse à la non-normalité a été démontrée par certains auteurs (Kim et Cai, 1993; Lubès-Niel *et al.*, 1998; Lemaitre, 2002). Il convient cependant de relativiser ce résultat, comme le montre la Figure 3-7 : les tests du type rapport de vraisemblances, pour lesquels moyenne et variance de la loi Gaussienne sont susceptibles de changer, sont très sensibles à l'asymétrie de la distribution. Dans une moindre mesure, le test de Hawkins tend également à devenir trop libéral pour des distributions fortement asymétriques. Les lois simulées ont ici des asymétries respectivement égales à 0, 1.14, 3.5, 4.72 et $+\infty$ (divergence du moment d'ordre 3). Les tests ont été appliqués avec un risque $\alpha=0.05$.

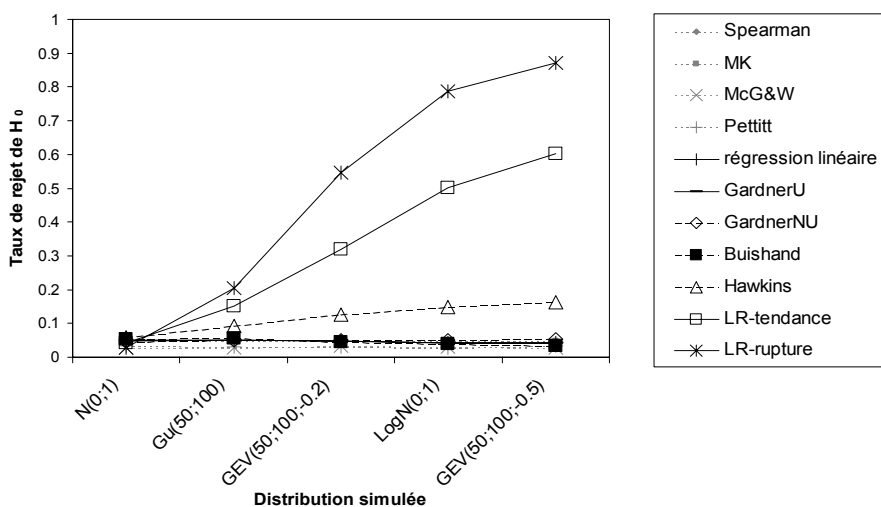


Figure 3-7. Robustesse des tests à la non-normalité ($\alpha=0.05$).

Pour remédier à ce problème, il est possible de modifier les données par une transformation de type « normal score ». Ceci revient alors à baser le test sur les rangs des observations. L'intérêt de cette manipulation est donc faible, et il semble préférable, lorsqu'un doute subsiste sur la distribution des valeurs, d'utiliser directement un test non paramétrique.

Le Tableau 3-3 résume différentes erreurs de distribution et y associe le taux de rejet de H_0 estimé par 1000 simulations stationnaires ($N=80$). Nous avons utilisé ici les tests LR-tendance ($\alpha=0.05$). Les asymétries des lois simulées sont respectivement de 0, 1.14, 2, 3.5 et 3.6.

Loi Modélisée Loi simulée	Normale	Gumbel	Expo	GEV	GPD
Normale(100, 10)	0.059	0.121	0	0.077	0.087
Gumbel(50, 100)	0.148	0.056	0	0.041	0.345
Expo(10)	0.252	0.01	0.053	0.026	0.058
GEV(50, 100, -0.2)	0.353	0.019	0.005	0.053	0.226
GPD(50, -0.2)	0.418	0.006	0.122	0.035	0.067

Tableau 3-3. Taux de rejet de H_0 par les tests LR-tendance ($\alpha=0.05$)

Sur la diagonale de ce tableau, les taux proches de 0.05 illustrent le fait que les tests sont correctement calés. Hors de la diagonale, les taux s'éloignent parfois significativement de cette valeur, ce qui montre que les erreurs de distribution ne sont pas forcément anodines. Plus particulièrement, on peut observer qu'en supposant une loi exponentielle alors que les données suivent une loi GPD, le test est beaucoup trop libéral (12.2% de rejet). Inversement, le test devient trop conservatif (1.9% de rejet) si l'on modélise une loi de Gumbel sur des données issues d'une loi GEV. Par contre, les taux de rejet restent proche de 5% lorsqu'on modélise une loi GPD (resp. GEV), même si les données sont issues d'une loi exponentielle (resp. Gumbel). Dans le cadre de l'analyse des valeurs extrêmes, ceci nous incite à utiliser des tests LR basés sur des lois à trois paramètres pour éviter de biaiser la significativité des tests par d'éventuels comportements hyper-exponentiels.

Ce manque de robustesse s'observe également sur les lois à valeurs entières. En hydrologie, ce type de distribution est utilisé par exemple pour le dénombrement des événements de crue. Dans le cadre de l'échantillonnage par valeurs supérieures à un seuil, il est fréquent de supposer que le nombre d'événements sélectionnés par unité de temps (l'année, en général) suit une loi de Poisson (Lang *et al.*, 1999), pour laquelle moyenne et variance sont égales. Cette contrainte étant parfois trop restrictive, d'autres processus peuvent être choisis (cf. Yagouti *et al.* (2001) pour une revue), notamment la loi Binomiale (variance inférieure à la moyenne) ou la loi Binomiale négative (variance supérieure à la moyenne). La Figure 3-8 illustre le manque de robustesse au non-respect de l'hypothèse Poissonienne (1000 simulations, $N=80$). Les distributions utilisées pour ces simulations sont décrites dans le Tableau 3-4, et sont toutes de moyenne 2. Nous avons utilisé un test LR-tendance ($\alpha=0.05$) basé sur une loi de Poisson.

Loi simulée	Binomiale				Poisson	Binomiale négative		
Paramètre(s)	(2.5, 0.8)	(5, 0.4)	(8, 0.25)	(10, 0.2)	2	(8, 0.8)	(2, 0.5)	(0.5, 0.2)
Indice de dispersion (Var/Moy)	0.2	0.6	0.75	0.8	1	1.25	2	5

Tableau 3-4. Distributions utilisées pour la simulation de données entières

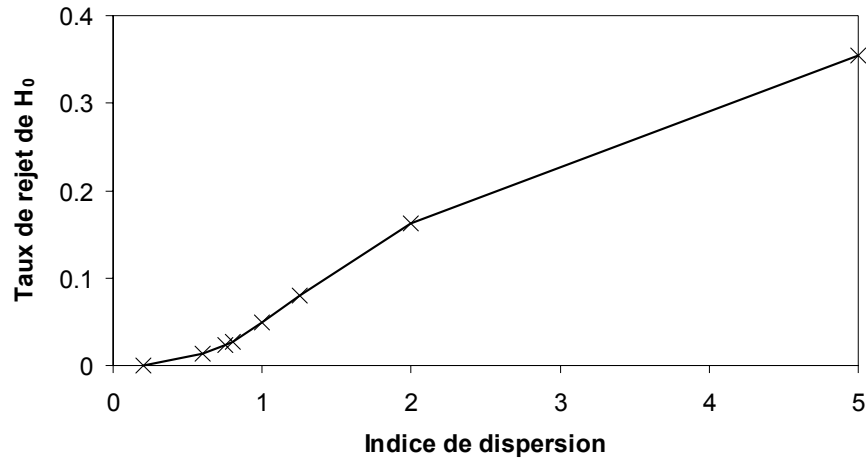


Figure 3-8. Taux de rejet de H_0 en fonction de la variance du processus ($\alpha=0.05$). Les données sont issues d'une loi Binomiale (dispersion <1), d'une loi de Poisson (dispersion $=1$) ou d'une loi Binomiale négative (dispersion >1).

I.2.3. Présence d'autocorrélation

L'autocorrélation est connue pour perturber significativement le déroulement des tests statistiques (Woodward et Gray, 1993; Von Storch et Navarra, 1999; Yue *et al.*, 2002c; Yue et Wang, 2002a; Yue et Pilon, 2003; Yue *et al.*, 2003; Yue et Wang, 2004). La Figure 3-9 illustre l'effet d'une autocorrélation positive d'ordre 1 sur le risque de première espèce. Nous avons simulé des données Gaussiennes centrées-réduites $(\varepsilon_i)_{i=1,\dots,80}$, puis appliqué le modèle AR(1) suivant :

$$\begin{cases} Y_1 = \varepsilon_1 \\ Y_i = \rho Y_{i-1} + \varepsilon_i, \quad i = 2, \dots, 80 \end{cases} \quad (3-21)$$

Les tests ont été appliqués à la série des Y_i au risque $\alpha=0.05$. Il apparaît clairement que l'utilisation d'un test sur ce type de données conduit à rejeter beaucoup trop souvent l'hypothèse nulle. Une autocorrélation négative induit l'effet inverse, c'est à dire que l'on ne rejette que très peu H_0 (Yue *et al.*, 2002c).

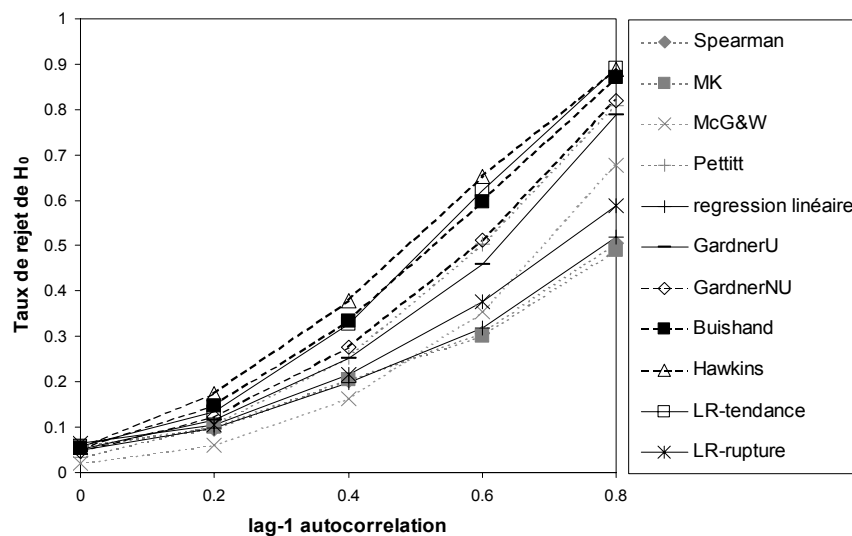


Figure 3-9. Effet de l'autocorrélation sur le risque de première espèce ($\alpha=0.05$).

Ce phénomène se comprend facilement en considérant l'exemple suivant, illustré par la Figure 3-10. Les séries générées sont un bruit blanc Gaussien, un processus AR(0.8) et un processus AR(-0.8). On observe que la présence d'une autocorrélation positive crée des tendances apparentes, alors qu'une autocorrélation négative semble rendre la série d'apparence plus stationnaire.

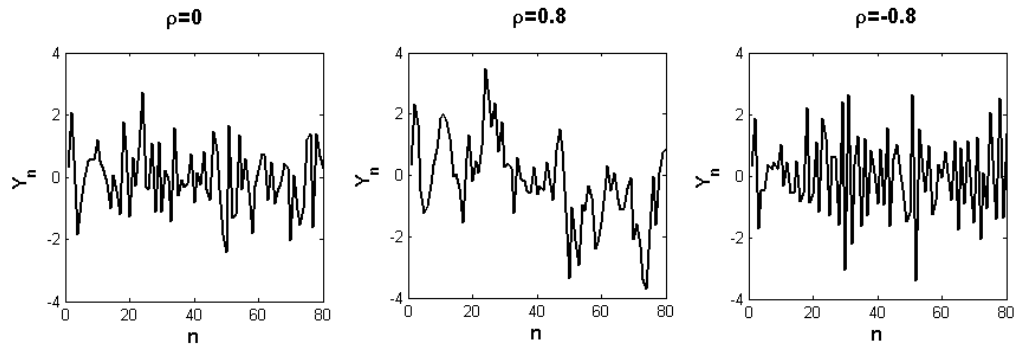


Figure 3-10. Influence de l'autocorrélation sur l'allure d'une série chronologique

Le problème de distinguer ce qui relève de l'autocorrélation ou d'une tendance déterministe dans la détection de non-stationnarité est un problème quasiment insoluble, comme le note Jaruskova (1997): « *if the finite part of a time series is observed, it is impossible to distinguish between a stationary series with the positive dependence between the neighbouring observations and a sequence of independent variables with the slowly changing mean* ». Ceci implique qu'il faut faire *a priori* le choix d'un modèle qui permet ou ne permet pas de prendre en compte l'autocorrélation, grâce à des considérations physiques du phénomène étudié notamment. Dans le domaine de l'hydrologie, les fortes crues sont en général considérées indépendantes les unes des autres. Par contre, d'autres variables relatives aux étiages, voire aux modules, peuvent être davantage autocorrélées, car elles sont directement reliées à des phénomènes météorologiques pouvant présenter des persistances plus importantes (circulation atmosphérique, mécanismes de recharge et vidange de nappes, etc.).

Dans le cas où la présence d'autocorrélation ne peut être écartée, il s'avère donc totalement indispensable de la prendre en compte dans l'évaluation de la significativité des changements. Plusieurs méthodes ont été proposées dans cette optique.

La transformation PW (pre-whitening)

Von Storch et Navarra (1999) proposent ainsi de « blanchir » les observations avant l'application du test. En effet, si l'on suppose que la série brute suit un modèle auto-régressif d'ordre 1 ($Y_i = \rho Y_{i-1} + \varepsilon_i$), alors la série transformée $X_i = Y_i - \rho Y_{i-1}$ est un bruit blanc. Dans la pratique, ρ étant inconnu, on le remplace par son estimateur classique :

$$\hat{\rho} = \frac{\frac{1}{n-1} \sum_{i=1}^{n-1} (Y_i - \bar{Y})(Y_{i+1} - \bar{Y})}{\frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (3-22)$$

En utilisant cet estimateur pour la transformation de données, on ajoute une part d'incertitude liée à l'estimation. Pour de petits échantillons, ceci pourrait avoir tendance à rendre le test trop libéral.

Cette procédure est connue sous le nom de pre-whitening (PW). Evidemment, elle est insuffisante si le modèle AR(1) est trop simpliste pour décrire la structure d'autocorrélation. Cependant, des modèles plus complexes (AR(p) ou ARMA(p,q)) poseraient des problèmes d'estimation des paramètres, surtout sur les séries hydrologiques que nous envisageons de tester, qui ne compteront au mieux qu'une centaine d'observations.

Cette technique est efficace pour conserver un taux de rejet de l'hypothèse nulle proche du risque de première espèce fixé, pour une taille d'échantillon de 80, comme le montre la Figure 3-11 ($\alpha=0.05$). Par contre, avec un échantillon de taille 20 (Figure 3-12), on observe que l'efficacité de cette technique diminue dans le cas de petits échantillons. Les tests deviennent notamment trop libéraux pour les fortes autocorrélations. On remarque également un mauvais calage du test LR-tendance, qui s'explique par le fait que la loi de sa statistique est asymptotique.

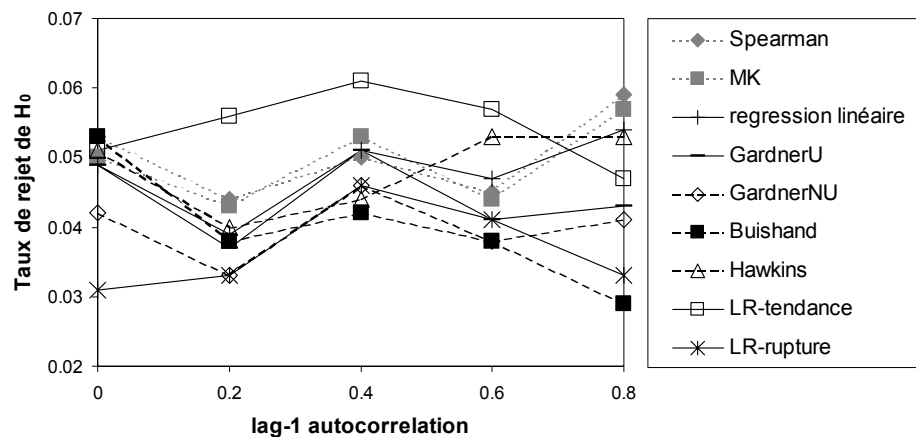


Figure 3-11. Taux de rejet de H_0 après transformation PW. $N=80$, $\alpha=0.05$.

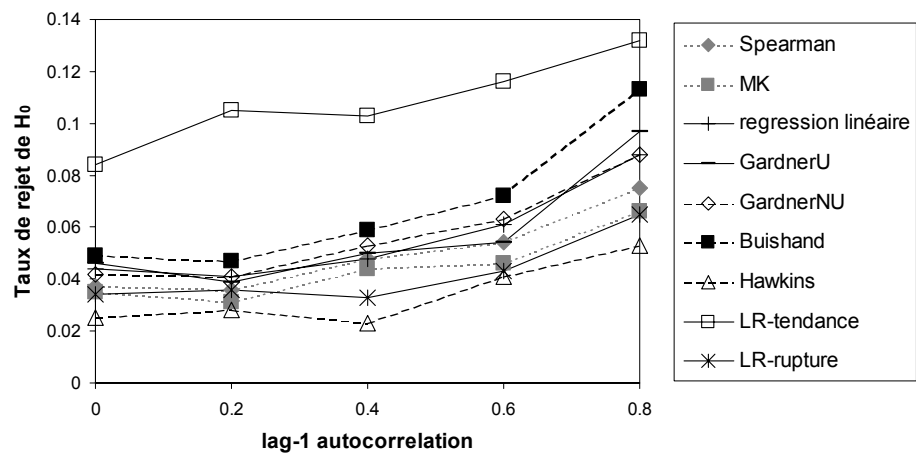


Figure 3-12. Taux de rejet de H_0 après transformation PW. $N=20$, $\alpha=0.05$

La transformation TFPW (trend-free pre-whitening)

Dans une série d'articles, Yue *et al.* (Yue *et al.*, 2002c; Yue et Wang, 2002a; Yue et Pilon, 2003; Yue *et al.*, 2003; Yue et Wang, 2004) critiquent cette procédure en démontrant que la transformation PW n'est pas une opération neutre dans le cas où autocorrélation et tendance sont présentes dans la série à traiter. Plus précisément, la présence d'une tendance tend à créer de l'autocorrélation, l'estimateur de ρ est donc biaisé. Ceci influe alors sur la série blanchie $X_i = Y_i - \rho Y_{i-1}$: si β est le coefficient directeur de la tendance sur la série des Y_i , alors la série des

X_i présentera une tendance apparente égale à $(1-\hat{\rho})\beta$. Inversement, la présence d'autocorrélation influe sur l'estimation du paramètre β , en augmentant la variance de l'estimateur. Dans le cas de l'estimateur non paramétrique proposé par Sen (1968), la moyenne de l'estimateur est par contre inchangée, ce qui signifie que l'autocorrélation diminue la précision mais n'induit pas de biais. Sur la base de ces résultats, Yue *et al.* proposent la procédure suivante, baptisée trend-free pre-whitening (TFPW) :

1. Estimer le coefficient directeur de la tendance sur les données brutes Y_i , par l'estimateur de Sen : $\hat{\beta} = \text{médiane} \left\{ \frac{Y_i - Y_j}{i - j}, i = 2, \dots, n, j = 1, \dots, i - 1 \right\}$
2. Créer une série homogénéisée, $X_i = Y_i - \hat{\beta}i$
3. Estimer l'autocorrélation $\hat{\rho}$ d'ordre 1 sur les X_i
4. Effectuer une transformation PW sur les X_i , $Z_i = X_i - \hat{\rho}X_{i-1}$
5. La série à tester sera $W_i = Z_i + \hat{\beta}i$

De la sorte, l'estimation de l'autocorrélation ne sera pas biaisée puisque effectuée sur une série sans tendance, et la série des W_i à tester est bien constituée d'une tendance plus d'un bruit blanc.

Comme précédemment, nous avons souhaité vérifier que cette technique préservait bien le risque de première espèce. Les résultats de 1000 simulations avec $\alpha=0.05$ et $N=80$ sont présentés dans la Figure 3-13. La technique du TFPW ne semble pas efficace pour conserver un taux de rejet proche du risque de première espèce souhaité. Ceci s'explique en fait par l'augmentation de la variance de l'estimateur de Sen due à l'autocorrélation : même si la série ne présente pas de tendance, l'estimateur de β pourra, avec une probabilité plus importante que sous l'hypothèse d'indépendance, être loin de zéro. En ajoutant cette tendance surestimée à un bruit blanc, les tests détecteront à tort un changement dans un trop grand nombre de cas.

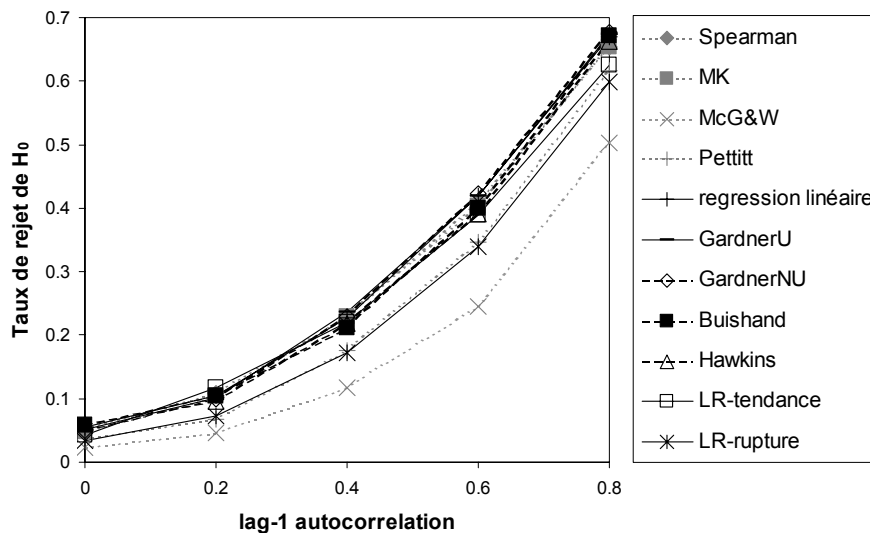


Figure 3-13. Taux de rejet de H_0 après transformation TFPW ($\alpha=0.05$)

Certains auteurs avaient déjà émis des doutes quant à l'approche proposée par Yue *et al.*, comme Zhang et Zwiers (2004) ou Bayazit et Onoz (2004). Un des arguments clé de Yue *et al.* est de constater que la tendance apparente d'une série après pre-whitening, égale à $(1-\hat{\rho})\beta$, est inférieure à la tendance réelle dans le cas d'une autocorrélation positive. Ceci

laisse présager une importante perte de puissance, qui est malheureusement inévitable, car l'autocorrélation retire de l'information aux données. Mais les auteurs ont effectué toutes leurs analyses sous l'hypothèse qu'autocorrélation et tendance coexistent dans la série à analyser. Or, dans la réalité, l'objet du test est justement de déterminer si une tendance est présente, et le respect du risque de première espèce doit être évalué sous l'hypothèse H_0 , c'est à dire sous hypothèse de stationnarité. Dans leur article de 2002, Yue *et al.* adressent le conseil suivant, à l'issue de la première étape d'estimation de la tendance : « *If the slope is almost equal to zero, then it is not necessary to continue to conduct trend analysis* ». Mais comment déterminer si la tendance est « presque » égale à zéro... si ce n'est en appliquant un test statistique ?

Le test de Mann-Kendall modifié

Une approche complémentaire à celles qui ont été présentées précédemment consiste à modifier le test à appliquer aux données plutôt que les données elles-mêmes. Hamed et Rao (1998) et Yue et Wang (2002b; 2004) ont ainsi proposé de corriger la statistique de test S de Mann-Kendall pour prendre en compte d'éventuelles autocorrélations. En effet, ces auteurs ont mis en évidence que S ne suit plus sous H_0 une loi Gaussienne centrée-réduite en présence d'autocorrélation, mais une loi $N(0, \sigma^2)$, où $\sigma^2 < 1$ si $\rho < 0$, et $\sigma^2 > 1$ si $\rho > 0$. Le problème est donc de calculer un facteur correctif γ tel que :

$$Var_{\rho}(S) = \gamma Var_{\rho=0}(S) \quad (3-23)$$

Dans cette optique, les auteurs utilisent la notion de « effective sample size », que nous traduirons par effectif efficace ou taille équivalente. Cette notion traduit le fait que l'autocorrélation positive retire de l'information. Tout se passe donc comme si l'échantillon autocorrélé de taille n se comportait comme un échantillon indépendant de taille $n^* < n$. Evidemment, le raisonnement dual peut être tenu dans le cas d'une autocorrélation négative. Le facteur correctif γ relatif à la statistique S doit pouvoir s'exprimer sous la forme $\gamma = n/n^*$.

Les approches proposées par Hamed et Rao (1998) et Yue et Wang (2002b; 2004) consistent alors à comparer la statistique $S/\sqrt{\gamma}$ calculée sur les données aux quantiles d'une loi normale centrée réduite. Les deux méthodes diffèrent par les formules utilisées pour le calcul de la taille équivalente.

Yue et Wang utilisent la relation établie par Bayley (1946) dans le cadre de l'effectif efficace relatif à la statistique \bar{X} :

$$\gamma = 1 + 2 \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \rho_k \quad (3-24)$$

où ρ_k désigne l'autocorrélation d'ordre k . Dans la pratique, il faut estimer la fonction d'autocorrélation pour calculer γ , ce qui conduit à $n-1$ estimations. Le problème est simplifié si l'on suppose que la série à tester est une réalisation d'un processus AR(1). Dans ce cas, on a en effet la relation $\rho_k = \rho_1^k$. Il suffit donc d'estimer le premier coefficient d'autocorrélation pour calculer γ . Sous cette hypothèse, l'équation (3-24) se simplifie de la manière suivante (Matalas et Langbein, 1962) :

$$\gamma = 1 + 2 \frac{\rho_1^{n+1} - n\rho_1^2 + (n-1)\rho_1}{n(\rho_1 - 1)^2} \quad (3-25)$$

Hamed et Rao proposent quant à eux une formule empirique spécifiquement calculée pour corriger la statistique de Mann-Kendall :

$$\gamma = 1 + \frac{2}{n(n-1)(n-2)} \sum_{k=1}^{n-1} (n-k)(n-k-1)(n-k-2)\rho_k \quad (3-26)$$

Les auteurs conseillent d'estimer la fonction d'autocorrélation à partir des rangs des observations. Ils suggèrent également de tester la significativité des autocorrélations estimées avant de les inclure dans la somme. Comme précédemment, il est possible de supposer directement que les données sont issues d'un processus AR(1), ce qui simplifie l'estimation grâce à la relation $\rho_k = \rho_1^k$.

Nous avons évalué l'aptitude de ces méthodes à conserver un taux de rejet proche du risque de première espèce spécifié ($\alpha=0.05$), grâce à des simulations de processus AR(1) d'effectif 60 (Figure 3-14). Les résultats montrent que ces méthodes sont assez performantes. L'hypothèse AR(1) donne de meilleurs résultats, mais on pourra cependant s'en passer si elle semble trop irréaliste, puisque les taux de rejets restent inférieurs à 0.11 si l'on prend en compte toute la fonction d'autocorrélation. La formule utilisée par Hamed et Rao semble légèrement meilleure sous l'hypothèse AR(1).

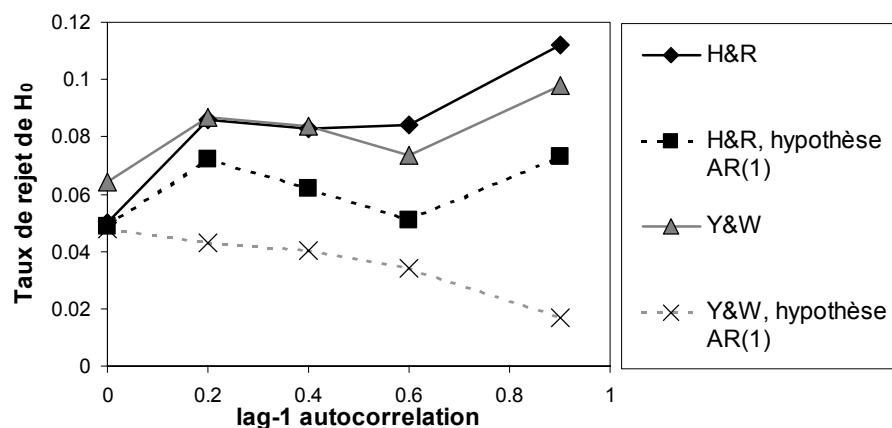


Figure 3-14. Taux de rejet de H_0 par les tests de Mann-Kendall modifiés présentés par Hamed et Rao (H&R) et Yue et Wang (Y&W) ($\alpha=0.05$)

Cependant, pour les mêmes raisons que celles invoquées par Yue *et al.* pour justifier l'utilisation de la transformation TFPW, les auteurs recommandent de retirer une estimation non paramétrique (estimateur de Sen) de la tendance aux données avant l'estimation des autocorrélations, conseil qui n'a pas été suivi ici pour la raison suivante : sous l'hypothèse H_0 , il n'y a pas de tendance sur les données, en retirer une aura donc pour effet d'ajouter inutilement une part de variance à la statistique de test, ce qui risque de décaler le test. Par contre, il est vrai que si H_0 est fautive, alors l'estimation des autocorrélations sera biaisée par l'existence d'une tendance, comme expliqué précédemment. Dans la pratique, nous ne savons pas si une tendance existe ou non (c'est justement l'objet du test), il semble donc préférable de privilégier un bon calage du test à un gain de puissance. Pour illustrer ce risque, nous avons appliqué le test de Mann-Kendall modifié par Hamed et Rao avec retrait de la tendance à des données simulées (Figure 3-15, les résultats sont identiques avec la formule de Yue et Wang). On voit clairement que le retrait d'une tendance conduit à rejeter trop souvent l'hypothèse H_0 . L'effet est moins sensible avec l'hypothèse AR(1), puisque le biais introduit en retirant à tort la tendance n'influe que sur l'estimation de l'autocorrélation d'ordre 1, et non sur toute la fonction d'autocorrélation.

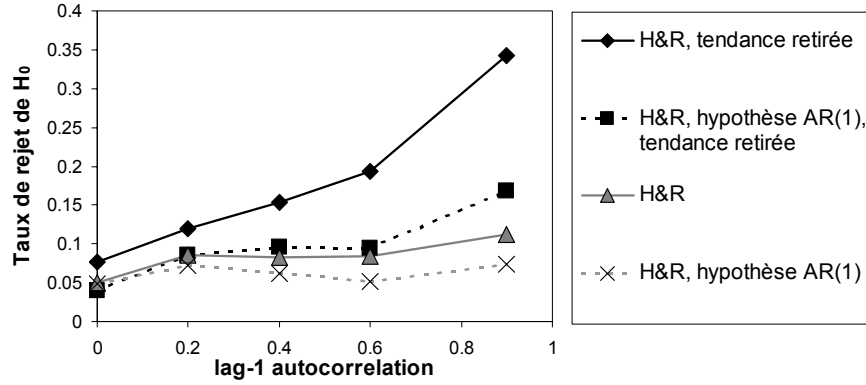


Figure 3-15. Effet du retrait d'une tendance sur le taux de rejet de H_0 ($\alpha=0.05$)

La régression linéaire modifiée

Le concept est en tout point identique à celui décrit précédemment pour le test de Mann-Kendall. Il s'agit d'adapter une statistique de test classique (ici, celle de la régression linéaire) au cas où les données présentent de l'autocorrélation. Rappelons que le modèle sous-jacent est le suivant :

$$Y_t = \alpha + \beta t + \varepsilon_t, \quad (3-27)$$

où les erreurs ε_t sont modélisées par une série temporelle Gaussienne, ici éventuellement autocorrélée (processus ARMA par exemple). Par la méthode des moindres carrés, les paramètres sont estimés de la manière suivante :

$$\hat{\beta} = \frac{\sum_{t=1}^n (t - \bar{t}) Y_t}{\sum_{t=1}^n (t - \bar{t})^2}, \quad \hat{\alpha} = \bar{Y} - \hat{\beta} \bar{t} \quad (3-28)$$

La variance de l'estimateur du paramètre de tendance β , dans le cas où les erreurs sont indépendantes, peut être estimée par :

$$Var_{\rho=0}(\hat{\beta}) = \frac{12 \sum_{t=1}^n (Y_t - \hat{\alpha} - \hat{\beta} t)^2}{(n-2)n(n^2-1)} \quad (3-29)$$

Dans le cas où les résidus ne sont pas indépendants, l'estimation de cette variance est biaisée. Woodward et Gray (1993) proposent d'utiliser la formule suivante :

$$Var(\hat{\beta}) = Var_{\rho=0}(\hat{\beta}) + \frac{288}{n^2(n^2-1)^2} \sum_{s=2}^n \sum_{t=s}^{s-1} (t - \bar{t})(s - \bar{t}) \hat{\gamma}_{s-t}, \quad (3-30)$$

où $\hat{\gamma}_k$ correspond à l'estimateur de l'autocovariance d'ordre k des résidus :

$$\hat{\gamma}_k = \frac{1}{n} \sum_{t=1}^{n-k} \hat{\varepsilon}_{t+k} \hat{\varepsilon}_t \quad (3-31)$$

La nouvelle statistique de test $\hat{\beta} / \sqrt{Var(\hat{\beta})}$ est alors comparée aux quantiles d'une loi normale centrée réduite.

De même que précédemment, il est possible de supposer directement que les données sont issues d'un processus AR(1), ce qui permet de ne pas estimer toute la fonction d'autocorrélation.

Les performances de cette approche sont illustrées dans la Figure 3-16, qui représente les résultats du test appliqué à 1000 simulations de processus AR(1) ($\alpha=0.05$). On voit dans cette figure que l'estimation de la fonction d'autocorrélation entière conduit à de très mauvaises performances. L'hypothèse AR(1) donne de résultats acceptables jusqu'à une autocorrélation de l'ordre de 0.6. On observe également que le taux de rejet est de plus en plus stable au fur et à mesure que la taille de l'échantillon augmente.

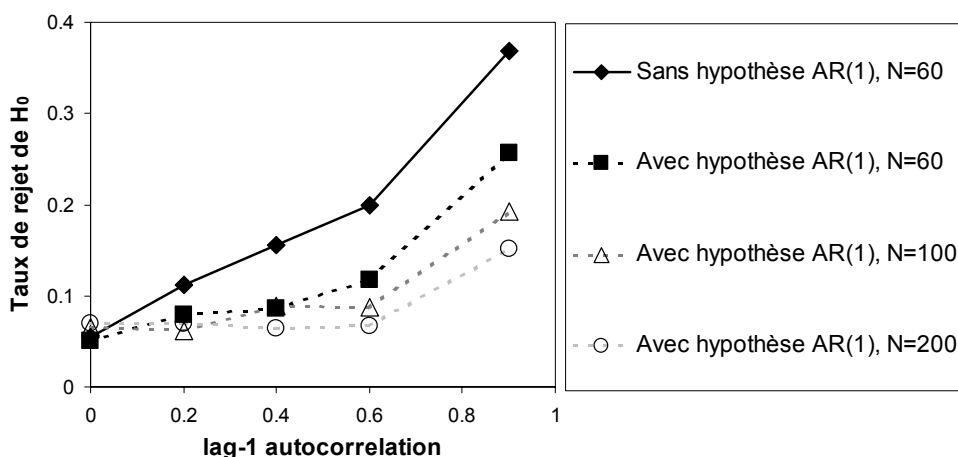


Figure 3-16. Taux de rejet de H_0 par le test de régression linéaire modifié ($\alpha=0.05$)

Quelle technique pour faire face à l'autocorrélation ?

En résumé, nous disposons finalement de trois méthodes pour prendre en compte la persistance dans les séries à traiter :

- ✓ La transformation PW : cette méthode s'est révélée très efficace pour conserver un taux de rejet acceptable. Elle présente de plus l'avantage de permettre l'utilisation de n'importe quel test après la transformation de données. Cependant, il est important de noter que la transformation $X_i = Y_i - \rho Y_{i-1}$ n'est pas neutre pour la distribution des données. Par exemple, si les Y_i sont issus d'une distribution GEV, les X_i suivront une distribution différente qu'il n'est pas possible à notre connaissance de déterminer théoriquement. Le test à utiliser après transformation PW devra donc être non paramétrique, sauf dans le cas Gaussien, puisque la distribution normale est stable par combinaison linéaire.
- ✓ Le test de Mann-Kendall modifié : nous préférons la formule proposée par Hamed et Rao, qui donne des résultats légèrement meilleurs que ceux issus de Yue et Wang. En utilisant l'hypothèse AR(1), le taux de rejet de H_0 reste très proche du risque spécifié, même avec des autocorrélations très fortes. Cette hypothèse peut cependant être abandonnée si nécessaire, l'augmentation du taux de rejet restant raisonnable.
- ✓ Le test de régression linéaire modifié : l'hypothèse AR(1) semble ici indispensable. Le taux de rejet reste proche du risque spécifié pour des autocorrélations modérées, par contre le biais semble important pour des autocorrélations élevées de l'ordre de 0.9, même avec un échantillon important.

La distinction entre ces méthodes se fera également sur la base de leur puissance, que nous étudierons au paragraphe I.3.4. Notons également que la liste de techniques proposée ici n'est pas exhaustive. Par exemple, dans le cadre de la régression linéaire modifiée, d'autres formules de correction de la variance ont été proposées, la plupart utilisant l'analyse spectrale de Fourier pour estimer la fonction d'autocorrélation (Brillinger, 1989; Bloomfield et Nychka, 1992). Quelle que soit la méthode utilisée, il ne faut cependant pas espérer une estimation précise de la totalité de la fonction d'autocorrélation avec des échantillons de taille réduite (une centaine d'observations tout au plus). Nous avons donc privilégié les méthodes les plus simples, quitte à faire des hypothèses réductrices du type AR(1) sur la structure de l'autocorrélation.

I.3. Etude de puissance

L'étude de robustesse a démontré que le respect des hypothèses était souvent indispensable pour conserver un taux de rejet de H_0 proche du risque de première espèce spécifié. Dans un second temps, le choix du test à utiliser se fera en général sur des considérations de puissance. Rappelons que la puissance est la probabilité de rejeter à raison H_0 . Dans le cadre de la détection de non-stationnarités, il s'agit donc de la probabilité de détecter correctement une tendance ou une rupture. Evidemment, cette probabilité dépend d'un certain nombre de facteurs, comme la taille de l'échantillon, l'importance du changement, la position de la rupture, la distribution de la série à traiter, etc. Nous allons dans cette partie tenter de répondre aux questions suivantes :

- ✓ Les tests paramétriques sont-ils réellement plus puissants que les non paramétriques ?
- ✓ Quels sont les tests les plus puissants dans chaque catégorie ? Quel est le rôle de l'hypothèse alternative (tendance/rupture) ?
- ✓ Certains tests sont-ils plus performants pour des ruptures situées aux extrémités de la chronique ?
- ✓ Quelle est la méthode la plus puissante dans le cas de données autocorrélées ?

Certaines de ces questions ont déjà été traitées dans la littérature, nous allons donc présenter un état de l'art en la matière, complété si nécessaire par quelques simulations complémentaires. Dans les figures des paragraphes suivants, l'amplitude des ruptures a été « normalisée » en divisant la valeur de la rupture par la durée d'observation. Cette valeur normalisée est ainsi plus facilement comparable avec le coefficient directeur d'une tendance. De plus, les ruptures sont placées au centre de la chronique, sauf indication contraire.

I.3.1. Comparaison paramétrique (P) / non paramétrique (NP)

De nombreux auteurs se sont déjà penchés sur cette question dans le cadre des tests de détection de tendance ou de rupture. Hirsch *et al.* (1991) ainsi que Yue *et al.* (2004) ont ainsi comparé les performances des tests de Mann-Kendall (NP) et de régression linéaire (P). Le test paramétrique s'avère ainsi légèrement plus puissant si les données sont Gaussiennes. Cependant, cet avantage disparaît rapidement en utilisant des distributions asymétriques ou des mélanges de populations. Lubès-Niel *et al.* (1998) obtiennent le même type de résultats en observant que le test de Buishand (P) est légèrement plus efficace pour détecter des ruptures dans des échantillons Gaussiens que les tests de Mann-Kendall ou de Pettitt (NP). Enfin, Lemaitre (2002) montre que les tests de Buishand, des écarts cumulés et de régression linéaire (P) sont un peu plus puissants que le test de Pettitt (NP) sur des données Gaussiennes.

La supériorité des tests paramétriques semble donc réelle mais limitée en cas de données Gaussiennes. Cet avantage disparaît si les données ne sont plus normales, ce qui explique pourquoi la plupart des études de stationnarité menées sur des variables hydrologiques ont

utilisé des tests non paramétriques (Mann-Kendall en majorité). Cependant, l'avantage des tests paramétriques pourrait être plus important dans le cadre de valeurs extrêmes. Zhang *et al.* (2004) ont récemment comparé la puissance des tests LR-tendance (P) pour les lois des valeurs extrêmes avec les tests classiques de régression linéaire (P) et de Mann-Kendall (NP). Il apparaît que les tests LR basés sur l'hypothèse GEV sont significativement plus puissants que le test de Mann-Kendall. Le test de régression linéaire est le moins efficace, à cause de la forte asymétrie de la distribution des données. Les auteurs recommandent de construire le test avec une hypothèse alternative de tendance sur le seul paramètre de position, afin de gagner en puissance. Le danger de cette restriction est de ne pas pouvoir détecter certains types de changements (en variance par exemple), nous préférons donc autoriser une tendance sur le paramètre d'échelle. Enfin, les auteurs démontrent l'intérêt de s'intéresser aux r plus grandes valeurs annuelles afin d'utiliser au mieux l'information extrême présente dans une chronique. Ce gain d'information peut également être obtenu en procédant à un échantillonnage par valeurs supérieures à un seuil, avec en moyenne plus d'un événement par an. L'avantage de cette dernière méthode est qu'elle assure une plus grande homogénéité de l'échantillon.

Afin d'illustrer et de résumer ces conclusions, nous avons simulé 1000 séries de taille 80 et appliqué les tests de Mann-Kendall et LR-tendance. Les séries, issues d'une loi normale ou d'une loi GEV, ont ensuite été perturbées par l'ajout d'une tendance linéaire. Il apparaît alors sur la Figure 3-17 que dans le cas Gaussien, le test paramétrique n'est pas plus puissant que son homologue non paramétrique. Notons que le test LR-tendance autorise un changement en moyenne et en variance, ce qui requiert l'estimation d'un paramètre supplémentaire. Par contre, dans le cas GEV, le gain de puissance peut atteindre 15 points en utilisant un test paramétrique approprié, ce qui est loin d'être négligeable.

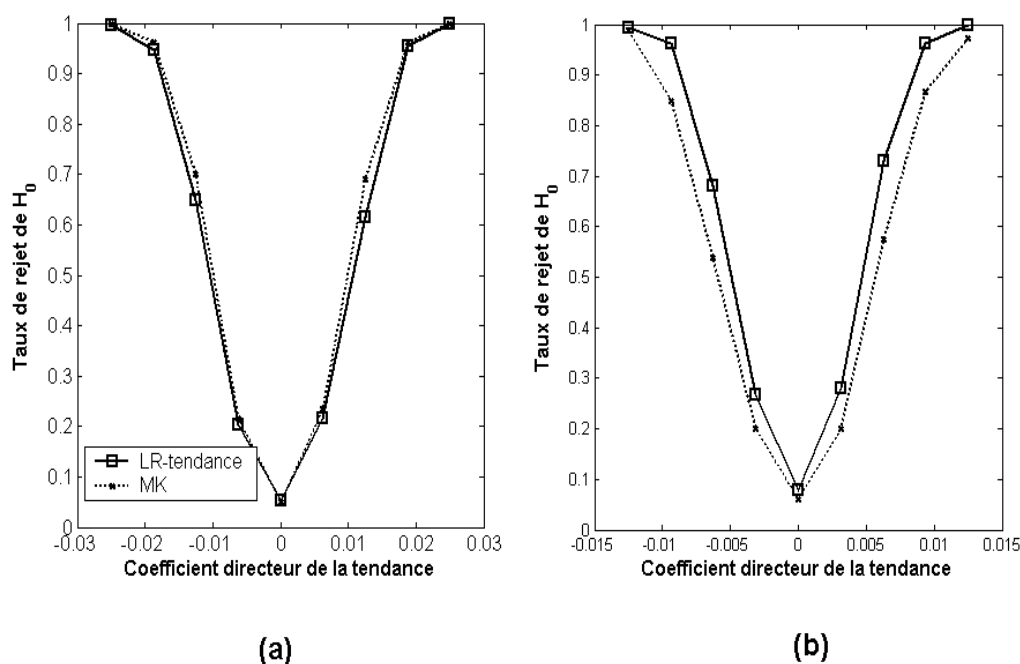


Figure 3-17. Puissance des tests de Mann-Kendall et LR-tendance ($\alpha=0.05$), (a) données Gaussiennes, (b) données GEV

I.3.2. Recherche des tests les plus puissants

Tests non paramétriques

Yue *et al.* (2002a; 2002b) ont étudié le comportement des deux tests non paramétriques les plus couramment utilisés en hydrologie : le test de Mann-Kendall et le test de Spearman. Il s'avère que ces deux tests sont pratiquement indiscernables en terme de puissance, et ce quelles que soient les conditions d'utilisation (taille de l'échantillon, coefficient de variation des données, distribution...). Lubès-Niel *et al.* (1998) observent quant à eux une légère supériorité du test de Mann-Kendall sur celui de Pettitt pour la puissance, sous hypothèse de rupture avec des données Gaussiennes. Les moins bonnes performances du test de Pettitt peuvent en partie s'expliquer par le fait que ce test est un peu trop conservatif pour de petits échantillons (cf. I.2.1.).

La Figure 3-18 résume les performances de quatre tests non paramétriques, avec deux types de distributions et deux types de changements (1000 simulations, $\alpha=0.05$, $N=80$). On y retrouve la similarité des tests de Mann-Kendall et de Spearman, qui semblent quasiment équivalents. Le test de McGilchrist et Woodyer semble moins puissant, tandis que le test de Pettitt n'est pas très efficace pour détecter une tendance, mais l'est d'avantage pour détecter une rupture. On retrouve d'ailleurs ici une conclusion qu'avait formulée Lemaitre (2002) : les tests de détection de tendance (MK, Spearman) sont aussi puissants que les tests de détection de rupture (Pettitt) pour détecter une rupture centrale, mais l'inverse n'est pas vrai.

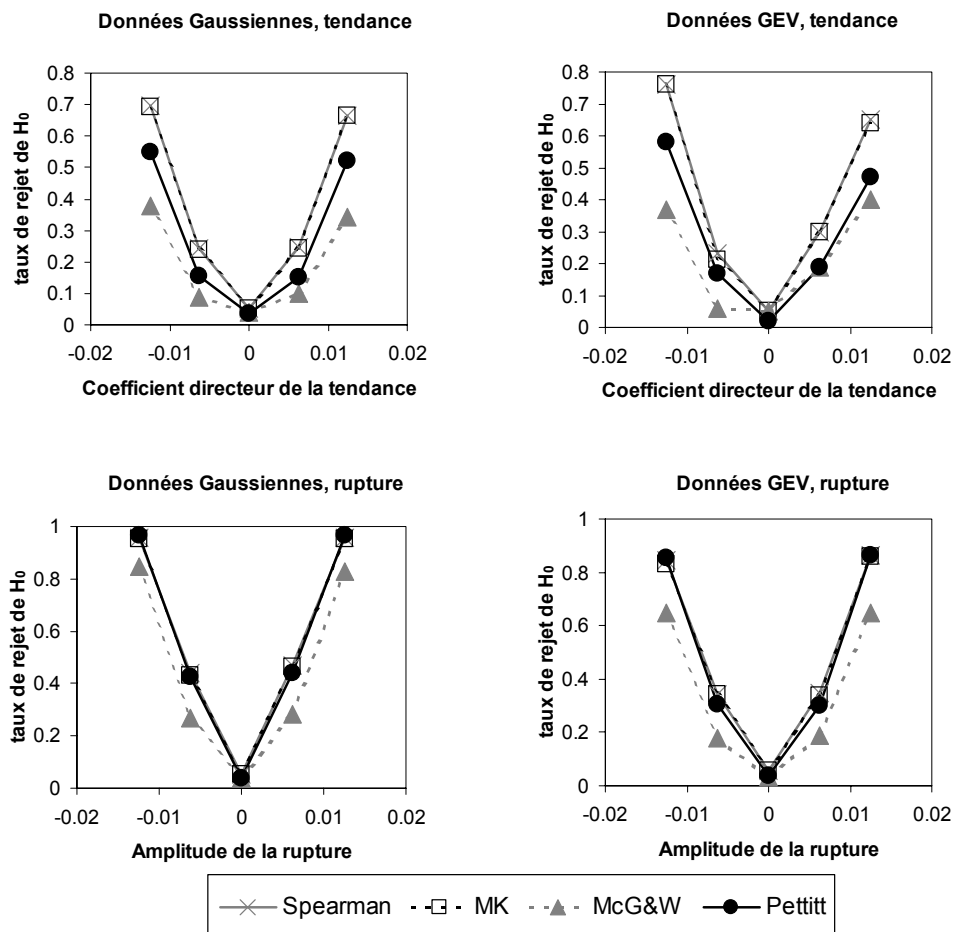


Figure 3-18. Puissance comparée de quatre tests non paramétriques ($\alpha=0.05$)

Tests paramétriques

Nous nous intéresserons seulement à la comparaison de tests paramétriques dans le cas de données Gaussiennes ou exponentielles à un paramètre. Pour d'autres types de distribution, le choix de tests est en effet très restreint, et l'utilisation de tests LR semble s'imposer.

La Figure 3-19 résume la performance des tests dans le cas de données Gaussiennes, en considérant deux types de changement (1000 simulations, $N=80$, $\alpha=0.05$). Dans le cas d'une tendance, les tests de régression linéaire et de Gardner sont les plus puissants. La perte de puissance liée à l'estimation d'un paramètre supplémentaire dans le cas d'un test LR-tendance n'est pas très importante (sept points au maximum), par contre le test LR-rupture est nettement moins performant. Dans le cas d'une rupture centrale, le test de Buishand semble être le plus puissant, mais la différence avec les tests de détection de tendance (régression et Gardner) est très faible. On retrouve encore une fois la conclusion de Lemaitre (2002) quant à la puissance des tests de tendance pour détecter des ruptures centrales. Étonnamment, dans le cas des tests LR, le test de tendance est même plus puissant que le test de rupture. L'estimation du paramètre supplémentaire est ici un peu plus coûteuse, puisque la perte de puissance maximale par rapport au test de Buishand atteint une quinzaine de points.

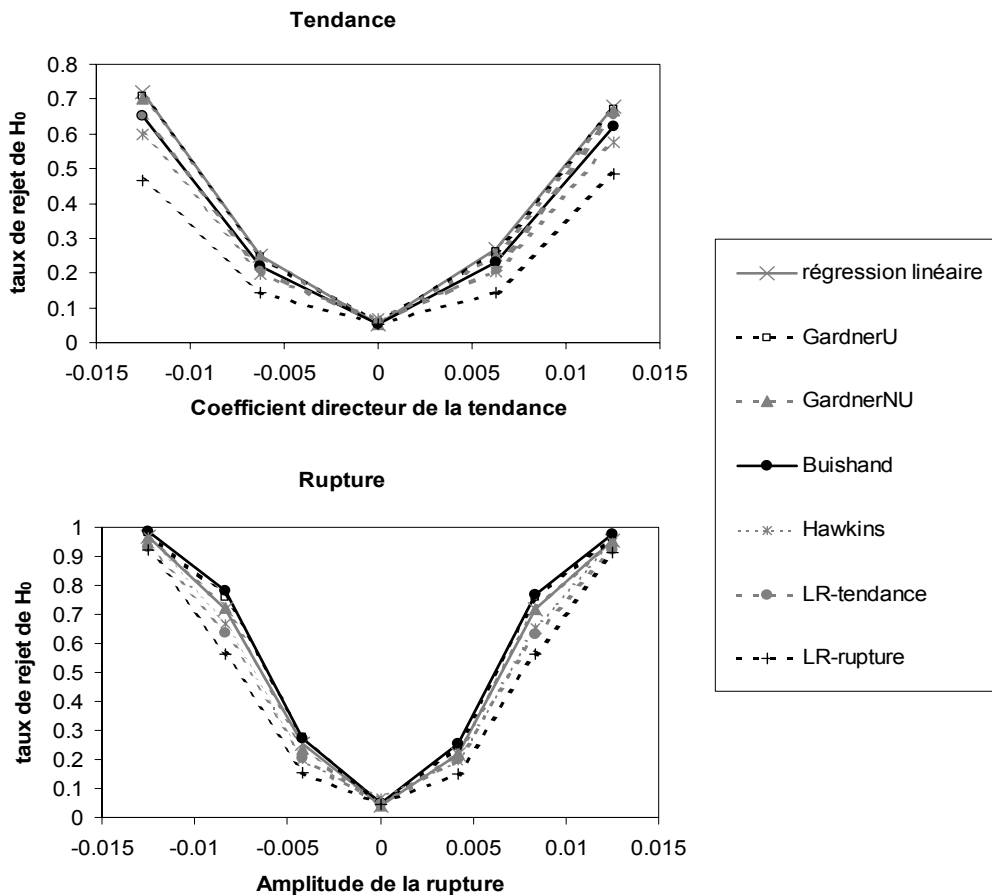


Figure 3-19. Puissance comparée de quelques tests paramétriques sur des données Gaussiennes ($\alpha=0.05$)

Il est parfois intéressant de rechercher un changement ayant affecté la fréquence d'événements extrêmes plutôt que leur intensité, ce qui conduit à s'intéresser au processus d'occurrence. Une hypothèse classique est que le temps d'attente entre deux événements suit une loi exponentielle. La stationnarité de cette loi peut alors être évaluée par les tests de Keim et Cruise (1998), par le test de Lang *et al.* (2004), par un test LR basé sur la loi exponentielle ou par tout test non paramétrique (nous choisirons Mann-Kendall). La Figure 3-20 présente

les résultats de 1000 simulations ($N=80$, $\alpha=0.05$, tendance et rupture centrale). Dans le cas d'une tendance, le test LR-tendance est largement le plus puissant. Dans le cas d'une rupture, il est rejoint par les tests de Lang *et al.* et LR-rupture.

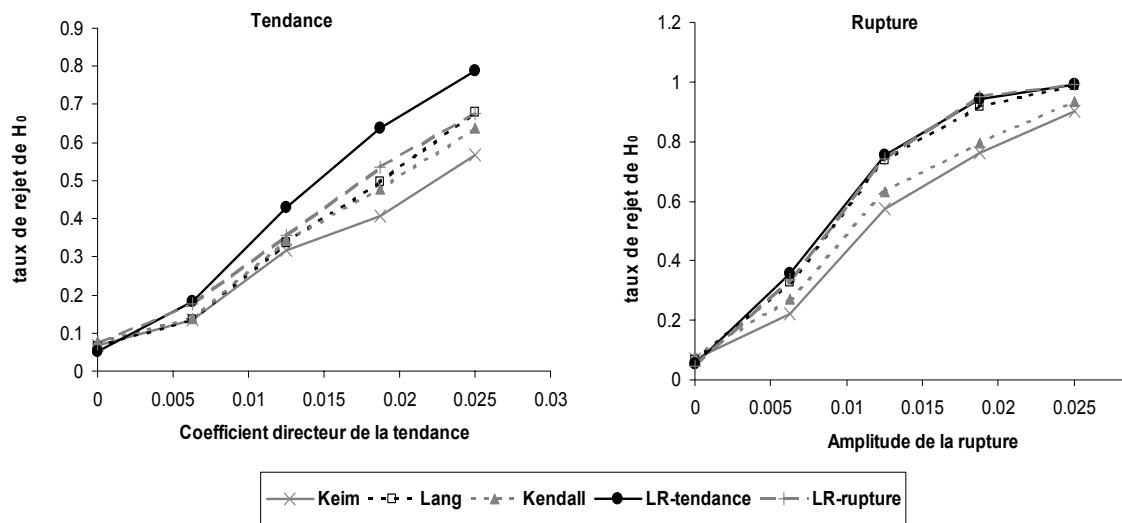


Figure 3-20. Puissance comparée des tests portant sur le processus d'occurrence d'événements ($\alpha=0.05$)

I.3.3. Influence de la position d'une rupture

Il peut paraître surprenant que les tests de détection de rupture ne soient pas plus puissants que les tests de tendance dans le cas d'un changement brusque. Nous avons souhaité préciser cette observation en étudiant sa validité dans le cas de ruptures non centrales, qui sont *a priori* plus difficiles à détecter.

Sen et Srivastava (1975) ont déjà effectué ce type d'étude, mais avec des tests qui supposent la variance des données connue. Plus récemment, Dindar (2003) a démontré que les tests de Hawkins et Buishand (qui sont des tests paramétriques de rupture) sont plus puissants pour détecter des changements situés aux extrémités de la chronique. La Figure 3-21 confirme cette observation. Elle a été réalisée en simulant un doublement du paramètre de la loi exponentielle à une date variable (1000 simulations, $N=80$, $\alpha=0.05$). Il apparaît que le test de rupture devient plus puissant pour les ruptures situées proche des extrémités de la série.

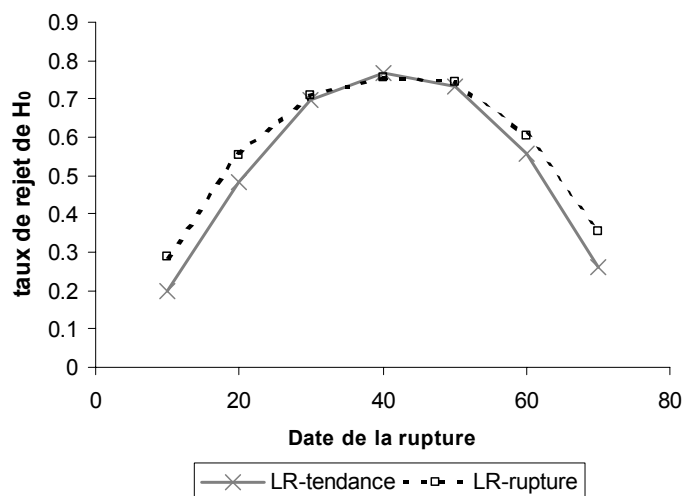


Figure 3-21. Influence de la date de rupture sur la puissance des tests ($\alpha=0.05$)

Dans la pratique, si l'on s'attend à un changement brusque à une date inconnue, les tests de rupture ne sont donc pas dénués d'intérêt. Ce qui nous intéresse en effet, plus que la puissance à date de rupture fixée, est la probabilité de parvenir à détecter le changement, quel que soit l'endroit où il se situe. Afin de nous mettre dans ce type de situation, nous avons simulé une rupture à date aléatoire, uniforme sur $1, \dots, N-1$. (données exponentielles, 1000 simulations, $N=80$, $\alpha=0.05$). Il apparaît alors que, pour ce critère de puissance modifiée, le test de rupture est plus performant que le test de tendance (Figure 3-22).

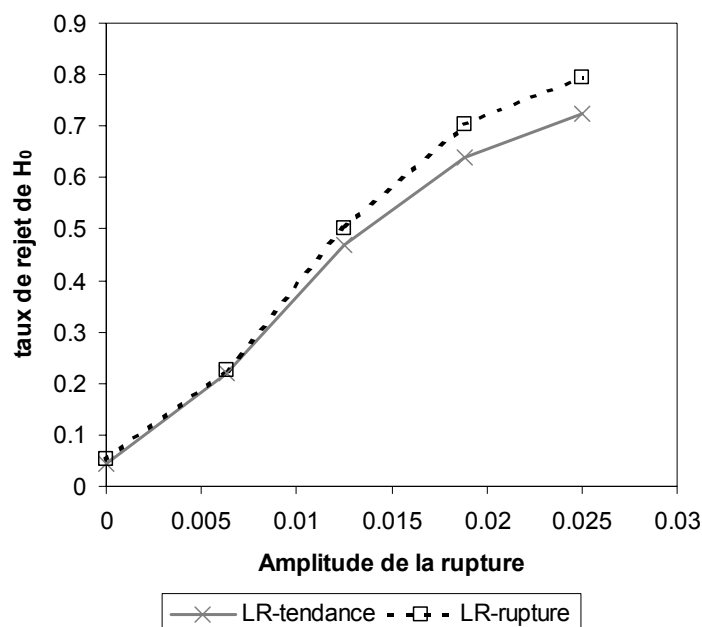


Figure 3-22. Taux de rejet de H_0 avec une date de rupture aléatoire ($\alpha=0.05$)

I.3.4. Le cas des données autocorrélées

Nous souhaitons à présent évaluer la performance des tests prenant en compte l'autocorrélation. Nous avons pour cela simulé des processus AR(1) Gaussiens perturbés par une tendance, auxquels nous avons appliqué les tests de Mann-Kendall modifié (MMK) et de régression linéaire modifié (W&G), avec l'hypothèse AR(1). Nous avons également appliqué ces tests non modifiés après transformation PW. Les résultats sont présentés dans la Figure 3-23 (1000 simulations, $N=80$, $\alpha=0.05$). On note tout d'abord l'inévitable perte de puissance qui accompagne la présence d'une forte autocorrélation. Pour des autocorrélations modérées ($\rho < 0.4$), le test de régression linéaire modifié est le plus puissant et reste correctement calé. Il devient par contre inutilisable pour des autocorrélations plus importantes (avec cette taille d'échantillon tout du moins), puisque le taux de rejet est beaucoup plus important que le risque spécifié. Dans ces conditions, le test de Mann-Kendall modifié est à recommander, puisqu'il semble beaucoup plus puissant que l'application d'un test après PW, tout en gardant sous H_0 un taux de rejet convenable.

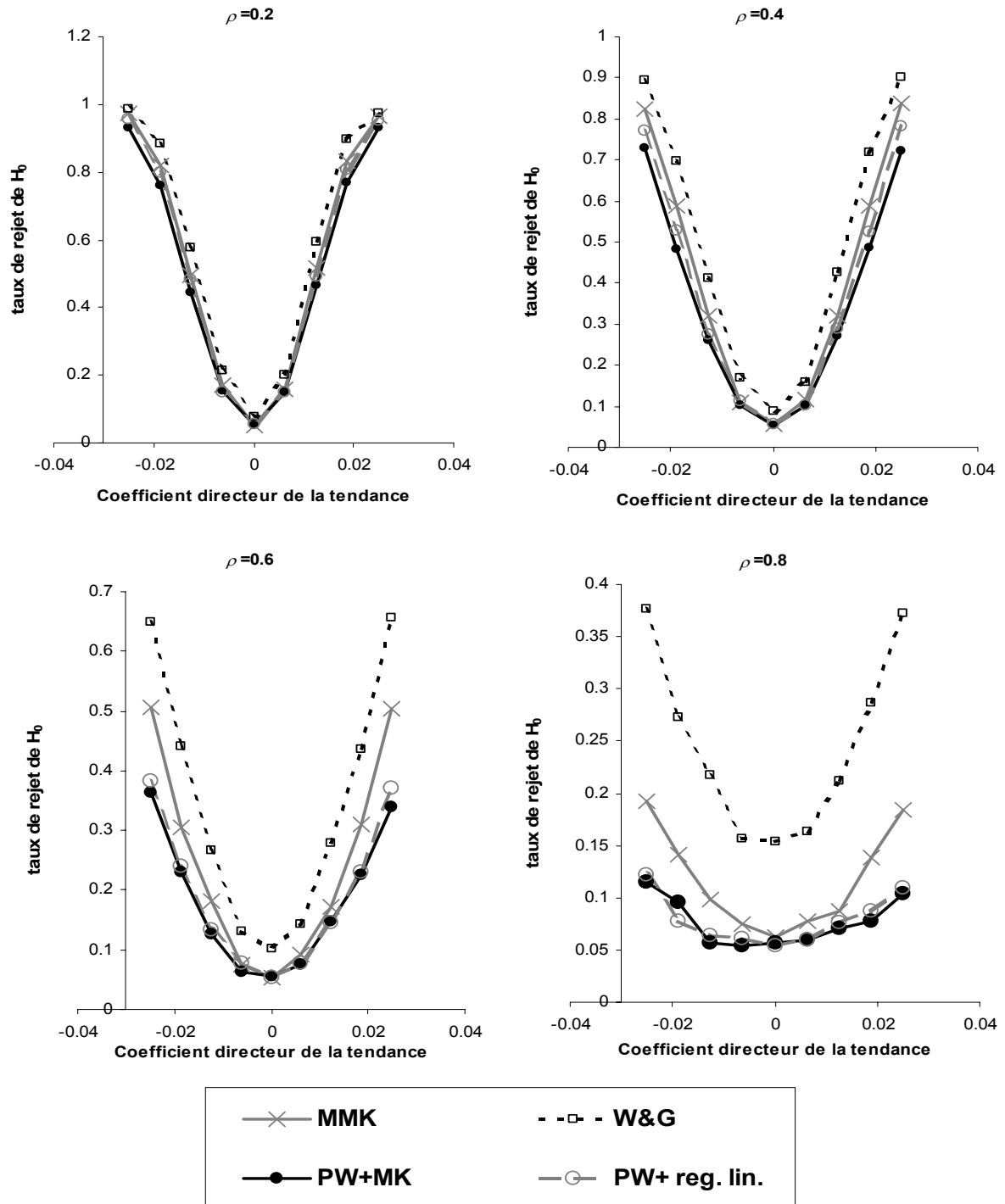


Figure 3-23. Puissance des tests prenant en compte l'autocorrélation ($\alpha=0.05$)

I.4. Conclusion sur les méthodes locales de détection

Cette étude nous a permis de dresser un état des lieux des méthodes disponibles pour la détection de changements, ainsi que des connaissances quant à leurs performances en terme de robustesse et de puissance. Parmi les pratiques les plus courantes dans le domaine de la détection de changements dans des séries hydro-météorologiques, un certain nombre de modifications devraient permettre d'améliorer l'efficacité de telles études. Tout d'abord, nous reprenons les conclusions de Zhang *et al.* (2004) : un net gain de puissance peut être espéré

en utilisant des tests paramétriques appropriés, particulièrement dans le domaine des valeurs extrêmes. Les tests LR devraient donc être plus utilisés, après vérification des hypothèses de distribution des données. Un autre avantage de ces tests est qu'ils permettent de détecter des changements en moyenne et en variance. Les changements en variance n'ont que rarement été évalués, alors que cette hypothèse est loin d'être incongrue physiquement (IPCC, 2001). Concernant les données autocorrélées, le retrait d'une tendance avant l'application d'un test semble être une opération dangereuse vis-à-vis du risque de première espèce. Par contre, dans le cas où une tendance significative a été détectée au préalable, l'estimation de la fonction de corrélation peut être effectuée après retrait de cette tendance afin de diminuer le biais. Enfin, la transformation PW, bien que correcte du point de vue de la conservation du risque, est bien moins puissante que le test de Mann-Kendall modifié proposé par Hamed et Rao (1998).

Sur la base de ces conclusions, nous proposons un schéma de sélection du test statistique le plus approprié à une situation donnée (Figure 3-24). Appliqué à l'étude que nous souhaitons mener sur la stationnarité des régimes hydrologiques en France, ce schéma de sélection permet de choisir un test pour chaque variable hydrologique à tester.

Pour les variables relatives aux crues, nous supposons que les valeurs de l'échantillon sont indépendantes. Les pointes de crues, obtenues par échantillonnage MAXAN ou SUPSEUIL, seront testées par les tests LR construits respectivement avec les lois GEV et GPD. De même, les variables de dénombrement d'événements seront supposées Poissoniennes, bien que cette hypothèse soit certainement abusive dans le cas de régimes à forte composante nivale ou glaciaire. En effet, pour de tels cours d'eau, la régularité du régime fait qu'il n'y a, la plupart des années, qu'une unique crue. Le test sera alors trop conservatif (Figure 3-8), ce qui n'est pas un inconvénient majeur puisqu'il semble de toute façon difficile de détecter un changement sur des séries présentant si peu de variabilité. Les variables relatives aux formes des hydrogrammes peuvent quant à elles avoir des distributions très variées, voire multimodales malgré les précautions que nous avons prises lors de l'échantillonnage (détection de mélange de populations). Il semble donc indispensable d'utiliser un test non paramétrique.

Concernant les variables relatives aux étiages ou aux variables moyennes annuelles, la principale question concerne la présence éventuelle d'autocorrélation. Les rivières contrôlées par une nappe importante ont plus de chance de présenter une persistance significative sur ces variables, à cause de l'inertie du mécanisme de remplissage/vidange de nappe. Faut-il alors tester toutes ces variables avec un test de Mann-Kendall modifié, au risque de diminuer fortement la puissance ? Il semble préférable de tester au préalable la significativité des autocorrélations. En effet, bien que la présence d'une tendance biaise l'estimation de la fonction d'autocorrélation, ce biais est toujours positif, c'est à dire que la tendance crée de l'autocorrélation. Nous ne risquons donc pas d'ignorer à tort la persistance, et nous préserverons ainsi le risque de première espèce. Concernant les distributions des différentes variables, seules les minima annuels (GEV) seront testés paramétriquement, la loi des autres variables étant assez hypothétique.

Notons enfin que nous appliquerons pour chaque variable un test de tendance et un test de rupture. En effet, dans l'optique d'un changement climatique, on s'attend plutôt à une variation graduelle. Cependant, les tests de rupture pourront nous être utiles pour détecter et dater des changements brusques dont l'origine n'a rien de climatique, comme les problèmes de métrologie (modification de la section de mesure, déplacement de l'échelle) ou des influences anthropiques directes (aménagement).

Les tests que nous utiliserons sont résumés dans le Tableau 3-5.

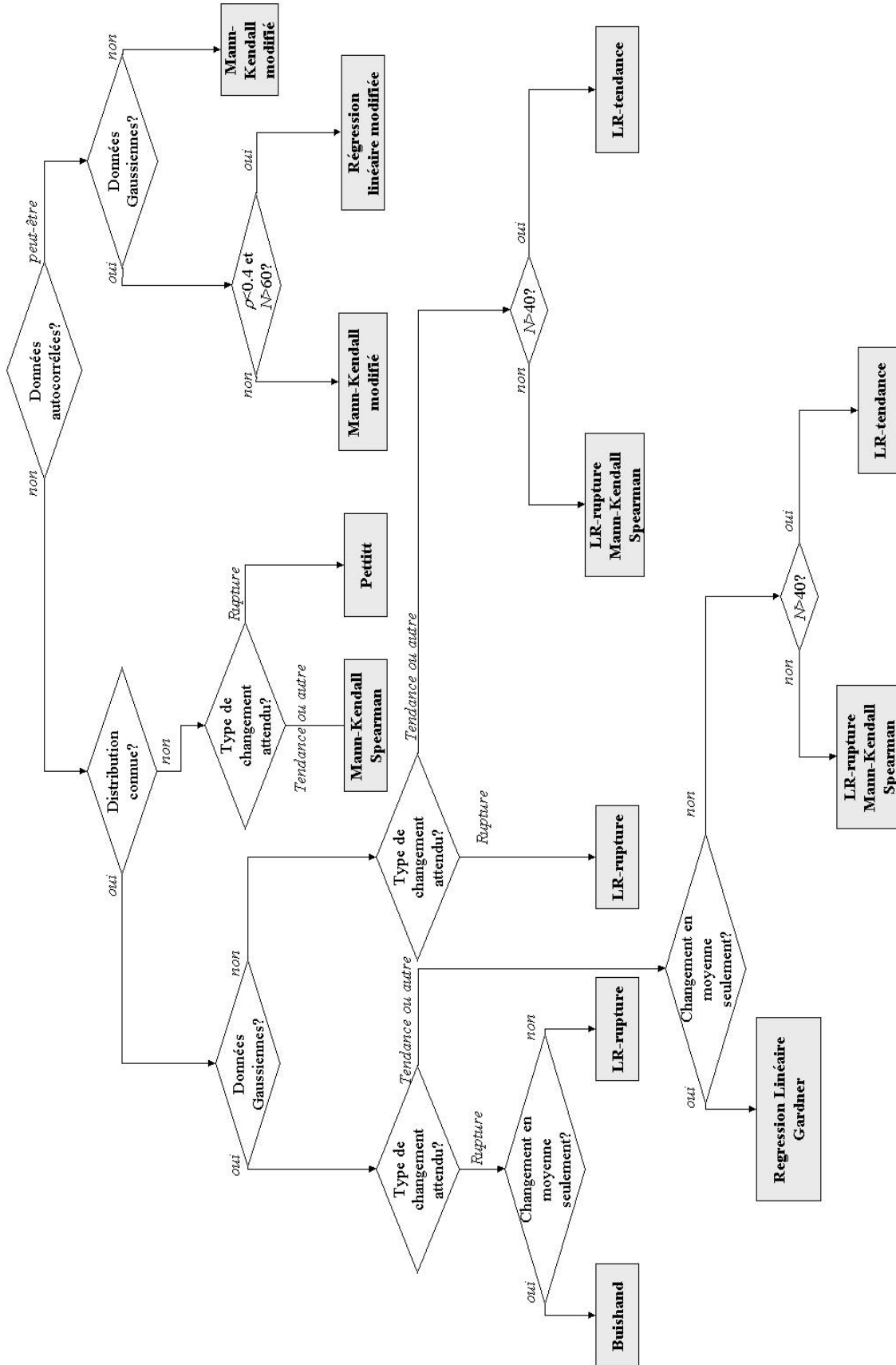


Figure 3-24. Schéma de sélection d'un test de stationnarité

Variable	Unité	Description	Echantillonnage	Distribution	Auto-corrélation	Test(s)
<i>Crues</i>						
q_{max}	$m^3.s^{-1}$	Valeur du débit Maximum Annuel	MAXAN	GEV	non	LR-tendance+LR-rupture
δ	jours	durée de l'événement (dépassement de la moitié du pic)	Sélection d'hydrogrammes	Inconnue	non	MK+Pettitt
γ	/	coefficient d'asymétrie	Sélection d'hydrogrammes	Inconnue	non	MK+Pettitt
m	jours	temps de montée entre la moitié du pic et le pic	Sélection d'hydrogrammes	Inconnue	non	MK+Pettitt
r	jours	temps de descente entre le pic et la moitié du pic	Sélection d'hydrogrammes	Inconnue	non	MK+Pettitt
σ	/	coefficient d'écart-type	Sélection d'hydrogrammes	Inconnue	non	MK+Pettitt
θ	/	Date angulaire de l'événement	Sélection d'hydrogrammes	Inconnue	non	MK+Pettitt
v	$10^6 m^3$	Volume écoulé pendant l'événement	Sélection d'hydrogrammes	Inconnue	non	MK+Pettitt
q_s	$m^3.s^{-1}$	Débit de pointe sup-seuil	SUPSEUIL	GPD	non	LR-tendance+LR-rupture
N_μ	/	Nombre d'événements sur l'année, sélectionnés avec en moyenne μ événements par an	variable annuelle	Poisson	non	LR-tendance+LR-rupture
<i>Etiages</i>						
q_{min}	$m^3.s^{-1}$	Valeur du débit Minimum Annuel	MINAN	GEV	A tester	LR-tendance+LR-rupture ou MMK
θ	/	Date angulaire de q_{min}	variable annuelle	Inconnue	A tester	MK+Pettitt ou MMK
δ	jours	Durée cumulée sur l'année au dessous du seuil	variable annuelle	Inconnue	A tester	MK+Pettitt ou MMK
v	jours	Déficit de volume par rapport au seuil, cumulé sur l'année	variable annuelle	Inconnue	A tester	MK+Pettitt ou MMK
θ_1	/	date angulaire du premier passage sous le seuil de l'année	variable annuelle	Inconnue	A tester	MK+Pettitt ou MMK
θ_2	/	date angulaire du dernier jour sous le seuil de l'année	variable annuelle	Inconnue	A tester	MK+Pettitt ou MMK
<i>Régime des eaux</i>						
$module$	$m^3.s^{-1}$	module (annuel ou saisonnier)	Moyenne annuelle	Inconnue	A tester	MK+Pettitt ou MMK
$dist$	/	Distance au régime inter-annuel	variable annuelle	Inconnue	A tester	MK+Pettitt ou MMK

Tableau 3-5. Récapitulatif des tests appliqués aux variables hydrologiques

II. Significativité régionale

Supposons que nous ayons réalisé un test statistique au risque $\alpha=0.05$ sur 100 stations. Supposons également que toutes ces stations soient stationnaires. On s'attend tout de même, à cause du risque d'erreur choisi, à détecter à tort environ 5 changements significatifs. Mais que dire si l'on en détecte 6 ? 10 ? 20 ? A partir de combien de résultats significatifs peut-on affirmer, avec un certain risque d'erreur α' , que tous ces changements ne peuvent être dus au hasard ? Il apparaît clairement qu'une nouvelle démarche de test statistique est nécessaire à l'échelle régionale, pour s'assurer que le nombre de résultats significatifs localement est suffisant. Nous parlerons de significativité régionale des changements observés. Insistons sur le fait qu'il ne s'agit pas ici d'évaluer la cohérence régionale de changements, mais bel et bien de vérifier que les résultats locaux sont significatifs. En particulier, des tendances en sens contraires peuvent très bien être observées sur deux stations proches (au sens hydro-climatique), nous ne nous intéresserons ici qu'au nombre de résultats significatifs.

Ce phénomène est connu et a déjà été étudié dans le domaine hydro-météorologique (Livezey et Chen, 1983; Lettenmaier *et al.*, 1994; Douglas *et al.*, 2000; Burn et Elnur, 2002; Yue et Wang, 2002b; Yue *et al.*, 2003). Cependant, les différentes méthodes proposées par ces auteurs pour prendre en compte cet aspect régional n'ont été que rarement utilisées dans les études de stationnarité. Les prochaines parties seront consacrées à l'étude de ces méthodes.

II.1. Quelques considérations théoriques

Nous nous intéressons donc à la loi de la variable aléatoire N , égale au nombre de changements significatifs localement, sous l'hypothèse que toutes les séries sont stationnaires. La matrice des données sera notée \mathbf{X} (n années en lignes, p stations en colonnes), les données de la station j seront notées $\mathbf{X}^{(j)}$. Nous supposons pour simplifier les notations que toutes les séries sont testées avec une même procédure, dont la significativité est évaluée en comparant la statistique de test $S_j=g(\mathbf{X}^{(j)})$ avec une valeur critique c . Notons enfin Ω_k l'ensemble des combinaisons de k éléments parmi p , qui représente l'ensemble des possibilités d'observer k stations significatives parmi les p stations de l'étude. On peut alors écrire :

$$\begin{aligned} \Pr(N=k) &= \sum_{\Omega_k} \Pr\left(S_{j_1} > c, \dots, S_{j_k} > c, S_{j_{k+1}} \leq c, \dots, S_{j_p} \leq c\right) \\ &= \sum_{\Omega_k} \Pr\left(g(\mathbf{X}^{(j_1)}) > c, \dots, g(\mathbf{X}^{(j_k)}) > c, g(\mathbf{X}^{(j_{k+1})}) \leq c, \dots, g(\mathbf{X}^{(j_p)}) \leq c\right) \\ &= \sum_{\Omega_k} \Pr\left(\mathbf{X}^{(j_1)} \in g^{-1}(]c; +\infty[), \dots, \mathbf{X}^{(j_k)} \in g^{-1}(]c; +\infty[), \mathbf{X}^{(j_{k+1})} \in g^{-1}(]-\infty; c]), \dots, \mathbf{X}^{(j_p)} \in g^{-1}(]-\infty; c])\right) \\ &= \sum_{\Omega_k} \Pr\left(\mathbf{X} \in \mathfrak{R}_{\Omega_k}\right) \end{aligned} \quad (3-32)$$

\mathfrak{R}_{Ω_k} représentant la partie de $\mathbf{R}^{n \times p}$ définie par :

$$\left\{ \mathbf{X} = (\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p)}) \mid \mathbf{X}^{(j_1)} \in g^{-1}(]c; +\infty[), \dots, \mathbf{X}^{(j_k)} \in g^{-1}(]c; +\infty[), \mathbf{X}^{(j_{k+1})} \in g^{-1}(]-\infty; c]), \dots, \mathbf{X}^{(j_p)} \in g^{-1}(]-\infty; c]) \right\} \quad (3-33)$$

Cette expression montre que la loi multivariée relative à la matrice de données \mathbf{X} est nécessaire à l'obtention de la loi de la variable N . Malheureusement, il semble difficile d'accéder à cette loi, qui doit intégrer les dépendances spatiales entre les stations. Dans le cas de données hydrologiques, cette corrélation spatiale est d'autant plus problématique qu'elle est la résultante de la dépendance spatiale du processus de génération des écoulements (en général la pluie), qui peut être considérée comme à peu près continue dans l'espace, et de la structuration du réseau hydrographique. Il apparaît donc que des approximations seront nécessaires pour tenter de prendre en compte la significativité régionale.

Ainsi, si l'on suppose que les stations sont indépendantes, la loi de N est aisément accessible :

$$\begin{aligned}
 \Pr(N = k) &= \sum_{\Omega_k} \Pr(S_{j_1} > c, \dots, S_{j_k} > c, S_{j_{k+1}} \leq c, \dots, S_{j_p} \leq c) \\
 &= \sum_{\Omega_k} \Pr(S_{j_1} > c) \dots \Pr(S_{j_k} > c) \Pr(S_{j_{k+1}} \leq c) \dots \Pr(S_{j_p} \leq c) \\
 &= \sum_{\Omega_k} \alpha^k (1 - \alpha)^{p-k} \\
 &= C_p^k \alpha^k (1 - \alpha)^{p-k}
 \end{aligned} \tag{3-34}$$

On reconnaît une loi binomiale, ce qui semble naturel car sous l'hypothèse d'indépendance, les p tests sur les stations peuvent être considérés comme p expériences aléatoires indépendantes, ayant chacune une probabilité de succès égale à α . Cette approximation binomiale permet de donner un premier ordre de grandeur du nombre minimum de stations localement significatives pour garantir la significativité régionale (Livezey et Chen, 1983). A titre d'exemple, avec 100 stations testées au risque $\alpha=0.05$, il faut au moins 10 stations significatives pour garantir la significativité régionale au risque $\alpha'=0.05$.

II.2. Les différentes méthodes disponibles

II.2.1. Utilisation du nombre de stations équivalent

De manière analogue au cas de données autocorrélées, une première approche consiste à utiliser la notion d'effectif efficace. En d'autres termes, les p stations corrélées vont être considérées comme p^* stations indépendantes. Il sera alors possible de calculer un pourcentage critique par la loi binomiale. Notons que cette approche est en grande partie empirique, car d'une part la valeur de p^* n'est pas définie de manière unique (elle est relative à une statistique), et d'autre part rien ne garantit théoriquement que le pourcentage critique calculé sur les données pseudo-indépendantes doit rester égal au pourcentage critique sur toutes les stations.

Douglas *et al.* (2000) ainsi que Yue et Wang (2002b) ont étudié l'augmentation de variance causée par la présence de corrélation sur une statistique particulière (statistique de Mann-Kendall moyennée sur plusieurs stations). Les auteurs ont mis en évidence la relation suivante :

$$\frac{Var_{\rho}(S)}{Var_{\rho=0}(S)} = \eta = 1 + (p-1)\bar{\rho}, \text{ où } \bar{\rho} = \frac{2 \sum_{k=1}^{p-1} \sum_{l=1}^{p-k} \rho_{k,k+l}}{p(p-1)} \tag{3-35}$$

De même que dans le cas de données autocorrélées, nous supposons que ce facteur correctif η est égal au ratio p/p^* .

Reprenons à titre d'exemple le cas de 100 stations testées au risque $\alpha=0.05$. Supposons que la corrélation moyenne entre les stations vaille 0.1. Le facteur correctif est alors donnée par $\eta=1+(p-1)\bar{\rho}$, ce qui donne un effectif de 9.17 stations pseudo-indépendantes. La loi binomiale fournit alors une valeur critique de 2.37 stations, soit un pourcentage critique de 25.8%, ce qui représente un minimum de 26 stations localement significatives pour obtenir la significativité régionale.

II.2.2. Le Bootstrap par année

Douglas *et al.* (2000) proposent d'évaluer la significativité régionale en utilisant une procédure de rééchantillonnage. En effet, la corrélation entre stations est invariante par permutation des lignes (i.e. des années) dans la matrice de données. Sous l'hypothèse que toutes les stations sont stationnaires, la création de nouveaux jeux de données par rééchantillonnage des années permet donc d'accéder à la distribution d'échantillonnage du nombre de stations localement significatives. Douglas *et al.* proposent en fait d'échantillonner les années par Bootstrap plutôt que par permutation, ce qui conduit à la procédure suivante :

Pour $m=1, \dots, n_{sim}$:

- ✓ Création d'un échantillon de n années par Bootstrap
- ✓ Pour chaque année, on considère les valeurs observées sur les p stations
- ✓ Application d'un test à chacune des p stations
- ✓ Calcul de $N^{(m)}$, le nombre de tests significatifs

L'échantillon ($N^{(m)}$) peut alors être utilisé pour calculer une valeur critique.

II.2.3. Utilisation d'une copule Gaussienne

Afin de modéliser la loi multivariée des observations, il est possible d'utiliser une copule Gaussienne. Par souci de clarté, cet outil statistique ne sera pas développé dans ce paragraphe. Une description détaillée est par contre fournie dans l'article figurant en Annexe 3 (Renard et Lang, 2006). Pour résumer, ce modèle suppose qu'après une transformation adéquate des données marginales, la loi normale multivariée peut modéliser les observations. Cet outil présente l'avantage d'être assez simple à utiliser. Il convient par contre de vérifier que le modèle Gaussien est adéquat pour décrire les dépendances entre stations. Une fois les paramètres de la copule estimés, il ne sera cependant pas directement possible d'accéder à la loi de la variable N , car la détermination des régions \mathfrak{R}_{Ω_k} (équation (3-33)) ne peut être effectuée de manière explicite. Il sera donc nécessaire d'avoir recours à la simulation, ce qui conduit à la procédure suivante :

- ✓ Estimation des paramètres de la copule sur les données
- ✓ Pour $m=1, \dots, n_{sim}$:
 - Simulation de la copule estimée
 - Application d'un test à chacune des p stations
 - Calcul de $N^{(m)}$, le nombre de tests significatifs

Encore une fois, le calcul d'une valeur critique est effectué avec l'échantillon ($N^{(m)}$). Cette approche ne peut être envisagée que pour un nombre modéré de sites : en effet, l'estimation d'une matrice de variance non singulière peut être impossible si le nombre d'observations est insuffisant par rapport au nombre de stations.

II.3. Etude par simulation

La simulation de données spatialisées est en soi un problème complexe. Dans le cas de données hydrométriques, et notamment de données extrêmes, nous ne pouvons pas espérer disposer d'un moyen de simulation à la fois simple et réaliste. Nous allons donc mener cette étude par différentes procédures de simulation, qui ne reflèteront qu'une partie des différentes structures de dépendance spatiale que l'on peut rencontrer avec des données réelles. Pour toutes ces simulations, nous choisirons des risques locaux et régionaux égaux à 0.1, et le test de Mann-Kendall sera appliqué localement.

II.3.1. Simulation à partir d'une loi normale multivariée

Les données X_θ seront issues d'une loi Gaussienne de dimension 50, centrée réduite, et de matrice de corrélation Σ_ρ définie par :

$$\Sigma_\rho = \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{p-1} \\ \rho & 1 & \rho & \ddots & \vdots \\ \rho^2 & \rho & \ddots & \ddots & \rho^2 \\ \vdots & \ddots & \ddots & 1 & \rho \\ \rho^{p-1} & \dots & \rho^2 & \rho & 1 \end{pmatrix} \quad (3-36)$$

Nous simulerons 80 années de données pour chaque variable. Il est possible d'estimer la distribution de N , en re-simulant un grand nombre de fois des données issues de cette loi Gaussienne, et en dénombrant à chaque simulation le nombre de tests localement significatifs. La Figure 3-25 illustre la déformation de cette distribution lorsque le paramètre ρ de la matrice de corrélation varie. Elle a été réalisée avec 1000 simulations, et en choisissant en risque local $\alpha=0.1$. Le trait vertical représente le quantile de probabilité 0.9, c'est à dire la valeur critique pour assurer la significativité régionale au risque 10%. Dans la suite de cette section, nous parlerons de la valeur critique « théorique », bien qu'il ne s'agisse que d'une estimation obtenue à partir d'un grand nombre de simulations. Remarquons enfin que la moyenne de ces distribution reste égale à 5 (sur 50 stations, soit 10%), conformément au risque de première espèce local.

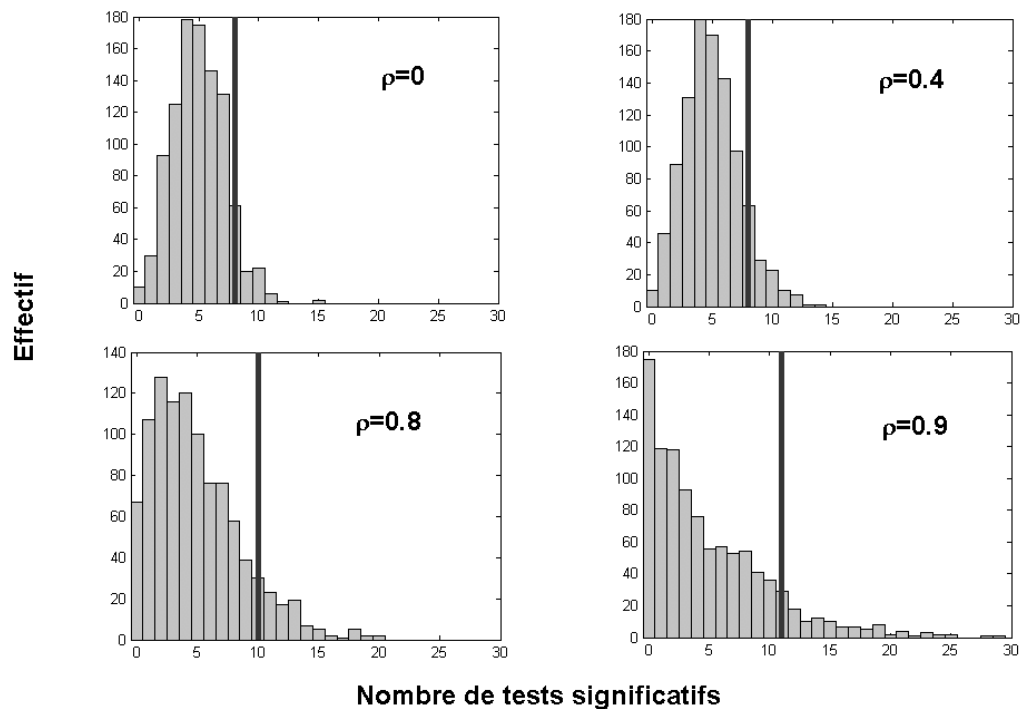


Figure 3-25. Distribution simulée de N en fonction de la corrélation des données (loi Gaussienne multivariée).

A partir de la matrice simulée X_θ , nous avons appliqué les méthodes décrites précédemment. Les résultats sont présentés dans le Tableau 3-6. Il apparaît que les méthodes par Bootstrap ou par copule donnent une bonne approximation de la valeur critique théorique de N . L'approximation binomiale, qui suppose l'indépendance des données, reste valable pour des corrélations modérées. Par contre, l'utilisation du nombre équivalent de stations conduit à

des valeurs critiques beaucoup trop élevées, ce qui tend à démontrer que la formule de Douglas *et al.* obtenue sur la statistique de Mann-Kendall moyenne ne peut pas être transférée trivialement à une autre statistique.

Corrélation ρ	0	0.2	0.4	0.6	0.8	0.9
Valeur « théorique »	8	8	8	9	10	11
Approximation binomiale	9	9	9	9	9	9
Bootstrap-années	8	9	9	9	11	12
Effectif efficace	9	10	11	13	18	25
Copule Gaussienne	8	8	8	9	11	12

Tableau 3-6. Estimation des valeurs critiques de N (loi Gaussienne multivariée)

La procédure de simulation que nous venons d'utiliser a le mérite de la simplicité, mais elle n'est clairement pas représentative des structures de dépendance que nous rencontrerons dans la réalité (la matrice Σ_ρ étant beaucoup trop simpliste).

II.3.2. Simulation par une copule Gaussienne

Nous allons utiliser des paramètres estimés sur des données réelles afin de rendre nos simulations plus réalistes. Sur 13 stations hydrométriques situées dans le Nord-Est de la France, les paramètres suivants ont été estimés : paramètres de position et d'échelle des lois de Gumbel marginales (Tableau 3-7) et matrice de corrélation des observations (Tableau 3-8)

Paramètre	Station												
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13
Position	34.96	196.8	265.8	111.2	427.4	226.1	120.3	143	122.9	134.6	104.8	165.7	186.3
Echelle	20.89	67.43	81.24	39.49	146.8	101.5	54.05	67.58	45.08	41.4	44.12	57.29	50.11

Tableau 3-7. Paramètres des lois marginales

$$\Sigma = \begin{pmatrix} 1 & 0.574 & 0.63 & 0.663 & 0.75 & 0.886 & 0.69 & 0.409 & 0.481 & 0.517 & 0.431 & 0.424 & 0.388 \\ 0.574 & 1 & 0.976 & 0.608 & 0.778 & 0.635 & 0.217 & 0.2 & 0.539 & 0.245 & 0.218 & 0.609 & 0.621 \\ 0.63 & 0.976 & 1 & 0.66 & 0.839 & 0.713 & 0.273 & 0.207 & 0.58 & 0.276 & 0.319 & 0.631 & 0.646 \\ 0.663 & 0.608 & 0.66 & 1 & 0.834 & 0.754 & 0.449 & 0.135 & 0.597 & 0.393 & 0.45 & 0.335 & 0.384 \\ 0.75 & 0.778 & 0.839 & 0.834 & 1 & 0.831 & 0.436 & 0.055 & 0.621 & 0.262 & 0.359 & 0.514 & 0.531 \\ 0.886 & 0.635 & 0.713 & 0.754 & 0.831 & 1 & 0.621 & 0.394 & 0.52 & 0.415 & 0.379 & 0.404 & 0.393 \\ 0.69 & 0.217 & 0.273 & 0.449 & 0.436 & 0.621 & 1 & 0.487 & 0.251 & 0.292 & 0.216 & 0.364 & 0.307 \\ 0.409 & 0.2 & 0.207 & 0.135 & 0.055 & 0.394 & 0.487 & 1 & 0.362 & 0.558 & 0.237 & 0.175 & 0.051 \\ 0.481 & 0.539 & 0.58 & 0.597 & 0.621 & 0.52 & 0.251 & 0.362 & 1 & 0.643 & 0.628 & 0.283 & 0.259 \\ 0.517 & 0.245 & 0.276 & 0.393 & 0.262 & 0.415 & 0.292 & 0.558 & 0.643 & 1 & 0.672 & 0.227 & 0.153 \\ 0.431 & 0.218 & 0.319 & 0.45 & 0.359 & 0.379 & 0.216 & 0.237 & 0.628 & 0.672 & 1 & 0.177 & 0.172 \\ 0.424 & 0.609 & 0.631 & 0.335 & 0.514 & 0.404 & 0.364 & 0.175 & 0.283 & 0.227 & 0.177 & 1 & 0.964 \\ 0.388 & 0.621 & 0.646 & 0.384 & 0.531 & 0.393 & 0.307 & 0.051 & 0.259 & 0.153 & 0.172 & 0.964 & 1 \end{pmatrix}$$

Tableau 3-8. Matrice de corrélation de la copule

La distribution du nombre de tests localement significatifs, N , peut être approchée par simulation et est représentée sur la Figure 3-26. Les valeurs critiques de N obtenues par les différentes méthodes disponibles sont reportées dans le Tableau 3-9. Les conclusions sont identiques à celles du cas précédent. Les méthodes du Bootstrap-années et par copule Gaussienne sont efficaces pour évaluer la significativité régionale, alors que l'utilisation d'un effectif efficace conduit à une nette surestimation de la valeur critique.

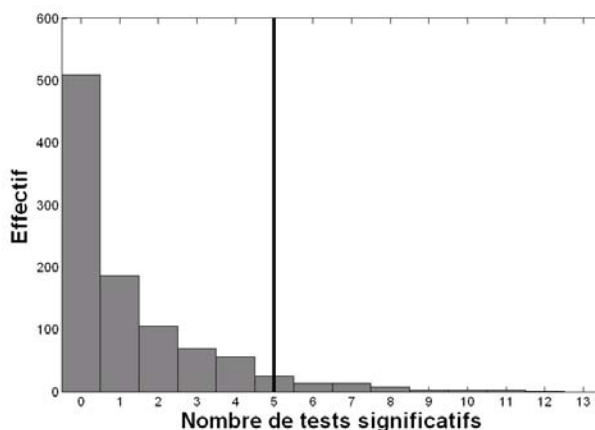


Figure 3-26. Distribution simulée de N (données simulées par une copule Gaussienne)

Méthode	Valeur critique
Valeur « théorique »	5
Approximation binomiale	4
Bootstrap-années	5
Effectif efficace	11
Copule Gaussienne	5

Tableau 3-9. Estimation des valeurs critiques de N (données simulées par une copule Gaussienne)

II.3.3. Simulation par un générateur de pluies

Dans le paragraphe précédent, l'utilisation d'une copule estimée sur des données réelles nous a permis de recréer une dépendance spatiale un peu plus réaliste que celle utilisée en II.3.1. , mais qui reste tributaire d'un certain nombre d'hypothèses sous-jacentes à l'utilisation de la copule. Nous allons à présent utiliser un générateur de pluies, dont le logiciel TBM (version 0.99) a été développé au Cemagref par M.H. Ramos et E. Leblois, afin de simuler des champs spatiaux de précipitations (Ramos, 2002). Nous allons nous intéresser aux valeurs ponctuelles de pluies, mais aussi aux valeurs cumulées par zones. Le domaine que nous simulons est un carré de 20 km de côté, avec une résolution de 200 m. La distribution des valeurs ponctuelles est une loi log-normale de paramètres (1;1). La structure spatiale des pluies est déterminée par un variogramme de type sphérique. La portée sera fixée à 5 km (cas 1) et 50 km (cas 2). La Figure 3-27 illustre un champ simulé, et montre la répartition des pluviomètres (carrés), ainsi que des zones sur lesquelles nous effectuons le cumul. Ces zones correspondent à des bassins versants réels, dans le bassin de l'Arc, mais les caractéristiques des pluies simulées ont par contre été choisies de manière arbitraire.

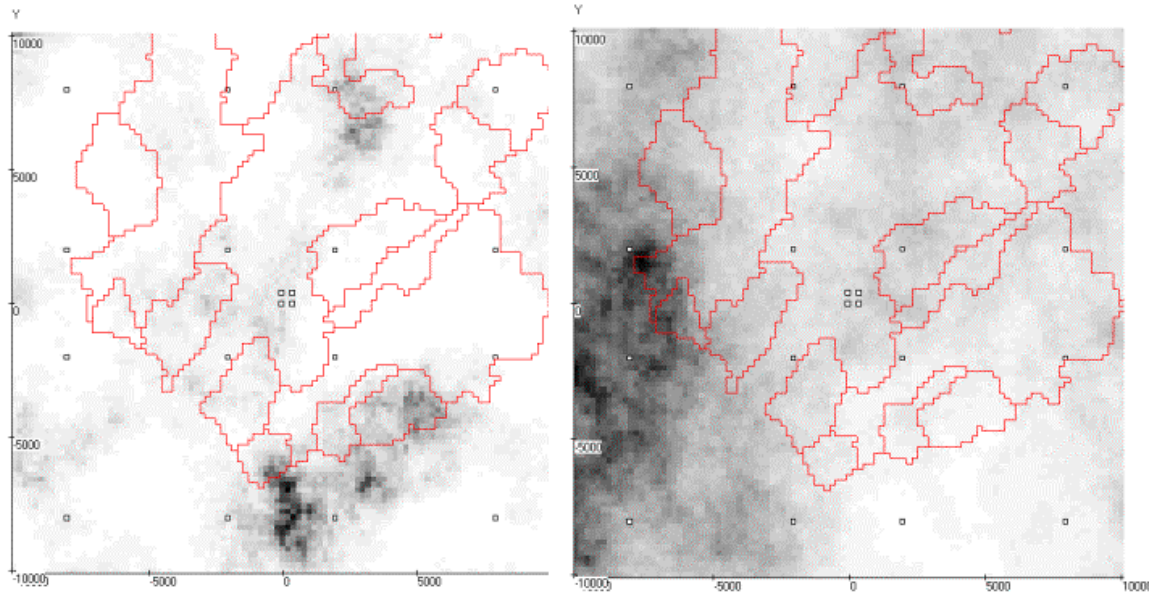


Figure 3-27. Exemples de champs simulés par TBM (unité : mètres, portée de 5 km à gauche, 50 km à droite). Les carrés représentent la localisation des pluviomètres, les contours représentent les limites des bassins versants (zones de cumul).

Le nombre de champs simulés a été fixé à 80. Nous obtenons ainsi des jeux de données multivariés, représentant 80 années de données sur 20 stations (ou 20 zones). Nous pouvons alors appliquer des tests de stationnarité à chacune de ces vingt stations, et dénombrer le nombre de résultats significatifs. En réitérant cette simulation 100 fois, nous pouvons approcher la distribution de N , et disposer ainsi d'une valeur critique de référence, qui peut alors être comparée aux résultats obtenus avec les différentes méthodes à notre disposition (Tableau 3-10). Avec ce jeu de données spatialisées que l'on peut considérer comme assez réaliste, les résultats sont identiques à ceux obtenus précédemment : les méthodes du Bootstrap-années et par copule Gaussienne fournissent des ordres de grandeurs acceptables pour la valeur critique de N , alors que le passage par un nombre équivalent de stations conduit à une très forte surestimation. Ces deux méthodes restent efficaces dans le cas des cumuls par zone, alors que la structure de la dépendance spatiale devrait être plus complexe que dans le cas des pluviomètres ponctuels. L'approximation binomiale, quant à elle, perd de sa validité lorsque la dépendance spatiale est plus importante (i.e. lorsque la portée augmente).

Portée	5 km		50 km	
	pluvio	cumul	pluvio	cumul
Valeur « théorique »	4	5	8	8
Approximation binomiale	5	5	5	5
Bootstrap-années	5	5	7	8
Effectif efficace	5	7	18	19
Copule Gaussienne	6	5	9	8

Tableau 3-10. Estimation des valeurs critiques de N (données simulées par TBM).

II.4. Conclusion sur la significativité régionale

Cette partie a permis de démontrer l'influence de la dépendance spatiale des phénomènes que nous étudions. Avant même de parler de cohérence régionale et d'attribution d'un changement observé à une cause climatique, il apparaît que cette dépendance peut induire en erreur quant à la significativité même de ce changement. L'intérêt d'utiliser des tests statistiques étant justement de pouvoir associer à une prise de décision un risque d'erreur, il semble indispensable de vérifier que le nombre de résultats localement significatifs est suffisant avant de tenter d'évaluer le déterminisme des tendances observées.

Sur la base des quelques simulations effectuées, il apparaît que nous ne sommes pas dénués d'outils dans ce contexte. La méthode de Bootstrap-années décrite par Douglas *et al.* (2000) semble être la plus recommandable, car elle permet d'obtenir de bonnes approximations des valeurs critiques de N , sans faire d'hypothèses quant à la distribution marginale des observations ou la structure de la dépendance.

Remarquons cependant que toutes les méthodes présentées ici ne s'appliquent qu'à des variables échantillonnées de manière annuelle. Les variables relatives à des dépassements de seuil sont plus délicates à traiter, pour la simple raison que les échantillons locaux ne peuvent pas être simplement concaténés afin de créer un jeu de données multivarié (les événements n'étant pas forcément concomitants). Pour les variables de débits, on pourra approcher la valeur critique relative aux pointes sup-seuil par celle obtenue sur les maxima annuels, en argumentant que si notre méthode d'échantillonnage change, la variable physique que l'on observe reste la même. Par contre, les variables relatives aux formes des hydrogrammes restent problématiques. Un minimum consistera donc à utiliser l'approximation binomiale pour évaluer la significativité régionale des changements observés sur ces variables.

III. Détection de changements régionalement cohérents

Bien que ses conséquences puissent être locales, le changement climatique est un phénomène à grande échelle. On s'attend donc à ce que ses effets s'expriment de manière cohérente pour des bassins versants aux fonctionnements hydrologiques similaires. Cependant, vérifier cette présomption à partir de données réelles soulève un certain nombre de problèmes.

En premier lieu, il convient de définir ce qu'est une région homogène du point de vue hydro-climatique. Cet aspect a été abordé au chapitre précédent, nous supposons donc que nous disposons d'une région homogène. Le second problème majeur est d'ordre méthodologique, et concerne la difficulté liée aux données extrêmes multivariées. La dépendance spatiale entre les stations joue ici encore un rôle prépondérant, comme c'était le cas pour la significativité régionale. Nous allons donc proposer quelques pistes envisageables afin d'évaluer et si possible de quantifier le plus rigoureusement possible cet aspect de cohérence régionale.

III.1. Principales approches utilisées

III.1.1. Approches descriptives

A l'issue d'une analyse locale de stationnarité effectuée sur un ensemble de sites, les résultats sont en général projetés sur une carte. La Figure 3-28 est un exemple d'une telle représentation. On s'aperçoit que des changements cohérents peuvent avoir tendance à se regrouper (Ouest du Massif Central, par exemple). Cette tendance au regroupement peut cependant être simplement due à la dépendance spatiale entre les stations. On s'aperçoit aussi qu'il est important de ne pas raisonner en terme de distances géographiques, mais bel et bien en terme de fonctionnement hydrologique : ainsi, les tendances à la hausse observées sur les Cévennes ne sont pas forcément incompatibles avec l'ensemble des tendances à la baisse du Massif Central, malgré la proximité géographique. En fait, ce type de représentation n'est pas suffisant pour juger de la cohérence des résultats. Par contre, il peut être utilisé pour détecter des changements d'origine non climatique : ainsi, des tendances opposées observées sur des sites très proches (au sens hydro-climatique) peuvent être le signe de changements liés à des problèmes météorologiques, par exemple.

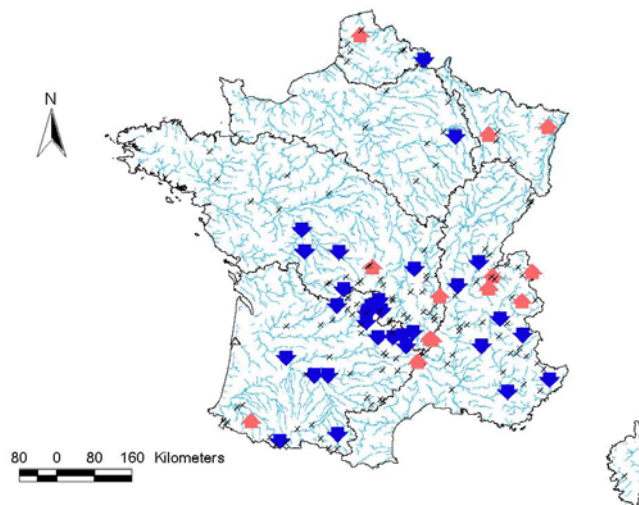


Figure 3-28. Exemple de cartographie des résultats locaux.

Un premier moyen simple de quantifier le déterminisme des changements observés localement est d'effectuer une analyse de variance (ANOVA). Supposons par exemple que toutes les stations aient été traitées avec un test de Mann-Kendall, conduisant à un échantillon de statistiques observées $(S_j)_{j=1,\dots,p}$. On peut alors rechercher l'influence de différents facteurs sur les valeurs de la statistique, comme par exemple le régime du cours d'eau, la zone climatique ou encore la superficie du bassin versant. Cependant, la significativité de ces influences devra être considérée avec prudence, car le modèle d'ANOVA suppose que les données sont indépendantes : ceci n'est malheureusement pas le cas, puisque la dépendance spatiale entre stations se répercute sur les statistiques de test.

III.1.2. Création de variables régionales

Afin de s'affranchir des problèmes liés à l'aspect multivarié de l'analyse, une possibilité consiste à se ramener au cas univarié en calculant des variables définies régionalement. Par exemple, pour une région donnée, il est possible de calculer le nombre annuel total d'événements de crues. Pour une région à la saisonnalité des événements de crues bien marquée (crues nivales, par exemple), la date angulaire moyenne des événements peut aussi être calculée annuellement. Le problème est un peu plus compliqué pour les variables de débit : les ordres des grandeurs n'étant pas homogènes d'une station à l'autre, la prise de moyenne n'a aucun sens puisqu'elle va privilégier les stations aux plus forts débits. Il convient donc dans un premier temps de normaliser les échantillons locaux. Le plus simple semble de considérer les rangs des observations plutôt que les valeurs brutes. Une autre solution consiste à ajuster une loi de probabilité sur chaque site, et à transformer les valeurs brutes en périodes de retour.

Quelle que soit la transformation adoptée, il convient de vérifier que la dépendance spatiale n'induit pas de biais dans le risque de première espèce, comme c'était le cas pour l'autocorrélation (cf. section I.2.3.). Plus précisément, l'agrégation de variables stationnaires dépendantes peut-elle créer des tendances « artificielles » sur la variable régionale ? Afin d'évaluer cet éventuel problème, nous avons repris l'ensemble des simulations effectuées dans la partie II.3. Les variables locales ont été agrégées en calculant annuellement le rang moyen des événements. Finalement, un test de Mann-Kendall a été appliqué au risque $\alpha=0.1$. Les résultats, résumés dans le Tableau 3-11, montrent que le taux de rejet reste cohérent avec le risque de première espèce spécifié. L'agrégation de variables dépendantes ne semble donc pas introduire de biais sur le risque de première espèce.

Simulation	Gaussienne multivariée, $\rho=$						Copule	TBM			
	0	0.2	0.4	0.6	0.8	0.9		pluvio 5 km	cumul 5 km	pluvio 50 km	cumul 50 km
Taux de rejet (%)	10.7	9.1	9.7	10.0	10.2	9.3	10.7	9.0	9.0	12.0	8.0

Tableau 3-11. Taux de rejet de H_0 pour la variable régionale "rang moyen des événements", avec un test de Mann-Kendall au risque 10%.

Toutefois, il convient de rester prudent sur l'interprétation des résultats du test, qui ne garantissent pas à proprement parler la cohérence des changements. Considérons en effet l'exemple caricatural suivant : 10 stations très fortement corrélées présentent une tendance à la hausse, et une onzième indépendante des dix autres présente une tendance à la baisse de même ampleur. La moyenne de ces variables risque de présenter une tendance à la hausse, car le poids des stations dépendantes sera très important. Ce faisant, on surestime très fortement l'information portée par ces dix stations, qui ne constituent en fait, en terme d'effectif efficace, guère plus d'une station.

III.1.3. Création de statistiques régionales

Une alternative à la création de variables régionales consiste à agréger non pas les données mais les statistiques de test calculées localement. Notons $S^{(1)}, \dots, S^{(p)}$ les statistiques de Mann-Kendall calculées sur chaque site. La statistique régionale de Mann-Kendall sera alors définie par (Douglas *et al.*, 2000; Yue et Wang, 2002b):

$$S_R = \frac{1}{p} \sum_{j=1}^p S^{(j)}, \text{ où } S^{(j)} = \sum_{k=1}^{n-1} \sum_{i=k+1}^n \text{signe}(X_i^{(j)} - X_k^{(j)}) \quad (3-37)$$

La statistique de test sera finalement donnée par l'équation (3-3) appliquée à S_R . Le problème vient du calcul de la variance de cette statistique régionale. Sous l'hypothèse d'indépendance spatiale, toutes les statistiques locales ayant la même distribution normale sous H_0 , cette variance est donc simplement égale à $\text{Var}(S^{(i)})/p$. Si les séries sont dépendantes, alors Douglas (2000) a mis en évidence la relation suivante :

$$\text{Var}(S_R) = \frac{n(n-1)(2n+5)}{18p} (1 + (p-1)\bar{\rho}), \text{ avec } \bar{\rho} = \frac{2 \sum_{i=2}^p \sum_{j=1}^{i-1} \rho_{i,j}}{p(p-1)} \quad (3-38)$$

où $\bar{\rho}$ est la corrélation moyenne entre statistiques locales. En pratique, les corrélations entre statistiques ne peuvent pas être évaluées directement, elles seront donc remplacées par les corrélations entre variables locales (Yue et Wang, 2002b):

$$r_{i,j} = \frac{\frac{1}{n} \sum_{k=1}^n (X_k^{(i)} - \bar{X}^{(i)}) (X_k^{(j)} - \bar{X}^{(j)})}{\sigma_{X^{(i)}} \sigma_{X^{(j)}}} \quad (3-39)$$

Notons qu'il s'agit là d'une hypothèse extrêmement forte : rien ne garantit que les corrélations entre statistiques locales soient égales aux corrélations entre variables locales. De plus, l'emploi d'un coefficient de corrélation se justifie pour modéliser la dépendance entre statistiques locales, dont la distribution jointe doit asymptotiquement être Gaussienne multivariée, mais la dépendance entre variables pourrait être bien plus complexe.

Nous avons souhaité vérifier que ce test de Mann-Kendall régional n'introduisait pas de biais vis-à-vis du risque de première espèce. Nous avons donc encore une fois repris l'ensemble des simulations de la partie II.3. Afin de mettre en évidence l'influence de la dépendance spatiale, nous l'avons également appliqué en faisant l'hypothèse d'indépendance des sites. Les résultats, présentés dans le Tableau 3-12, montrent que le correctif de variance apporté par l'équation (3-38) permet de garder un taux de rejet de H_0 proche du risque de première espèce spécifié. A l'inverse, ignorer la dépendance spatiale conduit à un test très fortement biaisé.

Simulation	Gaussienne multivariée, $\rho=$						Copule	TBM			
	0	0.2	0.4	0.6	0.8	0.9		pluvio 5 km	cumul 5 km	pluvio 50 km	cumul 50 km
dépendance ignorée	9.2	17.8	25.5	40.4	55.7	66.5	52.0	25.0	31.0	65.0	68.0
dépendance prise en compte	9.4	10.7	7.4	9.9	8.4	9.2	11.0	8.0	9.0	11.0	11.0

Tableau 3-12. Taux de rejet (%) de H_0 par le test régional de Mann-Kendall ($\alpha=0.1$).

Comme nous l'avons déjà signalé à la section III.1.2. , un résultat significatif doit être interprété avec prudence, car il ne garantit pas la cohérence des changements observés localement. Pour illustrer ce phénomène, nous avons simulé l'exemple donné précédemment : 11 stations avec 40 années de données sont simulées à partir d'une loi normale centrée de matrice de variance :

$$\Sigma = \begin{pmatrix} 1 & 0.99 & \dots & 0.99^9 & 0 \\ 0.99 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0.99 & \vdots \\ 0.99^9 & \dots & 0.99 & 1 & 0 \\ 0 & \dots & 0 & 0 & 1 \end{pmatrix} \quad (3-40)$$

Une tendance est ensuite ajoutée aux données ainsi simulées : dans le cas d'un changement cohérent, cette tendance est la même pour les 11 stations. Dans le cas d'un changement incohérent, la tendance appliquée à la onzième station est l'opposée de celle appliquée aux dix autres. Nous avons finalement appliqué un test de Mann-Kendall à la variable régionale « rangs moyens annuels », ainsi qu'un test régional de Mann-Kendall. En recommençant cette opération 1000 fois, il est possible d'évaluer la faculté des tests à détecter ces deux types de changements. Les résultats de ces simulations, présentés Figure 3-29, montrent tout d'abord que les deux approches sont équivalentes sur cet exemple en termes de taux de détection du changement. D'autre part, aucune des deux ne parvient à différencier un changement cohérent d'un changement incohérent, puisque le taux de détection dans ce dernier cas reste important.

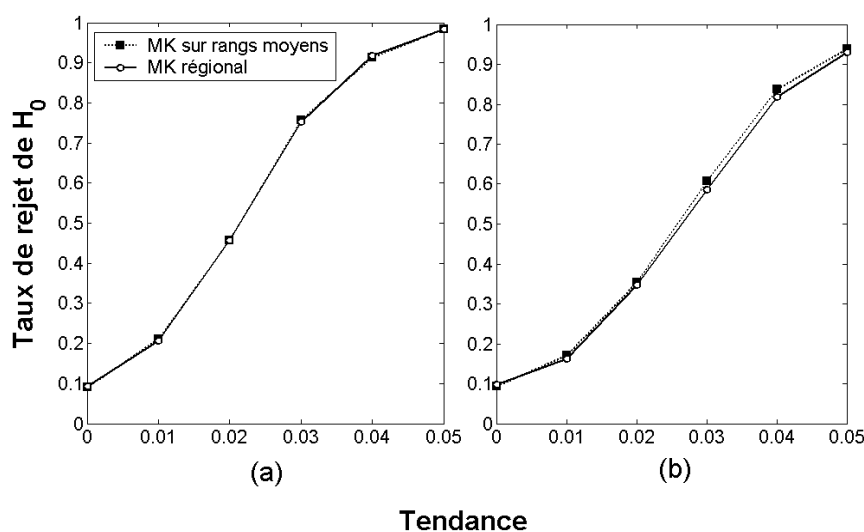


Figure 3-29. Puissance comparée du test de Mann-Kendall sur variable régionale et du test de Mann-Kendall régional. (a) Changement cohérent. (b) Changement incohérent.

III.2. Modèles régionaux non stationnaires

Face aux problèmes soulevés dans les parties précédentes, il convient de reconsidérer la définition que nous donnons à un changement cohérent. Dans les deux précédentes sections, les hypothèses H_0 pourraient être traduites par « la variable agrégée est stationnaire » (création d'une variable régionale), ou « toutes les stations sont stationnaires » (test de Mann-Kendall régional). En un certain sens, ces tests seraient donc plus propices à déterminer une certaine significativité régionale qu'une cohérence. La question à laquelle nous allons donc essayer de répondre à présent sera donc : « Un modèle régional de changement est-il plus adéquat que son équivalent stationnaire ? ».

III.2.1. Une approche semi-paramétrique

Etant donné la difficulté à prendre en compte explicitement des dépendances spatiales dans le cadre de phénomènes extrêmes, il semble naturel de se ramener dans un premier temps au contexte Gaussien, qui permet d'effectuer simplement des analyses multivariées. Dans cette optique, les données locales doivent être transformées pour suivre une loi normale. Une transformation de type « normal score » sera effectuée à cet effet. Dans un second temps, nous supposons que les données transformées suivent une loi Gaussienne multivariée, ce qui permettra de construire un test LR (que nous nommerons LR-Reg) afin de comparer un modèle stationnaire et un modèle à tendance régionale. Il s'agit là d'une hypothèse forte : en effet, les lois marginales seront bien Gaussiennes par construction, mais ceci ne constitue pas une condition suffisante pour assurer la normalité de la loi jointe. Il sera donc nécessaire de vérifier que l'hypothèse de normalité est acceptable. De nombreux outils existent à cet effet (Mardia, 1980), que nous ne développerons pas dans cette étude.

Description du test

La matrice de données est notée \mathbf{X} , et comporte n années en ligne et p stations en colonnes. $X_k^{(i)}$ représente la variable d'intérêt en une année k pour un site i . La transformation « normal score » peut être décrite comme suit :

$$\tilde{X}_k^{(i)} = \phi^{-1} \left(\hat{F}_i (X_k^{(i)}) \right) \quad (3-41)$$

Dans cette expression, ϕ est la fonction de répartition d'une loi Gaussienne centrée réduite, et \hat{F}_i est la fonction de répartition empirique pour le site i . Nous supposons donc que la matrice transformée $\tilde{\mathbf{X}}$ est un échantillon issu d'une loi normale de dimension p :

$$\tilde{\mathbf{X}}_k \sim N(\boldsymbol{\mu}; \boldsymbol{\Sigma}) \quad (3-42)$$

Sous l'hypothèse de stationnarité, le vecteur moyenne est identiquement nul par construction, et la matrice de variance peut être estimée par son estimateur classique de maximum de vraisemblance :

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n} \tilde{\mathbf{X}}^T \tilde{\mathbf{X}} \quad (3-43)$$

Supposons à présent que les données soient affectées par une tendance identique pour tous les sites. Cette situation est paramétrée de la manière suivante :

$$E(\tilde{X}_k^{(i)}) = \alpha y_k, \quad (3-44)$$

où y_k représente l'année correspondant à la k^e donnée. Puisque, par construction, les moyennes de tous les sites sont nulles, il conviendra de remplacer l'année brute y_k par une année centrée $\tilde{y}_k = y_k - \bar{y}$, de sorte que :

$$\sum_{k=1}^n \alpha \tilde{y}_k = 0, \quad (3-45)$$

L'estimation du paramètre de tendance peut se faire explicitement conditionnellement à la matrice de variance. Notons $(\gamma_{i,j})_{i,j=1,\dots,p}$ les termes de l'inverse de la matrice de variance $\boldsymbol{\Sigma}^{-1}$. Il est aisé de montrer que la dérivée de la log-vraisemblance par rapport à α s'annule au point suivant :

$$\hat{\alpha} = \frac{\sum_{k=1}^n \left\{ \tilde{y}_k \sum_{i,j=1}^p \gamma_{i,j} \tilde{x}_k^{(j)} \right\}}{\left(\sum_{k=1}^n \tilde{y}_k^2 \right) \left(\sum_{i,j=1}^p \gamma_{i,j} \right)} \quad (3-46)$$

En conséquence, si la matrice de variance est estimée comme décrit dans l'équation (3-43), l'estimateur de tendance régionale peut s'écrire sous la forme matricielle suivante :

$$\hat{\alpha} = \frac{\mathbf{I}_p^T (\tilde{\mathbf{X}}^T \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^T \tilde{\mathbf{Y}}}{\tilde{\mathbf{Y}}^T \tilde{\mathbf{Y}} \mathbf{I}_p^T (\tilde{\mathbf{X}}^T \tilde{\mathbf{X}})^{-1} \mathbf{I}_p}, \quad (3-47)$$

où $\tilde{\mathbf{Y}}$ est le vecteur colonne des années centrées et \mathbf{I}_p est un vecteur colonne de dimension p dont toutes les valeurs sont égales à un.

Il est alors possible de calculer la statistique du test LR, comme expliqué en section I.1.4. :

$$\begin{aligned} T &= -2 \left(L_0(\mathbf{X}; \hat{\theta}_0) - L_1(\mathbf{X}; \hat{\theta}_1) \right) \\ &= -2 \left(\sum_{k=1}^n \log \left(N(\tilde{\mathbf{x}}_k; \mathbf{0}, \hat{\Sigma}) \right) - \sum_{k=1}^n \log \left(N(\tilde{\mathbf{x}}_k; \hat{\alpha} \tilde{\mathbf{y}}_k, \hat{\Sigma}) \right) \right), \end{aligned} \quad (3-48)$$

et de comparer la valeur obtenue aux valeurs critiques issues d'une distribution du χ^2 à 1 degré de liberté.

Notons que l'estimateur proposé dans l'équation (3-47) a le mérite d'être explicite, mais ne correspond pas forcément au maximum de la vraisemblance, puisqu'il a été obtenu en utilisant une procédure conditionnelle. Ceci devrait avoir pour effet de donner des déviations plus faibles, et donc un test trop conservatif. En toute rigueur, l'ensemble des paramètres (tendance + matrice de variance) devrait être estimé simultanément par une méthode d'optimisation. On peut cependant s'attendre à des problèmes numériques, étant donnée l'augmentation rapide du nombre de paramètres à estimer avec la dimension, alors que la formule matricielle de l'équation (3-47) peut se calculer en toute dimension. L'impact de cette approximation sur le comportement du test doit toutefois être évalué.

Respect du risque de première espèce

Dans un premier temps, il convient de vérifier que ce test n'est pas biaisé. Ceci pourrait être le cas d'une part car l'estimateur de tendance ne maximise pas forcément la vraisemblance, d'autre part car l'approximation par la distribution du χ^2 n'est valable qu'asymptotiquement. Le respect du risque de première espèce a donc été évalué par simulation de 5000 échantillons stationnaires, provenant d'une loi Gaussienne de dimension d et de taille n . La matrice de variance est donnée par l'équation (3-36) avec $\rho=0.8$. Les résultats, présentés en Figure 3-30, montrent que le test semble correctement calé quelles que soient la dimension et la taille de l'échantillon, puisque les taux de rejet restent compris entre 9% et 11.4%, pour un risque de première espèce égal à 0.1.

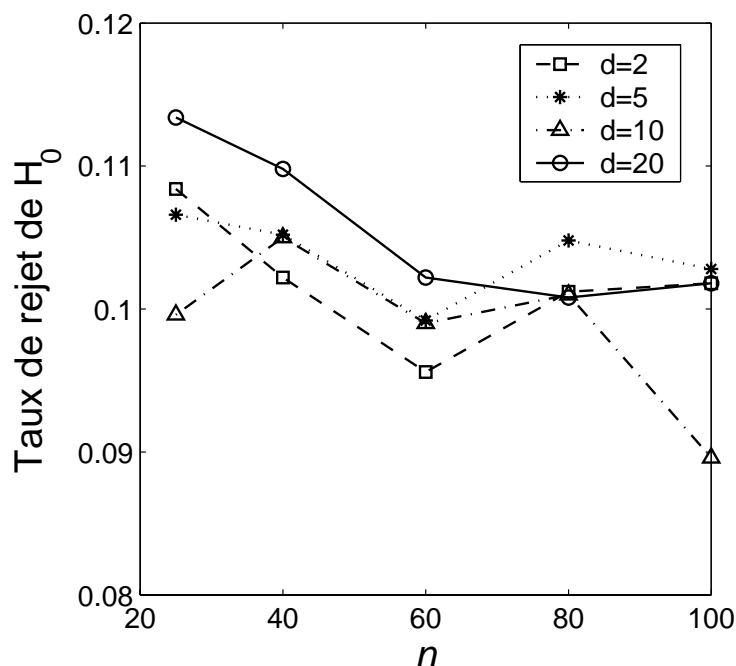


Figure 3-30. Taux de rejet de H_0 par le test LR-Reg ($\alpha=0.1$), pour des données stationnaires Gaussiennes de dimension d et de taille n

Des analyses complémentaires ont également été effectuées : nous avons repris l'ensemble des simulations de la section II.3. (loi Gaussienne, copule Gaussienne et TBM) afin de simuler des données multivariées stationnaires. Il est ainsi possible de vérifier que les taux de rejet sont proches du risque α spécifié (en l'occurrence, $\alpha=0.1$). Ces données stationnaires ont également été testées en supposant l'indépendance spatiale, c'est à dire que la matrice de variance a été fixée égale à la matrice identité dans le test LR-Reg.

Simulation	Gaussienne multivariée, $\rho=$						Copule	TBM			
	0	0.2	0.4	0.6	0.8	0.9		pluvio 5 km	cumul 5 km	pluvio 50 km	cumul 50 km
Dépendance ignorée	4.2	7.6	16.4	24.6	37.2	44.8	38.8	14.0	23.0	54.0	48.0
LR-Reg	9.1	10.1	11.2	9.9	10.8	10.6	10.7	12.0	11.0	10.0	10.0

Tableau 3-13. Taux de rejet (%) de H_0 par le test LR-Reg ($\alpha=0.1$)

Les résultats, présentés dans le Tableau 3-13, montrent tout d'abord que la prise en compte des dépendances est indispensable au respect du risque de première espèce. De plus, les taux de rejet obtenus avec le test LR-Reg sont encore une fois acceptables pour toutes les simulations, le test ne semble donc pas présenter un quelconque biais. Dans le cas d'une loi Gaussienne ou d'une copule Gaussienne, ce résultat n'est pas étonnant, puisque les hypothèses formulées pour la construction du test sont respectées. Le cas du générateur de pluies est plus général, et il s'avère que les données simulées étaient en bonne adéquation avec ces hypothèses. Dans la pratique, insistons sur le fait que l'hypothèse de normalité jointe des données transformées n'est pas garantie, et qu'il conviendra donc de vérifier que cette description est en adéquation avec les observations.

Aptitude à détecter des changements cohérents

Les simulations présentées en section III.1.3. ont été reprises, afin d'illustrer la différence de comportement entre le test que nous venons de construire, et des tests régionaux basés sur la création de variables ou de statistiques régionales. La Figure 3-31a montre que dans le cas d'un changement cohérent sur 11 stations, dont 10 sont très fortement dépendantes, tous les tests parviennent à détecter une tendance de manière satisfaisante. On note cependant un gain de puissance non négligeable avec le test LR-Reg. De plus, dans le cas d'un changement incohérent (Figure 3-31b), le test LR-Reg ne détecte pas de changement (il aurait même tendance à devenir conservatif), contrairement aux deux autres tests, qui ne parviennent pas à prendre en compte correctement la redondance d'information liée à la très forte dépendance entre les 10 premières stations.

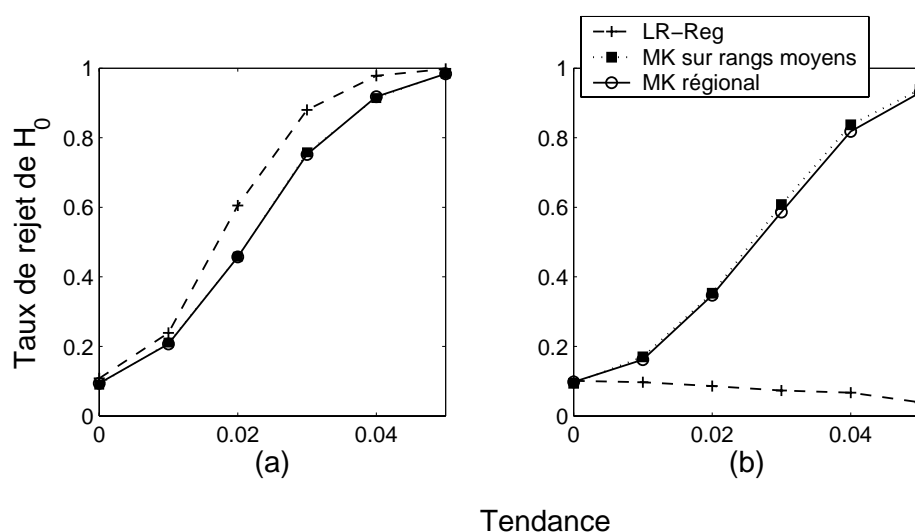


Figure 3-31. Puissance comparée du test LR-Reg, du test de Mann-Kendall sur variable régionale et du test de Mann-Kendall régional. (a) Changement cohérent, (b) changement incohérent

Le test LR-Reg semble donc répondre à l'objectif principal que nous nous étions fixés, à savoir la construction d'un test capable de différencier un changement cohérent d'un changement incohérent. De plus, ce test a l'avantage de la simplicité, puisque la statistique de test peut s'écrire de manière explicite, sans recourir à des méthodes d'optimisation numérique. Néanmoins, deux approximations peuvent limiter le champ d'application du test : tout d'abord, la normalité de la matrice de données transformée n'est pas garantie, il convient donc de la vérifier soigneusement. D'autre part, le test du rapport des vraisemblances maximales est construit avec un estimateur qui n'est pas forcément celui qui maximise la vraisemblance du modèle à tendance. Cependant, sur la base des simulations que nous avons effectuées, cette approximation ne semble pas induire de biais.

III.2.2. Approches paramétriques

Dans le cas univarié, les tests paramétriques se sont révélés plus puissants que leurs homologues non paramétriques pour détecter des tendances ou des ruptures. Il devrait logiquement en être de même dans le cas multivarié. Cependant, un certain nombre de problèmes techniques apparaissent, liés encore une fois à la difficulté de prendre en compte les dépendances entre valeurs extrêmes. Nous allons nous intéresser dans cette section à quelques modèles à tendance régionale, et tenter de construire des tests LR régionaux paramétriques.

Quelques modèles possibles

Nous allons nous placer dans le cadre d'une variable extrême suivant localement une loi GEV (maxima annuels, par exemple). Le modèle le plus simple consiste à supposer que les distributions locales sont affectées par un changement de même ampleur, paramétré de la manière suivante :

$$\text{Modèle M1 : } X^{(i)} \square GEV(\alpha_i, \beta_i (1 + \delta t), \xi_i), \quad (3-49)$$

On peut d'ores et déjà anticiper un certain nombre de problèmes liés à l'estimation de ces paramètres, surtout si l'on considère un nombre de sites élevé. Nous allons donc chercher à simplifier ce premier modèle. Une première hypothèse consiste à supposer que le paramètre de forme, qui est en partie relié au comportement des pluies extrêmes, est homogène à l'intérieur d'une même région hydro-climatique. Il s'agit d'une hypothèse relativement forte, qui pourrait être prise en défaut dans une région trop hétérogène, et qui doit être validée sur les observations. Néanmoins, cette hypothèse est effectuée, explicitement ou implicitement, dans la plupart des méthodes de régionalisation.

$$\text{Modèle M2 : } X^{(i)} \square GEV(\alpha_i, \beta_i (1 + \delta t), \xi), \quad (3-50)$$

Un modèle encore plus simplifié peut être dérivé à l'aide de la méthode de l'indice de crue (Dalrymple, 1960). Brièvement, cette méthode consiste à supposer qu'il existe une constante I telle que les variables $X^{(i)}/I$ aient toutes la même distribution. Il est possible de démontrer que ceci implique que, d'une part, le paramètre de forme est identique pour toutes les distributions locales, et d'autre part, que le rapport α_i/β_i reste également constant. Ceci peut se traduire par le modèle à tendance régionale suivant :

$$\text{Modèle M3 : } X^{(i)} \square GEV(C\beta_i, \beta_i (1 + \delta t), \xi), \quad (3-51)$$

Notons qu'avec le modèle M3, le rapport $\alpha_i(t)/\beta_i(t)$ reste bien égal sur toutes les stations à t fixé, mais évolue avec le temps. Pour forcer ce rapport à rester constant aussi bien temporellement que spatialement, il faudra considérer le modèle suivant :

$$\text{Modèle M4 : } X^{(i)} \square GEV(C\beta_i (1 + \delta t), \beta_i (1 + \delta t), \xi), \quad (3-52)$$

Pour finir, l'hypothèse que le changement s'exprime exactement de la même façon sur toutes les stations à l'intérieur d'une même région hydro-climatique peut sembler trop contraignante, malgré tout le soin apporté à la définition de ces régions. Afin de rendre le modèle plus souple, une solution consisterait à considérer que l'effet du changement climatique est en partie aléatoire, et donc de supposer que la tendance δ n'est pas fixe, mais suit une certaine loi de probabilité ($N(m, \sigma^2)$ par exemple). Malheureusement, il n'est alors pas possible d'écrire analytiquement la densité de probabilité des observations, et on peut redouter un certain nombre de problèmes numériques.

Dans la suite de ce paragraphe, nous travaillerons avec le modèle M2 (équation (3-50)).

Vraisemblances

La première chose à faire pour construire un test LR est de calculer la vraisemblance des observations. On se heurte donc immédiatement à une difficulté majeure : quelle est la loi d'un vecteur d'observations ? Le seul cas où cette difficulté disparaît est le cas où les stations sont indépendantes spatialement. On peut alors écrire :

$$\begin{aligned}
 p(\mathbf{X}; \boldsymbol{\theta}) &= p\left(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p)}; (\alpha_1, \beta_1, \dots, \alpha_p, \beta_p, \xi, \delta)\right) \\
 &= \prod_{i=1}^n p\left(x_i^{(1)}, \dots, x_i^{(p)}; (\alpha_1, \beta_1, \dots, \alpha_p, \beta_p, \xi, \delta)\right) \\
 &= \prod_{i=1}^n \prod_{j=1}^p GEV\left(x_i^{(j)}; (\alpha_j, \beta_j (1 + \delta i), \xi)\right)
 \end{aligned} \tag{3-53}$$

Les estimateurs seront donc les paramètres maximisant cette quantité. Dans la pratique, les méthodes numériques de maximisation de la vraisemblance semblent fonctionner jusqu'à une dizaine de stations (soit 22 paramètres, avec la fonction *fminsearch* de matlab). Il est indispensable par contre d'être très attentif aux points de départ : une première étape d'estimations locales est recommandable.

Dans le cas de données dépendantes, le modèle de vraisemblance doit être modifié. En effet, bien que négliger la dépendance spatiale n'ait que peu d'effet sur les estimations ponctuelles (Katz *et al.*, 2002), cela peut entraîner une sous-estimation de la variance des estimateurs. En conséquence, le test de déviance pourrait également être biaisé.

Afin de tenter de corriger les effets de la dépendance spatiale, nous allons utiliser une copule Gaussienne. Encore une fois, l'emploi d'un tel outil doit être conditionné à la vérification d'un certain nombre d'hypothèses, parmi lesquelles la plus importante est qu'il est possible de décrire la dépendance par des corrélations. Même si une telle approximation n'est pas optimale, nous ne disposons de toutes façons que de très peu d'outils dès lors que le nombre de stations est supérieur à deux ou trois.

Le densité de probabilité d'un vecteur est modélisée de la façon suivante par une copule Gaussienne :

$$\begin{aligned}
 f(x^{(1)}, \dots, x^{(p)}) &= f_1(x^{(1)}) \dots f_p(x^{(p)}) \times \\
 &|\boldsymbol{\Sigma}|^{-1/2} \exp\left\{-\frac{\left[\phi^{-1}\left(F_1(x^{(1)})\right), \dots, \phi^{-1}\left(F_p(x^{(p)})\right)\right] \left[\boldsymbol{\Sigma}^{-1} - \mathbf{I}\right] \left[\phi^{-1}\left(F_1(x^{(1)})\right), \dots, \phi^{-1}\left(F_p(x^{(p)})\right)\right]^T}{2}\right\}
 \end{aligned} \tag{3-54}$$

Malgré son aspect très peu engageant, cette formule peut être utilisée dans l'équation (3-53) pour calculer la vraisemblance des observations. En toute rigueur, les paramètres de la copule (c'est à dire la matrice de corrélation $\boldsymbol{\Sigma}$) doivent être estimés en même temps que les autres paramètres du modèle régional. Cependant, il est totalement illusoire d'espérer que les méthodes numériques puissent estimer un si grand nombre de paramètres. Nous allons donc procéder de manière analogue au cas du test LR-Reg non paramétrique : dans une première étape, la matrice de variance sera estimée sur les données transformées par « normal score ». Ceci revient à calculer une corrélation basée sur les rangs, ce qui peut induire un biais. Phoon *et al.* (2004) proposent une correction pour tenter de minimiser ce biais. Dans un second temps, les paramètres des lois marginales seront estimés, conditionnellement à l'estimation de la matrice de variance effectuée dans la première étape, avec une méthode d'optimisation numérique. Finalement, la déviance entre le modèle à tendance et le modèle stationnaire peut être calculée et comparée aux quantiles d'une distribution du χ^2 .

D'un point de vue pratique, les méthodes numériques de maximisation de la vraisemblance (programmées avec la fonction *fminsearch* de matlab) sont très fortement ralenties par l'introduction de la copule dans la vraisemblance. De plus, à grande dimension (une dizaine de stations), le choix des valeurs initiales devient critique, et doit être effectué manuellement de manière empirique, ce qui est très fastidieux puisqu'il faut plusieurs dizaines de minutes pour éventuellement converger vers un minimum local.

Nous avons souhaité vérifier que les approximations effectuées pour construire le test LR ne conduisaient pas à un test biaisé. Cependant, à cause des problèmes pratiques soulevés précédemment, les simulations effectuées sont restées très limitées : il est en effet difficilement concevable de gérer manuellement les problèmes de convergence sur un grand nombre de données simulées. Seules 500 simulations de données bidimensionnelles stationnaires ont été effectuées, à l'aide d'une copule Gaussienne, avec des lois marginales de type Gumbel. La matrice de variance est donnée par l'équation (3-36) avec $\rho=0, 0.5$ et 0.8 . Les résultats sont présentés dans le Tableau 3-14. On peut observer que construire un test LR basé sur une hypothèse d'indépendance alors que les données sont dépendantes conduit à un test trop libéral. L'utilisation d'une copule Gaussienne semble permettre de conserver un taux de rejet de H_0 plus proche du risque de première espèce spécifié. Cependant, la validité de cette observation pourrait être remise en cause en dimension supérieure.

ρ	Taux de rejet de H_0	
	hypothèse d'indépendance	copule Gaussienne
0	13.4%	11.8%
0.5	16.2%	11.8%
0.8	24.6%	13.2%

Tableau 3-14. Taux de rejet de H_0 par un test LR-tendance régional paramétrique ($\alpha=0.1$).

III.3. Conclusion sur la cohérence régionale

Evaluer la cohérence des changements observés localement constituera une de nos principales armes pour l'attribution du changement à une cause climatique. En ce sens, il semble primordial de bien étudier la méthodologie dont nous disposons. Dans un premier temps, nous avons évalué deux approches permettant de tester un changement à l'échelle régionale. Malheureusement, ces techniques ne règlent pas complètement le problème induit par la dépendance spatiale, et doivent plutôt être interprétées comme des indicateurs d'une certaine significativité régionale. L'approche la plus prometteuse semble être de raisonner en terme de modèle régional de changement. Si un certain nombre de problèmes théoriques comme pratiques persistent, il est au moins possible d'évaluer la pertinence d'un tel modèle face à son homologue stationnaire.

La principale difficulté est liée à la modélisation de la dépendance entre des variables extrêmes. Il y a certainement beaucoup à attendre des futurs développements méthodologiques dans ce domaine, qui semble constituer un des principaux champs de recherche des statisticiens spécialisés dans les extrêmes. En attendant, nous devons effectuer un certain nombre de simplifications pour parvenir à prendre en compte ce phénomène. Dans cette optique, le test semi-paramétrique LR-Reg semble présenter de bonnes performances, tout en étant très simple d'utilisation, puisqu'il ne requiert pas d'effectuer des optimisations numériques. La construction de tests LR-Reg paramétriques est également possible, mais leur mise en place soulève un certain nombre de problèmes pratiques, ce qui peut s'avérer contraignant si un grand nombre de régions doit être testé.

Quel que soit le test utilisé, il conviendra de rester modéré sur le nombre de stations à étudier régionalement. Plus précisément, il semble inutile d'incorporer dans le jeu de données régional deux stations très corrélées, puisque le surplus d'information risque d'être négligeable face à l'augmentation de complexité du problème. De plus, l'estimation des covariances peut être impossible si le nombre de sites est trop important par rapport au nombre d'observations. Une étude préalable pourrait donc être bénéfique, afin de déterminer quelles stations sont « suffisamment indépendantes » pour intégrer le jeu de données régional.

L'analyse en composantes principales est un outil adapté à ce type d'étude, puisqu'elle permet de visualiser simplement la redondance d'information entre stations.

En conclusion, signalons que l'approche Bayésienne pourrait être un outil adéquat pour l'estimation de modèles régionaux complexes. Elle n'a pas été développée dans cette partie, car elle ne conduit pas à la construction de tests statistiques, mais plutôt à des méthodes de comparaison de modèles. Cependant, nous y reviendrons en détail ultérieurement, dans le cadre de l'analyse fréquentielle en contexte non stationnaire.

SYNTHESE DU CHAPITRE 3

L'objet de ce chapitre était la mise en place d'un cadre méthodologique permettant d'évaluer la stationnarité de plusieurs descripteurs hydrologiques sur un grand nombre de stations, à la fois à l'échelle locale et à l'échelle régionale.

Au niveau local, une revue bibliographique couplée à une étude par simulation a permis de mettre en avant les principaux facteurs influant sur le choix d'un test statistique. Certains points ont pu être clarifiés au cours de cette étude, comme le calcul des valeurs critiques pour un test de rupture à date inconnue basé sur le rapport des vraisemblances maximales, ou encore la mise en place de tests permettant de prendre en compte l'autocorrélation. Finalement, un schéma de sélection d'un test de stationnarité a été établi sur la base de ces résultats. Ce schéma a permis de choisir le test adapté à chacune des variables échantillonnées au Chapitre 2.

A l'échelle régionale, la première question traitée a été celle de la significativité régionale du nombre de changements localement significatifs. Trois méthodes ont été évaluées : la première est basée sur le concept d'effectif efficace, la seconde utilise une procédure bootstrap, et la dernière est issue de l'utilisation d'une copule Gaussienne. Des jeux de données multivariés ont été simulés afin de vérifier que ces méthodes respectaient bien le risque de première espèce. L'utilisation de l'effectif efficace s'est avéré inappropriée, puisqu'elle conduit à fortement surestimer les valeurs critiques. Les deux autres méthodes ont paru plus acceptables sur la base des simulations que nous avons effectuées, qui ne reflètent malheureusement pas toute la diversité des structures de dépendances spatiales que l'on peut rencontrer dans la réalité. La procédure basée sur le bootstrap est cependant apparue plus robuste et plus simple à mettre en œuvre.

Le second aspect analysé a concerné la cohérence spatiale des évolutions. Dans l'optique de la mise en évidence d'impacts du changement climatique, il est naturel de rechercher des tendances cohérentes au sein d'une région hydro-climatique homogène. Des approches basées sur la création de variables ou de statistiques régionales ont ainsi été évaluées. Bien que non biaisées, ces techniques ne prennent pas en compte la redondance d'information liée à la dépendance spatiale, de sorte que des changements non cohérents peuvent également être détectés. Pour remédier à ce problème, un test semi-paramétrique a été mis au point, et s'est révélé, sur la base de quelques simulations multivariées, à la fois non biaisé et capable de distinguer un changement cohérent d'un changement incohérent. Enfin, des approches totalement paramétriques peuvent être envisagées, mais les difficultés numériques afférentes limitent leur utilisation.

CHAPITRE 4. ANALYSE DE STATIONNARITE EN FRANCE METROPOLITAINE

« Depuis quelques années, nous sommes témoins de refroidissements sensibles dans l'atmosphère, de variations subites dans les saisons et d'ouragans ou d'inondations extraordinaires auxquels la France semble devenir de plus en plus sujette ».
Circulaire n° 18 du ministère de l'intérieur du 25 avril 1821, cité par Bainville et Ladoy (1995).

Les données et les méthodes nécessaires à l'analyse de stationnarité des extrêmes hydrologiques en France ont été décrites dans les chapitres précédents. Nous allons à présent en présenter les principaux résultats. Encore une fois, un important travail de critique sera effectué, afin de ne pas déconnecter les résultats statistiques de la réalité physique des phénomènes étudiés. Ce chapitre est organisé en cinq parties. Tout d'abord, nous commentons en détail quelques exemples qui nous permettront d'une part d'expliquer la mise en forme des résultats et d'autre part de vérifier que les tests sont aptes à détecter des changements avérés. Dans une seconde partie, nous présentons les résultats de l'analyse sur l'ensemble des stations. Ces résultats seront ensuite critiqués, ce qui nous conduira à définir un nouvel échantillon réduit de stations, pour lequel les ruptures observées n'ont pu être reliées à des problèmes météorologiques ou à des influences directes. La stationnarité sera ensuite évaluée à l'échelle régionale, à la fois en terme de significativité et de cohérence. Enfin, nous aborderons le problème de l'attribution d'un changement à une cause climatique, en mettant en parallèle les débits et des variables physiques explicatives (pluies et températures), à la fois de manière empirique et en utilisant la modélisation hydrologique.

I. Deux exemples commentés

I.1. Mise en forme et diffusion des résultats

Un site Internet a été créé afin de diffuser les résultats de l'étude de stationnarité auprès notamment des services responsables de l'acquisition des données. Ceci nous a permis de dialoguer plus efficacement avec les gestionnaires, et d'engager avec eux un travail de critique des résultats, qui sera détaillé ultérieurement.

<http://www.lyon.cemagref.fr/hh/PNRH-NS/>

Pour chaque station, les résultats ont été regroupés dans des « fiches-stations ». A l'intérieur de chacune de ces fiches, des vignettes représentent le résultat des tests de stationnarité pour quelques variables (voir Figure 4-1). Les valeurs de la variable y sont tracées chronologiquement, ainsi que la tendance et la rupture estimées. En cas de changement significatif, un cadre coloré apparaît, et indique la significativité du changement (1%, 5% ou 10%), ainsi que la date de rupture estimée.

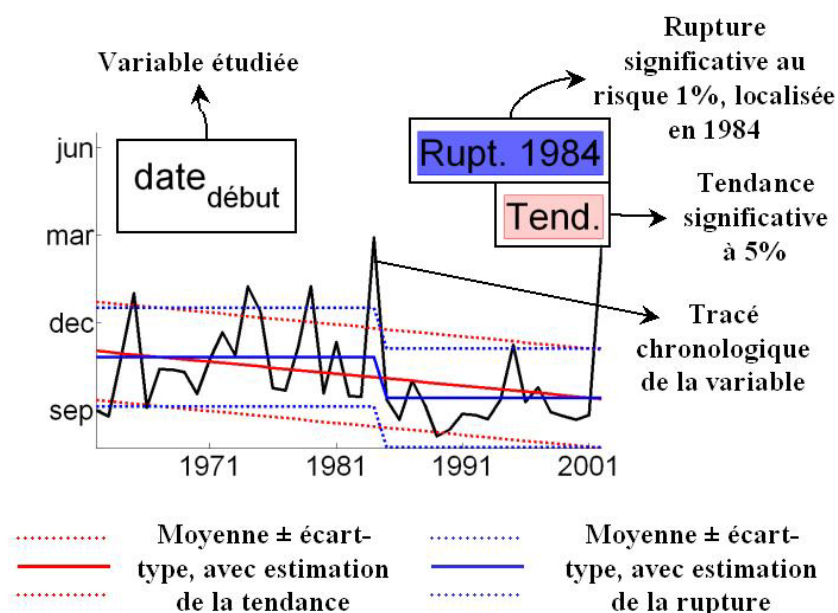


Figure 4-1. Vignette résumant les changements détectés sur une variable

La Figure 4-2 est un exemple de fiche-station pour les variables de crue. En haut de cette fiche figurent des informations générales sur le bassin versant (superficie, altitude, année hydrologique, saisonnalisation éventuelle, durée caractéristique de crue), ainsi que la légende des vignettes. Les variables illustrées concernent les pointes de crue (valeur et date) issues des échantillonnages MAXAN et SUPSEUIL, les descripteurs d'hydrogrammes (durée, volume, forme), ainsi que des indicateurs de fréquence (nombre annuel d'événements et intervalles inter-occurrences pour des seuils conduisant à 0.5, 1 et 2 événements par an en moyenne). Pour les bassins à influence pluvio-nivale, deux autres fiches du même type sont créées pour décrire chacune des deux saisons. Il existe enfin une fiche spécifique relative aux variables descriptives des ondes de fonte nivale (Figure 4-3).

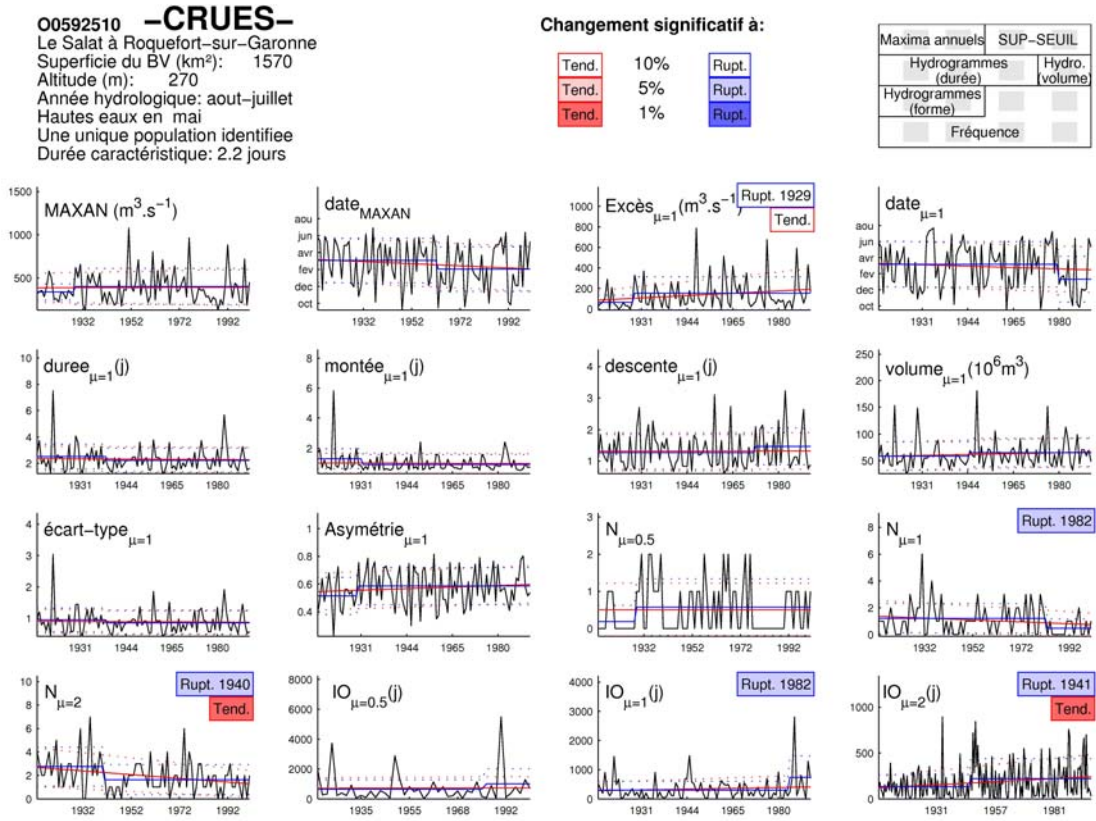


Figure 4-2. Exemple de fiche-station pour les variables de crue.



Figure 4-3. Exemple de fiche-station pour la description des crues nivales.

Les variables d'étiage sont résumées dans la Figure 4-4. Les variables décrites sont la durée et le déficit de volume des basses eaux, ainsi que les dates de début et de fin d'étiage. Les minima annuels des débits journaliers, ainsi que des débits moyens et seuils sur 7 et 30 jours y sont également représentés. Enfin, le module annuel et la distance au régime inter-annuel complètent cette description.

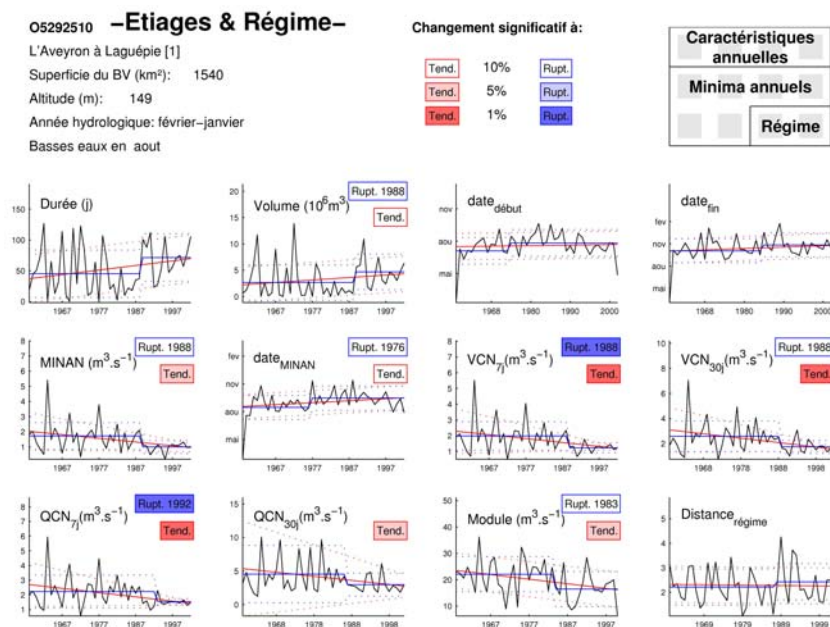


Figure 4-4. Exemple de fiche-station pour les variables d'étiage.

I.2. La Seine à Paris-Austerlitz

Nous avons volontairement inclus dans l'analyse quelques stations influencées, afin d'évaluer le comportement des tests dans des conditions d'application réelles. A titre d'illustration, les résultats concernant le bassin de la Seine à Paris (43800 km²) sont présentés. De nombreux aménagements ont été construits sur ce bassin, suite aux inondations de 1910 et 1924, et aux sécheresses des années 20 (<http://www.iibrbs.fr/>). Les quatre grands lacs-réservoirs mis en eaux à partir de 1949 ont un double rôle de soutien d'étiage et d'écroulement des crues. Il est donc intéressant d'étudier comment cette influence se traduit sur la distribution des variables décrivant les hautes et les basses eaux.

Cette chronique est remarquable par son ancienneté, puisque les données analysées vont de 1876 à 2003. La Figure 4-5 décrit les résultats en étiage. Certaines variables présentant une autocorrélation significative, elles n'ont été traitées qu'avec un test de tendance (Mann-Kendall modifié). La quasi-totalité des variables d'étiages a connu d'importants changements, à l'exception de la date du minimum annuel journalier. Tous ces changements vont dans le sens d'étiages moins sévères, ils reflètent donc l'influence positive des lacs réservoirs du bassin de la Seine. Les dates de ruptures estimées sont d'ailleurs à comparer avec les dates de mise en service des aménagements : rupture en 1948 pour la date de fin d'étiage (lac-réservoir de Pannecière mis en service en 1949), et ruptures en 1973, 1976 et 1979 pour la date de début d'étiage, le VCN_{30j} et le QCN_{30j} (lac-réservoir Marne mis en service en 1974). A l'inverse, le module ne montre aucun changement particulier, ce qui est cohérent, car à l'échelle de l'année, les lacs-réservoirs ne prélèvent ni n'injectent d'eau au bassin.

La Figure 4-6 décrit les résultats pour les variables de crue. Le changement le plus important concerne les variables relatives à la fréquence des crues : le temps d'attente entre deux événements est en augmentation pour les événements courants (correspondant à $\mu=1$ ou

2). Deux tendances faiblement significatives (10%) sont également détectées sur les variables relatives à la forme des hydrogrammes de crue : la durée et le volume des hydrogrammes sont à la hausse. On peut penser que ces changements reflètent les aménagements successifs qu'a connu le bassin versant. L'augmentation des durées et des volumes serait ainsi due à l'écrêtement des crues (crues plus « aplaties »), et l'action des barrages conduirait à une diminution de la fréquence des événements, au moins pour les crues courantes. En terme de distribution, l'effet des lacs réservoirs est cependant beaucoup moins marqué sur les hautes eaux.

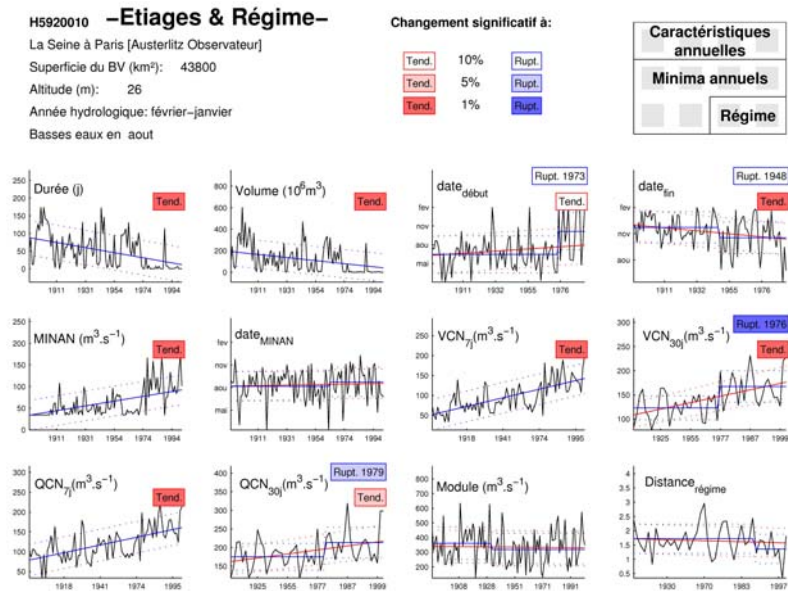


Figure 4-5. Résultats des tests de stationnarité en étiage pour la Seine à Paris-Austerlitz.

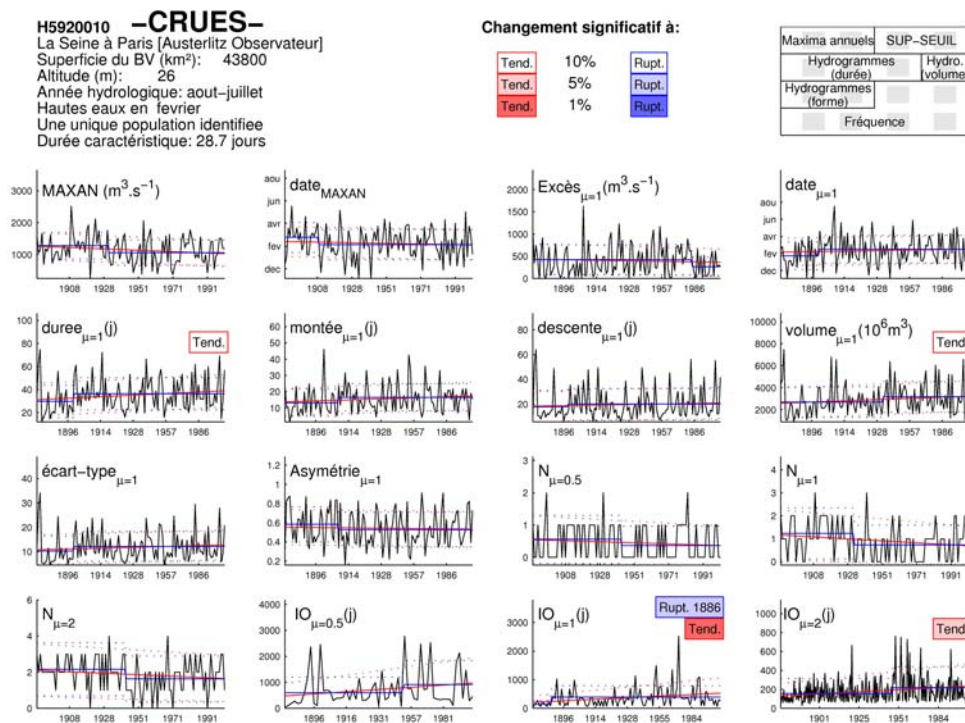


Figure 4-6. Résultats des tests de stationnarité en crue pour la Seine à Paris-Austerlitz.

I.3. La Neste de Cap de Long à Aragnouet

Cette station possède le plus petit bassin versant de l'échantillon analysé (5 km²). Il s'agit d'un petit cours d'eau d'altitude, au régime pluvio-nival non influencé. Les résultats pour les variables de crue non saisonnalisées sont décrits dans la Figure 4-7. Les variables saisonnalisées présentent le même type de changements. On observe sur cette station un changement marqué sur le régime des crues, daté vers 1957 pour l'intensité (MAXAN et SUP-SEUIL) et vers 1961 pour la fréquence. Tous ces changements vont dans le sens de crues moins fréquentes et moins intenses.

Ce type de résultat est représentatif de changements difficilement imputables au changement climatique : une rupture aussi marquée a très certainement été causée par un problème météorologique, un déplacement de station, ou la mise en eau d'un aménagement que nous n'avions pas relevée *a priori*. Après discussion avec les gestionnaires et consultation de l'historique de la station de mesure, il s'avère que l'échelle limnimétrique a été déplacée en septembre 1961, et un déversoir a été construit. La courbe de tarage a donc été modifiée mais l'extrapolation de la courbe de tarage vers les forts débits semble avoir été revue à la baisse. Avant 1961, l'extrapolation des débits commençait à partir de 1.5 m³.s⁻¹, alors que pour le barème postérieur à la construction du déversoir, l'extrapolation de la courbe commence vers 2.2 m³.s⁻¹. Ce changement d'extrapolation explique le changement observé sur la chronique, la surestimation des forts débits avant 1961 créant une tendance apparente à la baisse des crues.

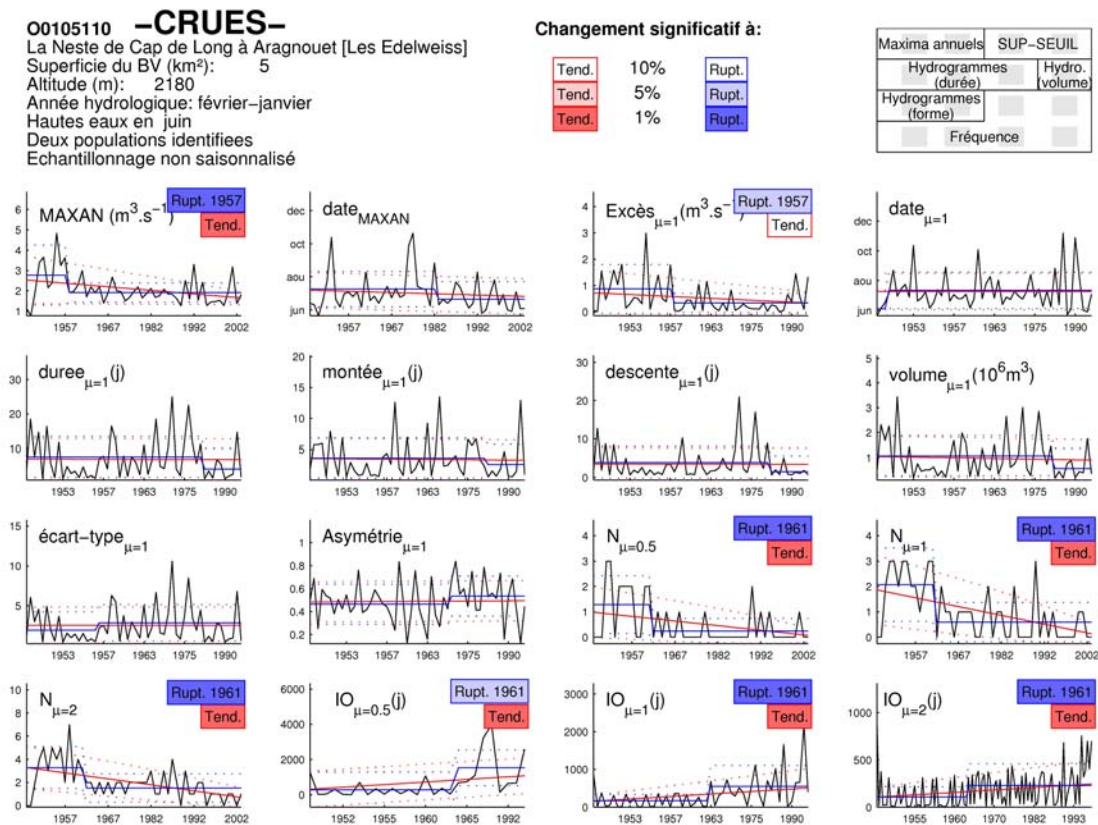


Figure 4-7. Résultats des tests de stationnarité en crue pour la Neste de Cap de Long à Aragnouet.

Ce type de résultat, qui ne représente pas un cas isolé, démontre l'importance d'une nouvelle étape de critique des données, afin de tenter d'expliquer les ruptures détectées. Dans l'optique de l'attribution des évolutions au changement climatique, il est en effet important d'écarter ce type de stations, pour lesquelles les changements n'ont rien de climatique. En conséquence, l'analyse de stationnarité sera conduite de la manière suivante : tout d'abord, elle portera sur l'ensemble des stations réputées non influencées. Cette première étape permettra de créer les « fiches stations », qui seront ensuite étudiées une par une, afin de rechercher des causes d'évolutions non climatiques. Dans un second temps, le jeu de données épuré sera analysé à l'échelle régionale, et nous tenterons de lier les changements observés sur les débits au comportement de variables explicatives (les pluies et les températures), à la fois de manière empirique, et en utilisant la modélisation pluie-débit.

Cette démarche pour tenter d'attribuer d'éventuelles évolutions à une cause climatique est en grande partie empirique. Des approches plus quantitatives, basées sur la méthode du *fingerprint*, existent et ont été appliquées sur les températures (Hasselmann, 1993, 1998; Barnett *et al.*, 1999; Spagnoli *et al.*, 2002; Zwiers et Zhang, 2003; Tank *et al.*, 2005). Cependant, ce type d'approche n'est pas utilisable dans le contexte des données de débit que nous utilisons : le principe général est en effet de comparer l'évolution observée à celle prédite par un modèle de circulation générale. Malheureusement, la précision des projections des GCM n'est pas suffisante à l'échelle de la plupart des bassins versants analysés. Des procédures de désagrégation spatiale peuvent être utilisées, mais sont elles-mêmes sujettes à de fortes incertitudes.

II. Analyse locale

Comme expliqué précédemment, l'analyse locale a été effectuée en trois étapes, qui sont détaillées dans cette section : une première analyse concerne l'ensemble des stations réputées non influencées ; puis une critique des résultats obtenus est effectuée station par station ; enfin une nouvelle analyse est restreinte aux stations pour lesquelles aucune explication aux changements observés n'a pu être trouvée.

II.1. Première analyse sur l'échantillon complet de stations

II.1.1. Résultats en crues

179 stations hydrométriques réputées non influencées en hautes eaux ont été utilisées. Pour chaque variable, le test statistique est choisi en appliquant le schéma de sélection décrit au Chapitre 3. Nous privilégions, dans l'optique de détecter des changements liés au climat, les tests de tendance aux tests de rupture. Concernant les variables échantillonnées par valeurs maximales annuelles, seules les années présentant moins de 10% de lacunes ont été conservées. Pour les variables échantillonnées par sélection d'hydrogrammes ou de pointes supérieures à un seuil, seuls les événements exempts de lacunes ont été conservés.

Les résultats concernant quatre variables descriptives des crues sont détaillés dans la Figure 4-8. On peut observer qu'un nombre important de changements est détecté, certains étant nettement significatifs. Les maxima annuels et le nombre annuel d'événements semblent présenter plus de changements que les pointes supérieures à un seuil ou que les dates d'événement. De plus, il est difficile de dégager une réelle cohérence spatiale dans ces résultats, même si certaines tendances semblent se regrouper (par exemple, à l'ouest du Massif Central pour la variable N). Le déterminisme climatique des changements observés est donc douteux. Une analyse détaillée station par station confirme visuellement que dans de nombreux cas, certaines ruptures très marquées apparaissent, ce qui pourrait traduire des problèmes météorologiques ou des influences anthropiques directes.

Le Tableau 4-1 présente de manière plus exhaustive les résultats pour l'ensemble des variables considérées. De manière générale, on peut observer que pour de nombreuses variables, environ 25% des séries sont significativement non stationnaires au risque 10%. La seule exception concerne les variables décrivant la saisonnalité des crues, qui sont en général moins soumises au changement. Ce pourcentage de 25% semble *a priori* important, mais il convient de garder à l'esprit qu'en toute rigueur, seule l'évaluation de sa significativité régionale permet de garantir que tous ces changements ne sont pas dus au hasard. Dans un premier temps, nous n'avons pas évalué cette caractéristique régionale sur ce jeu de données, car une rapide inspection des résultats sur quelques stations suffit à se convaincre que certains changements ne doivent rien au hasard (cf. Figure 4-7). Cependant, nous signalons en gras dans le Tableau 4-1 les résultats régionalement significatifs sous hypothèse d'indépendance spatiale des données. Cette hypothèse est évidemment abusive, mais permet de fournir aisément une grille de comparaison homogène pour les trois risques locaux de 0.01, 0.05 et 0.1, en choisissant un risque régional identique que nous prenons égal à 0.05. En utilisant l'approximation binomiale détaillée dans le Chapitre 3, il faudrait au moins 2.8%, 8.4% et 14.5% de changements significatifs aux risques locaux 0.01, 0.05 et 0.1 pour garantir la significativité régionale, si les stations étaient indépendantes. Globalement, on observe que quel que soit le risque adopté, la conclusion de l'analyse est similaire, c'est à dire que la conclusion quant à la significativité régionale d'une variable est généralement identique pour les trois risques.

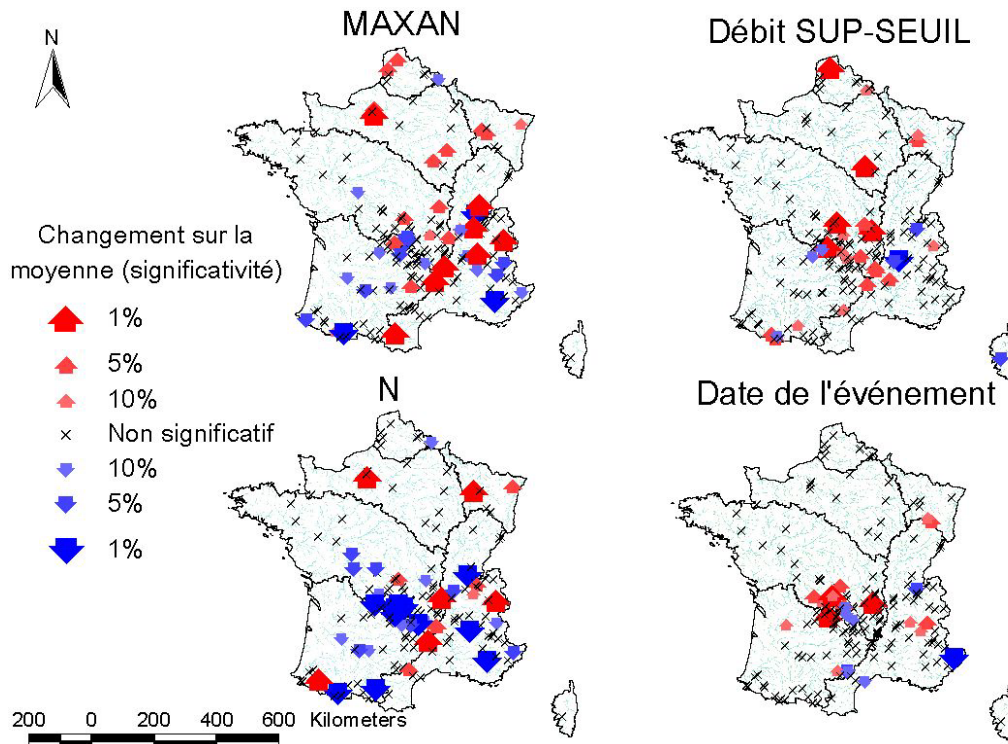


Figure 4-8. Résultat des tests de stationnarité pour quatre variables de crue: MAXAN (maximum annuel des débits journaliers), Débit SUP-SEUIL (pointe de débit journalier supérieure à un seuil), N (nombre annuel d'événements supérieurs au seuil) et date des événements supérieurs au seuil. Le seuil est choisi pour obtenir un événement par an en moyenne.

Echantillonnage	Variable	% de résultats significatifs au risque		
		1.0%	5.0%	10.0%
MAXAN	$(VCX_{1i})_A$	6.1	18.4	26.8
	θ_1	0.6	4.5	7.8
	$(VCX_d)_A$	8.4	19.0	26.3
	θ_1	0.0	3.4	7.3
	$(QCX_d)_A$	5.6	12.8	25.7
	θ_2	0.6	3.4	7.3
SUP-SEUIL	$N_3 (\mu=0.5)$	6.7	17.9	24.0
	$N_3 (\mu=1)$	9.5	17.3	22.3
	$N_3 (\mu=2)$	10.1	18.4	26.8
	$(VCX_{1i})_{\mu=1}$	3.4	10.1	16.2
	θ_3	2.2	6.1	12.8
	$(VCX_d)_{i=1}$	5.0	12.8	24.0
	θ_3	2.8	6.7	13.4
	$(QCX_d)_{i=1}$	4.5	15.1	22.3
	θ_4	2.8	8.9	14.5
	$IO (\mu=0.5)$	3.9	11.2	15.6
	$IO (\mu=1)$	6.7	15.1	20.7
$IO (\mu=2)$	8.9	19.6	25.1	
Sélection d'hydrogrammes	δ	7.3	15.1	23.5
	m	2.8	9.5	13.4
	r	6.1	19.0	26.8
	σ	6.7	17.3	24.0
	γ	1.7	8.4	14.5
	ν	5.6	14.5	20.7

Tableau 4-1. Résultats des tests de stationnarité pour les variables de crue. Les chiffres en gras représentent les résultats significatifs au niveau régional (France entière, risque régional de 5%), sous hypothèse d'indépendance spatiale.

II.1.2. Résultats en étiages

Seules 128 stations réputées non influencées ont été utilisées pour l'évaluation des basses eaux. Comme précédemment, seules les années présentant moins de 10% de lacunes ont été conservées. Le schéma de sélection d'un test a été appliqué à chaque variable, en utilisant les tests de tendance. Rappelons que pour les variables d'étiages, la présence d'autocorrélation est préalablement testée, et le cas échéant, un test adapté est appliqué (Mann-Kendall modifié).

Les résultats sont présentés dans la Figure 4-9, et sont comparables à ceux observés sur les hautes eaux : les changements semblent très nombreux, et la cohérence spatiale est assez faible. Plus quantitativement, on observe que plus d'un tiers des stations présentent un changement significatif à 10% pour les variables décrivant les minima annuels. Ce pourcentage est même de 44.5% au pas de temps journalier ! Les variables décrivant la saisonnalité, le déficit de volume ou la durée des étiages sont moins sensibles, mais le pourcentage de résultats significatifs reste élevé. Les trois risques conduisent encore une fois à des conclusions similaires, à l'exception des variables « date de fin d'étiage » et « déficit de volume ».

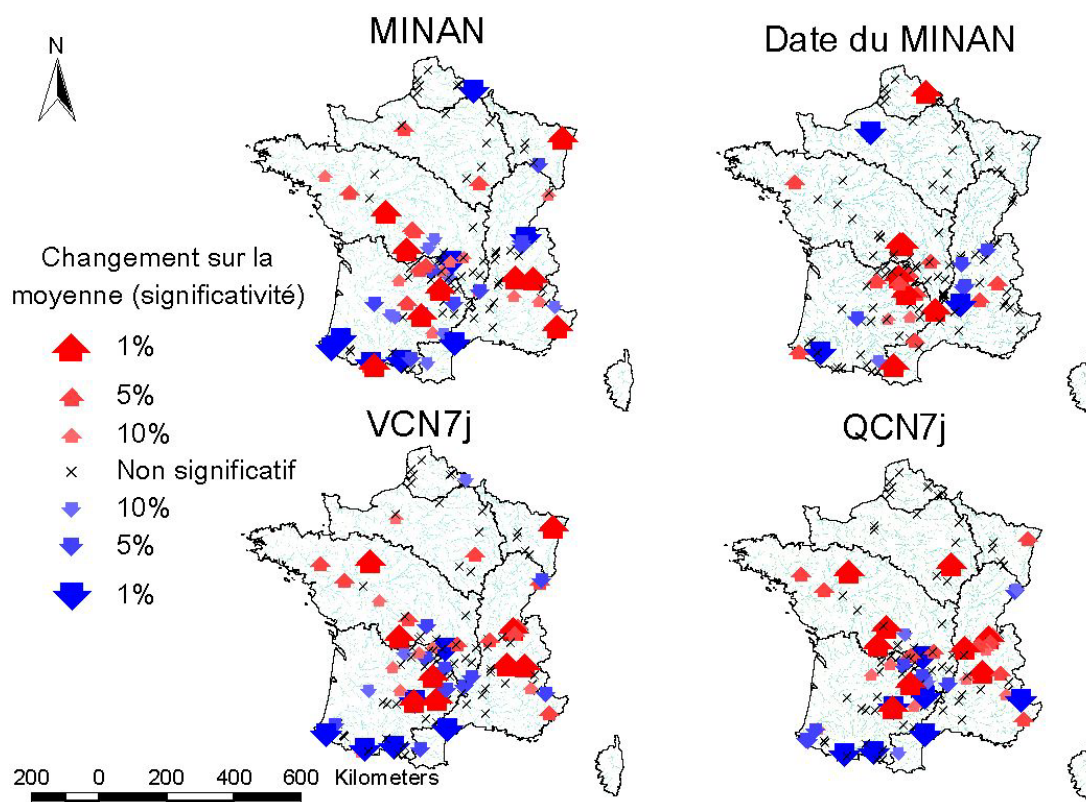


Figure 4-9. Résultat des tests de stationnarité pour quatre variables de basses eaux: MINAN (minima annuels des débits journaliers), Date du MINAN, VCN_{7j} (minima annuels des débits moyens sur 7 jours) et QCN_{7j} (minima annuels des débits seuils sur 7 jours).

Variable	% de résultats significatifs au risque		
	1.0%	5.0%	10.0%
$(VCN_{1j})_A$	13.3	28.1	44.5
θ_1	6.2	19.0	24.1
$(VCN_{7j})_A$	11.7	28.1	37.5
$(VCN_{30j})_A$	13.3	25.0	35.2
$(QCN_{7j})_A$	12.5	25.0	35.2
$(QCN_{30j})_A$	13.3	25.0	33.6
δ	7.0	11.7	17.2
v	6.3	11.7	15.6
θ_{deb}	6.3	17.2	20.3
θ_{fin}	3.9	8.6	18.8

Tableau 4-2. Résultats des tests de stationnarité pour les variables d'étiage. Les chiffres en gras représentent les résultats significatifs au niveau régional (France entière, risque régional de 5%), sous hypothèse d'indépendance spatiale.

II.1.3. Résultats en régime

Les variables « module » et « distance au régime inter-annuel » ont été évaluées pour l'ensemble des 195 stations, dans les mêmes conditions que celles décrites pour les étiages. Le nombre de changements semble cette fois plus modéré (respectivement 16.9% et 18.5% au risque 0.1 pour les modules et les distances, cf. Tableau 4-3). Dans le cas des modules, on note une meilleure cohérence spatiale, les tendances à la baisse étant plutôt localisées dans le sud du pays (Figure 4-10). Concernant les distances au régime inter-annuel, nous avons seulement indiqué la significativité des changements, puisque le sens d'un changement en distance n'a pas réellement d'interprétation hydrologique simple. On remarque cependant que la majorité des changements observés est située dans les Alpes.

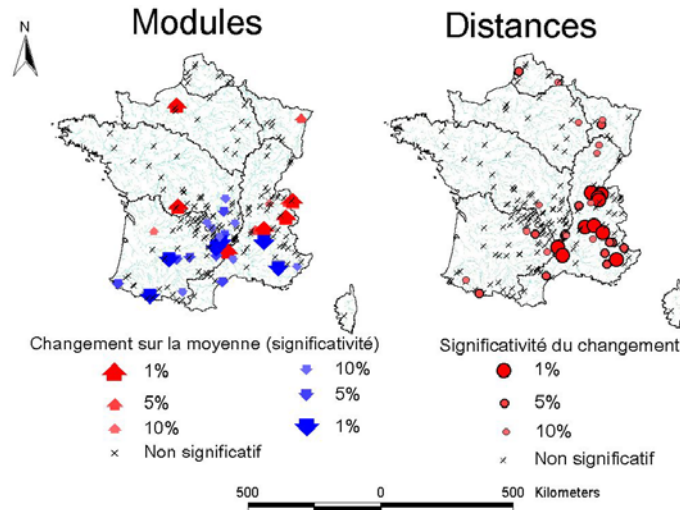


Figure 4-10. Résultat des tests de stationnarité pour les modules et les distances au régime inter-annuel.

Variable	% de résultats significatifs au risque		
	1.0%	5.0%	10.0%
Module	6.2	11.3	16.9
Distance au régime	5.1	11.8	18.5

Tableau 4-3. Résultats des tests de stationnarité pour les variables de régime. Les chiffres en gras représentent les résultats significatifs au niveau régional (France entière, risque régional de 5%), sous hypothèse d'indépendance spatiale.

II.2. Critique des résultats

Les résultats détaillés précédemment semblent indiquer que la stationnarité des séries hydrométriques en France n'est pas une hypothèse raisonnable pour de nombreuses stations. Cependant, la cause des changements observés reste à déterminer, et rien ne prouve à ce stade de l'analyse que la variabilité climatique soit impliquée. Dans un premier temps, nous avons donc souhaité éliminer les stations pour lesquelles les changements sont d'origine anthropique directe (problème métrologique, influence directe, etc.).

II.2.1. Organisation de la critique

Sur l'ensemble des stations hydrométriques étudiées, rares sont celles ne présentant aucun changement douteux sur aucune des variables analysées. Plus précisément, sur l'ensemble des 195 stations initiales, 87 stations doivent être critiquées en hautes eaux, 100 en basses eaux et 67 en régime. Dans la pratique, cette évaluation a été basée sur les fiches-stations (cf. Figure 4-2). Les tests de rupture constituent ici une aide précieuse, car ils permettent de dater approximativement les changements, et donc de préciser la période vers laquelle il faut rechercher un événement particulier à incriminer. La critique a été effectuée directement par les gestionnaires, dans le cas où le nombre de station à étudier n'était pas très important (DIREN Aquitaine, Bourgogne, Bretagne, Pays de la Loire, Alsace, Haute-Normandie, Limousin, Franche-Comté, Lorraine). Dans les autres cas, nous sommes allés rendre visite aux gestionnaires et avons effectué des recherches dans les archives (EDF-DTG Grenoble, Brive-la-Gaillarde et Toulouse, CNR, DIREN Centre – antennes du Puy et d'Orléans, Nord-Pas-De-Calais, Midi-Pyrénées, Languedoc-Roussillon, DDE Dordogne). Notons que la réorganisation actuelle des services d'hydrométrie ne nous a pas facilité la tâche : certains dossiers ont effectivement été transférés d'une DIREN à une autre entre le début et la fin de la thèse, et quelques stations n'ont finalement pas été évaluées pour cette raison. Elles ont donc été retirées de l'analyse.

II.2.2. Causes supposées des ruptures observées

Les conclusions de cette critique des résultats sont résumées dans le Tableau 4-4 pour les crues, le Tableau 4-5 pour les étiages et le Tableau 4-6 pour les variables descriptives du régime. Il est bien entendu quasiment impossible de relier avec certitude un changement observé sur les séries avec un événement particulier. Les causes supposées sont essentiellement justifiées par le fait qu'elles sont concomitantes et cohérentes avec les ruptures observées.

Cause des changements observés		%	Commentaires	%
Causalité non climatique	Hydrométrie	40.2	Changement de matériel	1.1
			Modification de la section de mesure	1.1
			Déplacement de l'échelle	3.4
			Changement de la méthode de calcul	4.6
			Changement de gestionnaire	4.6
			Déplacement de la station	5.7
			Qualité trop incertaine	8.0
			Changement de barème	11.5
	Influence	9.2	/	/
Causalité non déterminée	Manque d'information	20.7	Non évalué	4.6
			Pas d'informations (dossier incomplet)	16.1
	Informations disponibles, mais changement non expliqué	29.9	Cause climatique possible, analyse à poursuivre	/

Tableau 4-4. Causes supposées des changements observés en hautes eaux sur 87 stations hydrométriques.

Cause des changements observés		%	Commentaires	%
Causalité non climatique	Hydrométrie	48.0	Changement de matériel	1.0
			Modification de la section de mesure	2.0
			Déplacement de l'échelle	2.0
			Changement de la méthode de calcul	4.0
			Changement de gestionnaire	4.0
			Déplacement de la station	6.0
			Changement de barème	8.0
			Qualité trop incertaine	21.0
Influence	23.0	/	/	
Causalité non déterminée	Manque d'information	13.0	Non évalué	4.0
			Pas d'informations (dossier incomplet)	9.0
	Informations disponibles, mais changement non expliqué	16.0	Cause climatique possible, analyse à poursuivre	/

Tableau 4-5. Causes supposées des changements observés en basses eaux sur 100 stations hydrométriques.

Cause des changements observés		%	Commentaires	%
Causalité non climatique	Hydrométrie	46.3	Changement de matériel	1.5
			Modification de la section de mesure	1.5
			Déplacement de l'échelle	1.5
			Changement de la méthode de calcul	4.5
			Changement de gestionnaire	4.5
			Changement de barème	6.0
			Déplacement de la station	9.0
			Qualité trop incertaine	17.9
Influence	19.4	/	/	
Causalité non déterminée	Manque d'information	19.4	Non évalué	6.0
			Pas d'informations (dossier incomplet)	13.4
	Informations disponibles, mais changement non expliqué	14.9	Cause climatique possible, analyse à poursuivre	/

Tableau 4-6. Causes supposées des changements observés en régime sur 67 stations hydrométriques.

Le premier enseignement est que dans 40 à 48% des cas, les ruptures observées sont liées à la métrologie, ce qui témoigne de la réelle difficulté d'estimer des débits valables sur le long terme. La courbe de tarage est l'élément le plus sensible dans la chaîne de traitement menant au calcul des débits : ainsi, des changements de gestionnaires, des déplacements ou des modifications de la section de mesure se traduisent souvent par un changement de barème. Or, plusieurs années sont nécessaires pour effectuer des jaugeages sur une gamme satisfaisante de débits. Un problème spécifique concerne également l'extrapolation de la relation hauteur/débit. Par exemple, suite à un jaugeage en crue, il n'est pas rare de modifier la partie haute de la courbe pour prendre en compte ce point supplémentaire. Mais faut-il également recalculer les débits passés ? Cette question ne possède malheureusement pas de réponse simple. Il est en effet probable que l'extrapolation des débits avant le jaugeage soit erronée, mais le nouveau barème est-il pour autant applicable à toutes les crues observées antérieurement ? Il est ainsi difficile de savoir jusqu'où remonter dans le passé pour réévaluer ces débits. Ces changements de barème spécifiques aux parties extrêmes de la courbe de tarage sont les plus problématiques, et doivent être distingués des détarages liés à une modification de la section de mesure, qui sont justifiés et ne provoquent généralement pas de ruptures sur les débits couramment observés.

D'autres problèmes métrologiques indépendants de la courbe de tarage ont pu être identifiés comme causes de ruptures. Tout d'abord, les changements de méthode de calcul des

débites, qui concernent principalement les débits naturels reconstitués. Certains lissages, par corrélation avec des stations proches notamment, ont pu être introduits dans la reconstitution. Nous avons également inclus dans cette catégorie les ruptures liées au passage de données journalières « observateurs », relevées à heure fixe, à des données journalières calculées en moyennant les observations sur la journée. Cette différence, qui concerne essentiellement les stations possédant plus de 40 années de données, est minime pour les grands cours d'eau et les rivières lentes, dont le débit varie faiblement au cours d'une journée, mais peut créer des différences significatives pour des bassins au temps de réaction plus court. Enfin, les chroniques de « qualité trop incertaine » concernent principalement des débits calculés au droit d'aménagements en utilisant une conversion puissance/débit et des variations de réserve. Ces calculs se sont révélés particulièrement problématiques pour les bas débits alors qu'ils semblaient souvent acceptables en crue.

Les influences constituent également une part non négligeable des explications trouvées aux ruptures détectées, particulièrement en basses eaux (23%) et en module (19.4%). Inversement, de part les quantités d'eau impliquées, les crues sont plus rarement influencées, sauf dans le cas de grands aménagements qui sont en général signalés dans la banque HYDRO et que nous avons donc relevés.

Dans un nombre de cas non négligeable (entre 13% et 20.7%), nous n'avons pas trouvé d'informations permettant de critiquer les résultats. Comme indiqué précédemment, quelques stations n'ont pas été évaluées, suite notamment à la réorganisation des services d'hydrométrie. D'autre part, certains documents anciens (en général, antérieurs au début des années 1960) n'ont pas été retrouvés dans les dossiers, ce qui a rendu impossible la critique de ruptures survenues avant cette période. Nous avons donc décidé de retirer ces stations de l'analyse.

Enfin, certains changements n'ont pu être reliés à aucun événement particulier, et les stations ont donc été conservées dans l'analyse. Ainsi, environ 30% des changements observés en crue n'ont pu être expliqués, ce pourcentage n'étant que d'environ 15% pour les étiages et les modules.

II.3. Résultats sur les jeux réduits de stations

A l'issue de l'étape de critique des résultats, nous disposons d'échantillons de stations plus réduits, mais auxquels nous pouvons accorder une plus grande confiance. Bien que l'absence d'explication métrologique ou anthropique à un changement observé ne constitue en aucun cas une preuve de son déterminisme climatique, les données que nous allons à présent utiliser sont plus propices à une analyse climatologique. Le nombre de stations finalement utilisables est de 124 pour les hautes eaux, 25 pour la description des crues nivales, 90 pour les basses eaux et 130 pour la description du régime. Signalons que certaines stations pour lesquelles une rupture nette a pu être reliée à un événement particulier ont été conservées dans l'analyse : nous avons alors uniquement considéré les années situées avant ou après cet événement, dans le cas où le nombre d'années résultant restait supérieur à 40 ans.

II.3.1. Crues

Les résultats concernant les quatre variables descriptives des crues précédemment étudiées sont décrits dans la Figure 4-11. Le nombre de stations non stationnaires a nettement diminué, mais certains changements nettement significatifs demeurent (hausse des maxima annuels par exemple). Encore une fois, la cohérence spatiale semble limitée, même si les changements vers une baisse du risque de crue semblent plutôt se regrouper dans la moitié ouest du pays. La Figure 4-12 concerne les variables descriptives des hydrogrammes de crue. On peut notamment remarquer un regroupement de tendances vers des crues plus rapides, le long

d'une ligne Narbonne-Toulouse. Ces stations contrôlent des bassins versants répartis de part et d'autre de la ligne de partage des eaux Atlantique-Méditerranée. Les pluies occasionnant les crues peuvent donc être soit de type océanique, soit de type méditerranéen. Or, nous avons remarqué sur trois de ces stations que les forts événements tendaient à être de plus en plus souvent occasionnés par des pluies méditerranéennes, conduisant à des crues plus rapides. Des analyses complémentaires, basées sur les précipitations, seraient intéressantes pour confirmer ce phénomène. Dans une moindre mesure, signalons également une possible augmentation des volumes dans le nord-est de la France.

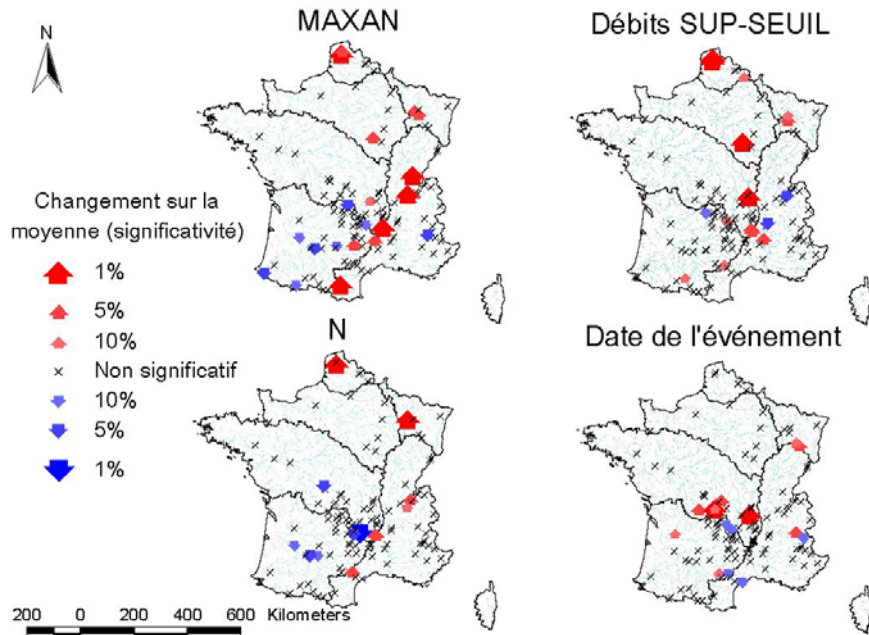


Figure 4-11. Résultat des tests de stationnarité pour quatre variables de crue: MAXAN, Débit SUP-SEUIL, N (nombre annuel d'événements supérieurs au seuil) et date des événements supérieurs au seuil (un événement par an en moyenne).

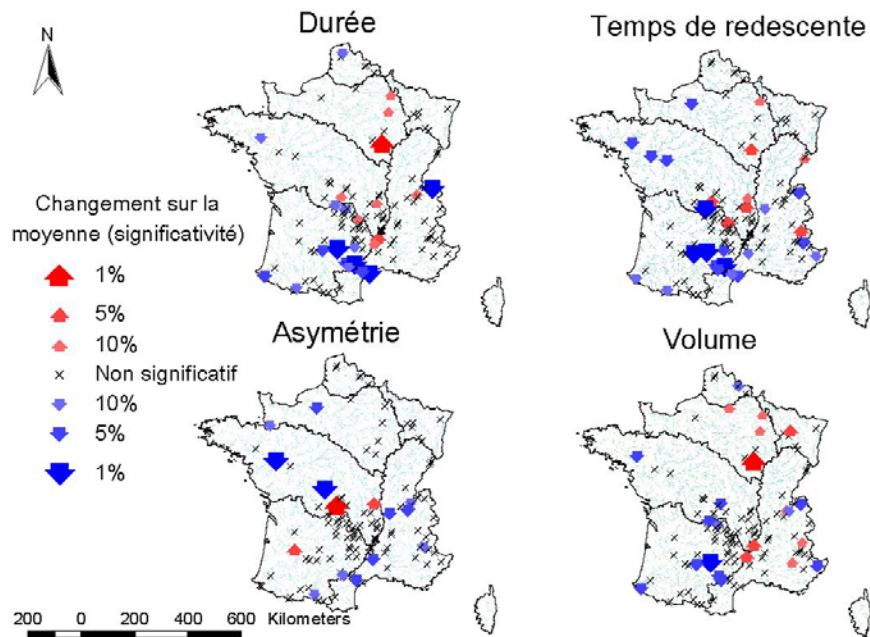


Figure 4-12. Résultat des tests de stationnarité pour quatre variables descriptives des hydrogrammes de crue (avec en moyenne un événement par an): Durée de dépassement de la moitié du pic, temps de redescente, asymétrie et volume généré.

Le Tableau 4-7 résume les résultats pour l'ensemble des variables étudiées. Le pourcentage de résultats localement significatifs a très nettement baissé après l'étape de critique des résultats. De plus, la pseudo-significativité régionale, évaluée sous hypothèse d'indépendance spatiale, semble plus souvent atteinte lorsque le risque local augmente. Ceci traduit le fait que les ruptures les plus significatives ont généralement été expliquées lors de l'étape de critique. A l'inverse, les changements non expliqués concernent majoritairement des évolutions modérées. Au risque 10%, les variables décrivant l'intensité des crues présentent des évolutions significatives pour 12.9% à 17.7% des stations. La fréquence des événements semble moins sujette au changement, sauf pour les crues sélectionnées avec un seuil conduisant à deux événements par an en moyenne. Enfin, les variables décrivant la forme des hydrogrammes semblent être les plus sensibles, avec plus de 20% de changements pour la durée et le temps de redescente.

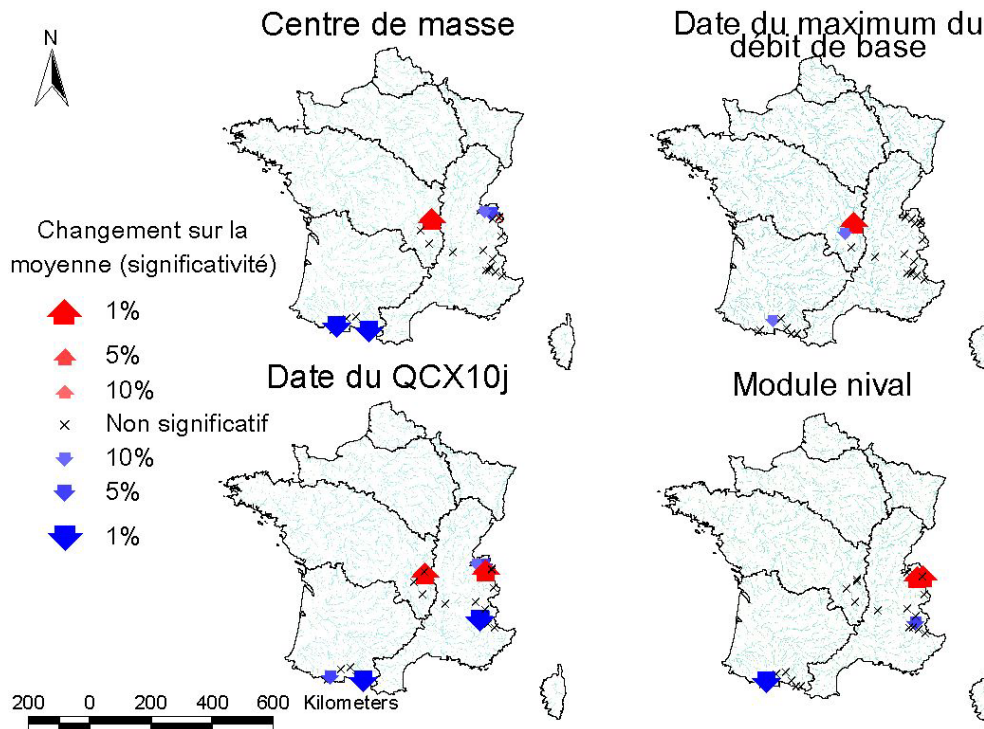
Echantillonnage	Variable	% de résultats significatifs au risque		
		1.0%	5.0%	10.0%
MAXAN	$(VCX_{1j})_A$	4.0	12.1	16.9
	θ_1	0.8	3.2	8.1
	$(VCX_d)_A$	4.0	11.3	16.1
	θ_1	0.0	2.4	5.6
	$(QCX_d)_A$	2.4	6.5	16.1
	θ_2	0.0	2.4	7.3
SUP-SEUIL	$N_3 (\mu=0.5)$	1.6	7.3	13.7
	$N_3 (\mu=1)$	2.4	8.1	10.5
	$N_3 (\mu=2)$	0.8	8.1	18.5
	$(VCX_{1j})_{\mu=1}$	3.2	8.1	12.9
	θ_3	1.6	4.8	12.1
	$(VCX_d)_{\mu=1}$	4.8	10.5	17.7
	θ_3	0.8	3.2	12.1
	$(QCX_d)_{\mu=1}$	2.4	10.5	16.1
	θ_4	2.4	8.9	12.9
	$IO (\mu=0.5)$	1.6	4.8	9.7
	$IO (\mu=1)$	1.6	4.0	12.1
	$IO (\mu=2)$	2.4	10.5	16.1
Sélection d'hydrogrammes	δ	4.8	8.9	20.2
	m	0.8	4.8	8.1
	r	4.0	16.1	23.4
	σ	3.2	12.1	19.4
	γ	2.4	8.1	12.9
	ν	1.6	12.1	19.4

Tableau 4-7. Résultats des tests de stationnarité pour les variables de crue. Les chiffres en gras représentent les résultats significatifs au niveau régional (France entière, risque régional de 5%), sous hypothèse d'indépendance spatiale.

II.3.2. Crues nivales

La composante nivale du régime des crues a été étudiée sur 25 stations situées dans les Alpes, le Massif Central et les Pyrénées. Les résultats sont présentés en Figure 4-13. On relève quelques changements nettement significatifs, qui ne sont pas *a priori* expliqués par des problèmes météorologiques ou des influences directes. Dans les Pyrénées, deux stations présentent une tendance vers une précocité accrue de l'onde de fonte nivale. Inversement, dans le massif Central, une station semble évoluer vers une fonte plus tardive. Dans les Alpes,

le résultat diffère suivant la variable choisie pour décrire la saisonnalité de la fonte. Enfin, le module de la saison nivale semble diminuer pour une station des Pyrénées et une station dans les Alpes du Sud, alors qu'il semble augmenter pour deux stations à régime glaciaire dans les Alpes du Nord. Au niveau global, les résultats présentés dans le Tableau 4-8 montrent que la date du QCX_{10j} est la variable la plus sensible. Les changements significatifs au risque 1% semblent également anormalement nombreux pour le centre de masse et le module nivale.



Variable	% de résultats significatifs au risque		
	1.0%	5.0%	10.0%
QCX_{10j}	4.0	8.0	12.0
θ_1 (Date du QCX_{10j})	16.0	28.0	36.0
q_2 (maximum saisonnier du débit de base)	4.0	12.0	16.0
θ_2 (date de q_2)	4.0	4.0	12.0
χ (centre de masse)	12.0	16.0	20.0
QS (Module saisonnier)	12.0	16.0	16.0

Tableau 4-8. Résultats des tests de stationnarité pour les variables relatives aux crues nivales. Les chiffres en gras représentent les résultats significatifs au niveau régional (France entière, risque régional de 5%), sous hypothèse d'indépendance spatiale.

II.3.3. Etiages

Les résultats concernant les quatre variables de basses eaux précédemment analysées sont présentés dans la Figure 4-14. Comme dans le cas des crues, le nombre de résultats significatifs a très nettement diminué. La cohérence spatiale des résultats est encore une fois assez faible, ce qui nous incite à croire que certains changements restent dus à des causes non climatiques que nous n'avons pas été capables de déterminer.

Ces résultats sur les étiages sont complétés avec les variables utilisant un seuil de bas débit, dont le comportement est décrit dans la Figure 4-15. Le nombre de résultats significatifs est ici très faible, ce qui peut s'expliquer par le fait que la puissance des tests est médiocre pour ces variables. En effet, le débit ne passe pas forcément sous le seuil tous les ans, et certaines années sont donc caractérisées par des durées d'étiage et des déficits de volumes égaux à zéro, entraînant la présence de plusieurs valeurs nulles. La présence d'*ex aequo* peut être prise en compte en intégrant un terme correctif dans le test de Mann-Kendall : le risque de première espèce est ainsi préservé, au détriment de la puissance de détection.

Les résultats du Tableau 4-9 résument les observations effectuées précédemment : les variables relatives à des seuils de bas débit sont nettement moins sensibles que celles décrivant les minima journaliers annuels. De plus, comme dans le cas des crues, les changements très significatifs (risque 1%) ont généralement été expliqués. En conséquence, la pseudo-significativité régionale n'est pas atteinte pour ce risque, à l'exception de la variable « date du $(VCN_{1j})_A$ ».

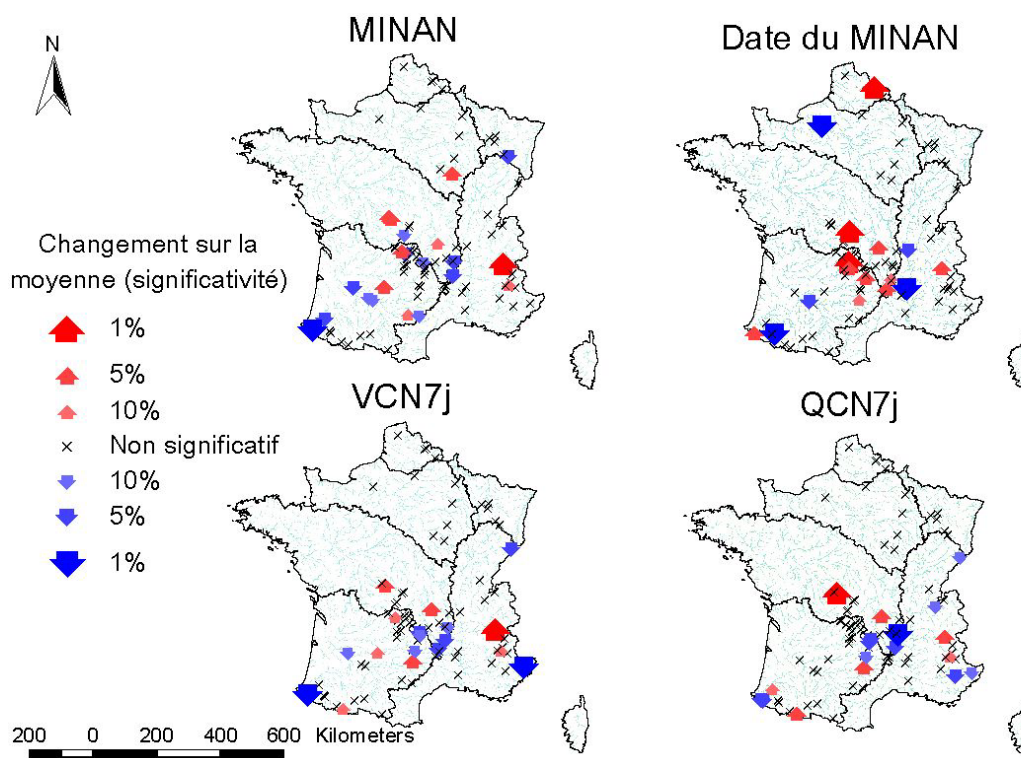


Figure 4-14. Résultat des tests de stationnarité pour quatre variables descriptives des étiages: MINAN (Minima annuels journaliers), date des MINAN, VCN7j (minima annuels des débits moyens sur 7 jours) et QCN7j (minima annuels des débits seuils sur 7 jours).

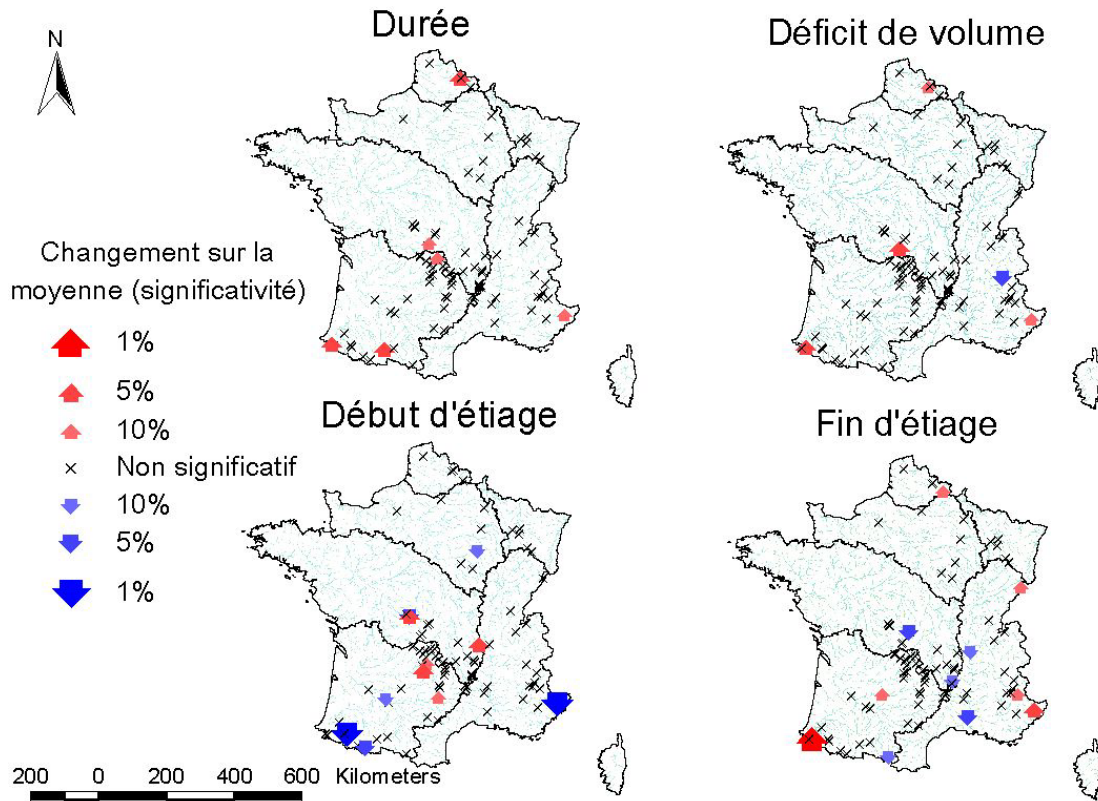


Figure 4-15. Résultat des tests de stationnarité pour quatre variables descriptives des étiages, relatives à un seuil de bas débit : durée annuelle de non-dépassement du seuil, déficit annuel de volume par rapport au seuil, date du premier et du dernier jour sous le seuil.

Variable	% de résultats significatifs au risque		
	1.0%	5.0%	10.0%
$(VCN_{1j})_A$	2.2	14.4	25.6
θ_1	6.7	15.6	17.8
$(VCN_{7j})_A$	3.3	12.2	22.2
$(VCN_{30j})_A$	2.2	13.3	21.1
$(QCN_{7j})_A$	2.2	13.3	20.0
$(QCN_{30j})_A$	3.3	11.1	20.0
δ	0.0	3.3	6.7
v	0.0	3.3	5.6
θ_{deb}	2.2	7.8	12.2
θ_{fin}	1.1	4.4	12.2

Tableau 4-9. Résultats des tests de stationnarité pour les variables d'étiage. Les chiffres en gras représentent les résultats significatifs au niveau régional (France entière, risque régional de 5%), sous hypothèse d'indépendance spatiale.

II.3.4. Régime

Les résultats concernant les deux variables décrivant le régime sont présentés dans la Figure 4-16 et le Tableau 4-10. Peu de changements apparaissent à l'échelle de la France. On peut cependant noter une baisse des modules assez nette sur trois stations du Sud-Ouest, ainsi qu'un regroupement de changement significatifs concernant la distance au régime inter-annuel dans les Alpes du Sud. La pseudo-significativité régionale n'est cependant jamais atteinte au risque 5%. *A fortiori*, elle ne le sera également pas si la dépendance est prise en compte, si l'on accepte que cette dépendance spatiale est positive.

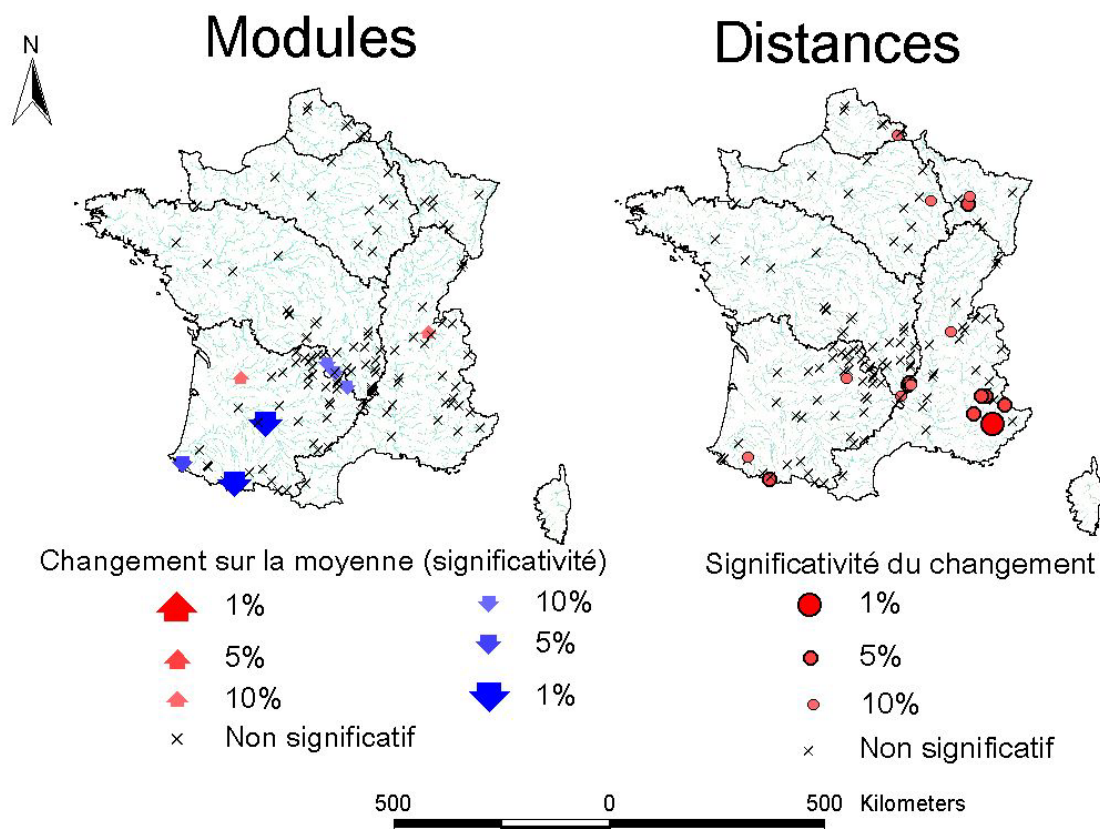


Figure 4-16. Résultat des tests de stationnarité pour les variables de régime : module annuel et distance au régime inter-annuel.

Variable	% de résultats significatifs au risque		
	1.0%	5.0%	10.0%
Module	1.5	2.3	6.2
Distance au régime	0.8	6.9	13.1

Tableau 4-10. Résultats des tests de stationnarité pour les variables de régime. Les chiffres en gras représentent les résultats significatifs au niveau régional (France entière, risque régional de 5%), sous hypothèse d'indépendance spatiale.

III. Analyse régionale

Après avoir écarté les stations influencées ou soumises à des biais météorologiques, l'analyse locale ne révèle aucun changement généralisé sur la base des séries hydrométriques étudiées, ni en hautes eaux, ni en basses eaux, ni en régime. Cela ne prouve pas l'absence d'évolutions climatiques, mais illustre simplement le fait que leurs éventuels impacts sur les débits sont trop modérés pour que nous soyons capables de les détecter. Le passage à l'échelle régionale pourrait permettre d'améliorer cette capacité de détection. Nous présentons dans cette section une évaluation des deux aspects régionaux discutés au Chapitre 3 : la significativité et la cohérence. Nous insisterons plus particulièrement sur ce dernier aspect, car il permettra de préciser quelles évolutions peuvent éventuellement être d'origine climatique. Les analyses présentées dans cette section ont été réalisées à partir des jeux réduits de stations, obtenus après critique des résultats initiaux.

III.1. Significativité régionale

Dans un premier temps, la significativité régionale des changements détectés localement doit être évaluée. Nous utilisons pour cela l'approche proposée par Douglas *et al.* (2000) basée sur des permutations Bootstrap des années. Cette évaluation n'est possible que pour les variables échantillonnées annuellement, puisque la création d'un jeu de données multivarié est problématique pour les variables événementielles, les événements n'étant pas forcément concomitants. De plus, la mise en œuvre pratique de cette analyse se heurte à plusieurs difficultés.

L'utilisation d'un large échantillon, à l'échelle de la France, rend critique le problème des lacunes. En effet, la méthode utilisée ne s'applique qu'aux années communes à toutes les stations. Sur les données que nous utilisons, ce nombre d'années communes est tout simplement nul. En conséquence, seul un sous-échantillon de stations peut être utilisé pour l'évaluation de la significativité régionale. Nous avons donc uniquement conservé les stations pour lesquelles toutes les années comprises entre 1965 et 2000 étaient disponibles.

Il est également possible de raisonner à l'échelle plus petite des régions hydro-climatiques définies au Chapitre 2. Après l'étape de critique des résultats, l'effectif de chacune de ces régions se trouve néanmoins nettement réduit, la plupart ne comptant pas plus d'une dizaine de stations. Évaluer la significativité régionale sur un si petit nombre de sites n'est alors pas très informatif : cette échelle est en effet plutôt adaptée à l'évaluation de la cohérence régionale.

Nous avons donc choisi d'évaluer la significativité régionale à l'échelle de la France, en utilisant des sous-échantillons de stations permettant de travailler avec un nombre suffisant d'années communes. Cette évaluation sera basée sur la comparaison du pourcentage de résultats significatifs observé sur les données, avec le pourcentage critique obtenu à partir de répliques Bootstrap. Signalons que cette approche comporte quelques désavantages, pouvant conduire à des biais dans l'analyse :

- ✓ Le pourcentage observé est relatif à toutes les stations, alors que le pourcentage critique calculé sur les répliques Bootstrap n'est obtenu qu'avec les stations permettant de travailler avec suffisamment d'années communes.
- ✓ Pour une station figurant à la fois dans l'échantillon total et dans l'échantillon utilisé pour les répliques Bootstrap, le test ne porte pas forcément sur les mêmes données, puisque seules les années comprises entre 1965 et 2000 sont conservées dans l'échantillon Bootstrap.

- ✓ La présence d'autocorrélation avait été préalablement testée pour les variables décrivant les basses eaux et les modules. Le cas échéant, un test de Mann-Kendall modifié avait été appliqué. Or, la permutation des années fait disparaître cette éventuelle autocorrélation dans les échantillons Bootstrap. Nous avons donc utilisé un test de Mann-Kendall classique pour évaluer localement ces variables, ce qui signifie que le test utilisé sur les données initiales et celui utilisé sur les données Bootstrap ne sont pas forcément identiques.

Malgré ces limitations, la procédure a donc été appliquée à l'échelle de la France. 10000 répliquions Bootstrap ont été effectuées. A chaque répliquion, les tests locaux sont appliqués au risque 5%, et le pourcentage de résultats significatifs est calculé. Le quantile à 95% de cet échantillon de pourcentages fournit une valeur critique qui permettra, par comparaison avec le pourcentage de changements observé, de conclure quant à la significativité régionale au risque régional de 5%.

Les résultats de cette analyse sont résumés dans le Tableau 4-11. Le nombre de stations utilisées dans les répliquions Bootstrap y est indiqué, et il apparaît qu'une importante réduction des effectifs a été nécessaire afin d'obtenir un nombre suffisant d'années communes. Les pourcentages critiques s'avèrent nettement plus importants que le risque local de 5%, ce qui n'est pas une surprise et confirme les résultats obtenus au Chapitre 3 à partir de simulations. Ces pourcentages sont également nettement plus importants que ceux que l'on obtiendrait par la loi binomiale sous hypothèse d'indépendance spatiale (resp. 10%, 20%, 12% et 10% pour les variables de crues, de crues nivales, d'étiages et de régime). Ceci démontre l'effet de la dépendance sur la significativité régionale. Enfin, lorsque l'on compare le pourcentage observé au pourcentage critique, le résultat concernant toutes les variables évaluées est identique : aucune ne semble présenter de changement significatif à l'échelle de la France. Autrement dit, le nombre de résultats significatifs détectés au risque 5% n'est pas suffisant pour garantir qu'ils ne soient pas tous dus au hasard.

Type de variable	Nombre de stations	Nombre d'années communes	Variable	Pourcentage critique	Pourcentage observé
Crues	83	36	$(VCX_{1j})_A$	19.3	12.1
			$(VCX_d)_A$	19.3	11.3
			$(QCX_d)_A$	20.5	6.5
			$(QCX_{10j})_A$	33.3	8.0
Crues nivales	15	41	θ_1 (Date du QCX_{10j})	33.3	28.0
			q_2 (maximum saisonnier du débit de base)	26.7	12.0
			θ_2 (date de q_2)	33.3	4.0
			χ (centre de masse)	53.3	16.0
			QS (Module saisonnier)	33.3	16.0
Etiages	58	36	$(VCN_{1j})_A$	24.1	14.4
			θ_1	20.7	15.6
			$(VCN_{7j})_A$	22.4	12.2
			$(QCN_{7j})_A$	25.9	13.3
			δ	31.0	3.3
			v	27.6	3.3
Régime	86	36	Module	26.7	2.3
			Distance au régime	25.6	6.9

Tableau 4-11. Significativité régionale des changements observés. Le pourcentage critique est obtenu à partir de la méthode Bootstrap proposée par Douglas et al. (2000). Les risques locaux et régionaux sont égaux à 0.05.

III.2. Cohérence régionale

L'analyse précédente confirme l'impression visuelle d'une absence d'évolution généralisée sur les débits extrêmes en France. Cependant, cette appréciation est limitée par la puissance à détecter des changements locaux. Il est ainsi possible, sur une région donnée, qu'aucun changement significatif localement n'apparaisse, mais qu'une fois les données agrégées à l'échelle régionale, la cohérence des changements locaux permette de distinguer une évolution significative. Cette section présente donc l'analyse de la cohérence régionale des changements, appliquée aux régions hydro-climatiques définies au Chapitre 2. Elle a été effectuée en utilisant l'approche semi-paramétrique (test LR-Reg) présentée au Chapitre 3.

III.2.1. Vérification des hypothèses du test LR-Reg

Nous avons insisté au Chapitre 3 sur la nécessité de vérifier l'hypothèse de normalité jointe des données transformées par normal score. A cet effet, deux outils graphiques ont été utilisés sur chaque station. Tout d'abord, nous avons vérifié la forme elliptique des nuages de points formés par les données transformées prises deux à deux. Nous avons également utilisé un graphique quantile-quantile basé sur l'observation suivante : si $(X^{(1)}, \dots, X^{(d)})$ suit une loi normale de moyenne θ et de matrice de variance Σ , alors la forme quadratique suivante doit suivre une loi du χ^2 à d degrés de liberté :

$$D = (X^{(1)}, \dots, X^{(d)}) \Sigma^{-1} (X^{(1)}, \dots, X^{(d)})^T \quad (4-1)$$

Les quantiles théoriques et empiriques peuvent donc être comparés sur un même graphique, et devraient, si le modèle est adapté, être proches de la première bissectrice.

A titre d'exemple, la Figure 4-17 illustre ces deux outils pour les maxima annuels des stations de la région Continentale. Sur la base de ces représentations, l'hypothèse de normalité jointe des données transformées ne semble pas ici irréaliste. Dans l'ensemble, l'hypothèse de dépendance Gaussienne s'est également révélée satisfaisante. Quelques exceptions ont cependant été rencontrées, et seront signalées au cours de l'analyse.

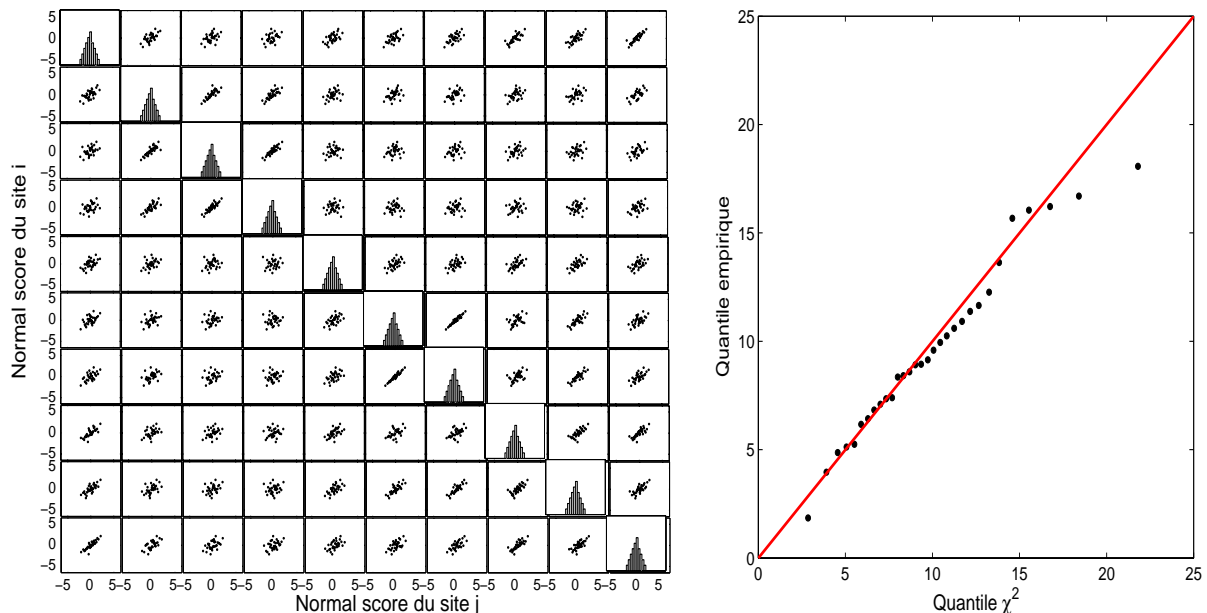


Figure 4-17. Outils graphiques utilisés pour évaluer la pertinence des hypothèses du test LR-Reg. Gauche : Matrice des nuages de points entre les données transformées par normal score ; Droite : Graphique quantile-quantile relatif à la forme quadratique D .

III.2.2. Crues pluviales

Les maxima annuels des débits journaliers, des débits moyens sur la durée caractéristique de crue d et des débits seuils sur la durée d ont été évalués avec le test LR-Reg. Pour les bassins à régime de crue pluvio-nival, seule la saison située hors de la période de fonte a été considérée. Certaines régions, contenant trop de stations par rapport au nombre d'années communes, n'ont pas été évaluées avec le test LR-Reg (Centre et Med). De plus, pour 10 stations situées en zone Océanique, le modèle de dépendance s'est avéré inapproprié. Seules les sous-régions situées dans ces zones ont donc été étudiées.

Le test LR-Reg n'a également pas pu être utilisé sur la variable « date du maximum annuel ». En effet, la forme des nuages de points obtenus n'était pas compatible avec une description Gaussienne des dépendances, particulièrement pour les régions à la saisonnalité des crues peu marquée. Nous avons donc appliqué l'approche par variable régionale, et calculé à l'échelle de la région la date moyenne du maximum annuel. Cette variable est évaluée avec un test de Mann-Kendall. De même, les dénombrements d'événements ne se prêtent pas au test LR-Reg, puisque les valeurs possibles sont entières, avec en pratique beaucoup de valeurs *ex aequo*. Nous avons donc calculé le nombre annuel total d'événements sur les stations de la région. Là encore, un test de Mann-Kendall est utilisé, avec le correctif permettant de prendre en compte les valeurs *ex aequo*. Il est en effet impossible d'utiliser un test paramétrique pour cette variable, puisque nous ignorons sa loi, qui dépend de la dépendance entre stations.

Région	Nombre de stations	Nombre d'années communes	Significativité				
			$(VC/1j)_A$	$(VC/d)_A$	$(QC/d)_A$	$N_3 (\mu=2)$	θ_1 (date du MA/AN)
AlpN-PN	3	40	x	x	x	x	x
AlpS-PN	2	31	x	x	x	x	x
Basque	3	35	x	x	x	x	x
Centre	26	18	/	/	/	x	x
CentreC	7	37	x	x	x	x	x
CentreP	12	37	x	x	-10%	x	x
Cont	10	31	x	+10%	x	x	x
Cont-Ouest	4	40	x	+5%	+10%	x	x
Cont-Vosges	5	37	+10%	+10%	x	x	x
Lozère	4	38	x	x	x	x	x
Med	20	22	/	/	/	x	x
Med (hors Alpes et Cevennes)	8	41	x	x	x	x	x
MedAlpes	3	28	x	x	x	x	x
MedCev	9	31	x	x	x	x	x
NivCent	4	41	x	x	x	x	x
Ocean	10	21	/	/	/	x	x
OceanBP	2	29	+5%	+10%	+10%	x	x
OceanNord	4	36	+5%	+10%	x	+5%	x
OceanOuest	4	63	x	x	x	x	x
Pyr-PN	4	32	-5%	-5%	-10%	x	x
Rhône	6	43	x	x	x	x	x
Roussillon	3	36	x	x	x	x	x
SudOuest	9	24	+5%	x	x	x	x

Tableau 4-12. Cohérence régionale des changements pour les crues pluviales. Le chiffre indique la significativité du test LR-Reg, un signe positif (resp. négatif) désignant un changement à la hausse (resp. à la baisse). Les changements non significatifs sont signalés par des croix.

Le Tableau 4-12 résume finalement les résultats obtenus à l'échelle régionale. Les régions des Alpes, du Pays Basque, de Lozère, de Méditerranée, du Centre (à régime pluvio-nival), du Rhône et du Roussillon ne présentent aucun changement cohérent significatif sur les variables étudiées. Un changement à la baisse, significatif à 10%, est détecté pour les débits seuils sur 12 stations pluviales de la région Centre. Dans le Sud-Ouest, on observe une hausse des maxima annuels journaliers significative à 5%. Les stations de la région Continentale connaissent des changements cohérents plus marqués, allant vers une hausse des maxima annuels. Cette évolution est également observée à l'échelle des sous-régions. Une hausse cohérente du risque de crue est également observée dans les sous-régions Nord et Bassin Parisien de la zone Océanique. Enfin, les stations pluvio-nivales des Pyrénées présentent une tendance à la baisse des intensités des crues pluviales.

III.2.3. Crues nivales

Les six variables décrivant l'onde de fonte nivale ont été évaluées sur cinq régions montagneuses. Le test LR-Reg a pu être appliqué à l'ensemble des variables, y compris les variables de date. Ceci s'explique par la saisonnalité très marquée des événements nivaux, contrairement à ce que l'on observe pour les crues pluviales. Les résultats, résumés dans le Tableau 4-13, montrent que des évolutions cohérentes sont observées sur les Alpes du Nord, à la fois pour cinq stations à influence nivale et pour les trois stations à régime glaciaire. Dans le premier cas, une précocité accrue de l'onde de fonte est détectée sur la base des variables « date du QCX_{10j} », « date du maximum du débit de base » et « Centre de masse ». Concernant les trois stations glaciaires, une légère précocité est également détectée sur la base de la variable « date du maximum du débit de base », accompagnée d'une hausse du module nival significative à 5%. Cependant, seules 17 années communes avec moins de 10% de lacunes ont pu être utilisées sur ces stations : les résultats sont donc à considérer avec précaution.

Région	Nombre de stations	Nombre d'années communes	Significativité					
			QCX_{10j}	θ_1 (Date du QCX_{10j})	q_2 (maximum saisonnier du débit de base)	θ_2 (date de q_2)	χ (centre de masse)	QS (Module saisonnier)
AlpesN	5	40	x	-10%	x	-10%	-5%	x
AlpesS	5	30	x	x	x	x	x	x
Pyr	6	30	x	x	x	x	x	x
NivCent	4	55	x	x	x	x	x	x
Glaciaire	3	17	x	x	x	-10%	x	+5%

Tableau 4-13. Cohérence régionale des changements pour les crues nivales.

III.2.4. Etiages

Six variables descriptives des basses eaux ont été évaluées sur les régions hydro-climatiques d'étiage. Le test LR-Reg a été appliqué, à l'exception de la variable « date du minimum annuel », pour laquelle nous avons testé l'évolution de la date moyenne sur la région (test de Mann-Kendall). Le modèle de dépendance Gaussienne s'appliquait en effet difficilement sur certaines stations. Précisons que les tests utilisés à l'échelle régionale ne prennent pas en compte l'autocorrélation, que l'on observe parfois sur ces variables : les résultats devront donc être interprétés avec prudence, particulièrement pour les régions où l'influence de la nappe d'accompagnement est importante.

Les résultats sont résumés dans le Tableau 4-14. Les stations des Alpes présentent globalement une tendance vers des étiages moins sévères. Cela se traduit par des minima annuels journaliers en hausse et des déficits de volume en baisse sur trois stations des Alpes

du Nord, et par des durées d'étiage en baisse sur quatre stations des Alpes du Sud. A l'inverse, on relève une augmentation des durées d'étiages sur les Pyrénées. Cette augmentation est également observée sur le Pays Basque, accompagnée d'une augmentation des déficits de volume. Les stations Océaniques présentent quant à elles des minima annuels en hausse. Parmi ces six stations, cinq sont situées dans la sous-région « Nord ». L'évolution des minima annuels n'y est plus significative, mais on relève une augmentation des durées d'étiage, qui semble à première vue contradictoire avec l'augmentation des minima. Précisons que dans cette région, l'influence des nappes d'accompagnement est prépondérante, et la plupart des séries présentent une autocorrélation significative. La significativité des évolutions observées est donc douteuse. Enfin, on relève sur trois stations du Rhône une augmentation significative à 10% des minima annuels des débits seuils sur sept jours. Les autres régions (Centre, Continentales, Méditerranée, Sud Ouest) ne présentent aucun changement régionalement cohérent.

Région	Nombre de stations	Nombre d'années communes	Significativité					
			$(VCN_{1j})_A$	θ_1 (date du MINAN)	$(VCN_{7j})_A$	$(QCN_{7j})_A$	δ	ν
AlpN	3	41	+10%	X	X	X	X	-5%
AlpS	4	32	X	X	X	X	-10%	X
Basque	5	34	X	X	X	X	+5%	+10%
BasqueC	3	35	X	X	X	X	+5%	+10%
BasqueP	2	90	X	X	X	X	X	X
Centre1	8	33	X	X	X	X	X	X
Centre2	13	35	X	X	X	X	X	X
Cont	8	31	X	X	X	X	X	X
ContC	2	38	X	X	X	X	X	X
ContOuest	4	40	X	X	X	X	X	X
Med	6	30	X	X	X	X	X	X
Ocean	6	26	+5%	X	+5%	+10%	X	X
OceanNord	5	34	X	X	X	X	+10%	X
Pyr	3	30	X	X	X	X	+5%	X
PyrC	2	33	X	X	X	X	X	X
Rhone	3	81	X	X	X	+10%	X	X
SudOuest	5	27	X	X	X	X	X	X

Tableau 4-14. Cohérence régionale des changements en étiage.

III.2.5. Régime

Le sens d'un changement sur la distance au régime inter-annuel n'ayant pas d'interprétation hydrologique simple, seul le module a été évalué avec le test LR-Reg. Nous évaluerons ultérieurement un possible glissement de régime observé sur les stations des Alpes du Sud (cf. section II.3.4.). Le module d'une station hydrométrique étant en grande partie contrôlé par la pluie annuelle que reçoit son bassin versant, nous avons utilisé la classification relative aux étiages pour définir les régions relatives aux modules : cette classification était en effet principalement basée sur le zonage pluviométrique. Les sous-régions n'ont cependant pas été considérées, puisque cette distinction supplémentaire reflétait des différences sur la saisonnalité des basses eaux.

Les résultats sont résumés dans le Tableau 4-15. Il apparaît que le seul changement détecté concerne une baisse des modules significative à 5% dans les Pyrénées. Aucune des autres régions ne semble présenter de changement cohérent sur cette variable.

Région	Nombre de stations	Nombre d'années communes	Significativité
			Module
AlpN	7	41	x
AlpS	4	32	x
Basque	5	34	x
Centre1	10	33	x
Centre2	17	27	x
Cont	9	31	x
Med	7	30	x
Ocean	9	29	x
Pyr	6	24	-5%
Rhone	4	81	x
SudOuest	6	27	x

Tableau 4-15. Cohérence régionale des changements en régime.

III.2.6. Récapitulatif des évolutions régionalement cohérentes

Les résultats précédents sont résumés par région hydro-climatique, afin de visualiser les évolutions observées à cette échelle.

Alpes du nord (Figure 4-18)

Les changements concernent les crues nivales et les étiages. L'onde de fonte semble être plus précoce, et les étiages semblent moins sévères, avec une augmentation significative des minima annuels journaliers et une diminution significative du déficit de volume. La durée des étiages semble également décroître, mais pas de manière significative.

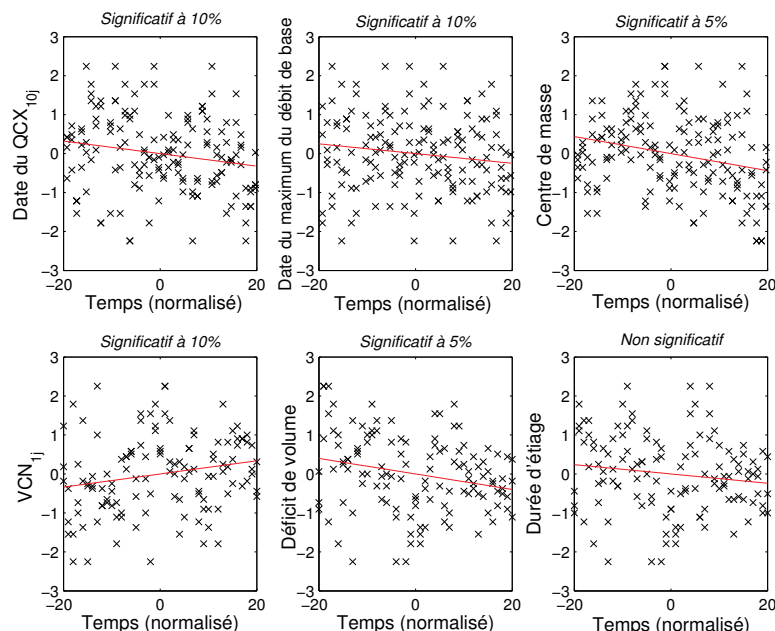


Figure 4-18. Evolution de six variables dans les Alpes du Nord. Les variables sont transformées par normal score, et le temps centré sur zéro. La droite représente une estimation de la tendance régionale.

Alpes du Sud

La sévérité des étiages semble diminuer, comme c'était le cas dans les Alpes du Nord (Figure 4-19, à droite). Par contre, l'onde de fonte nivale ne présente pas de changement cohérent sur ces cinq stations.

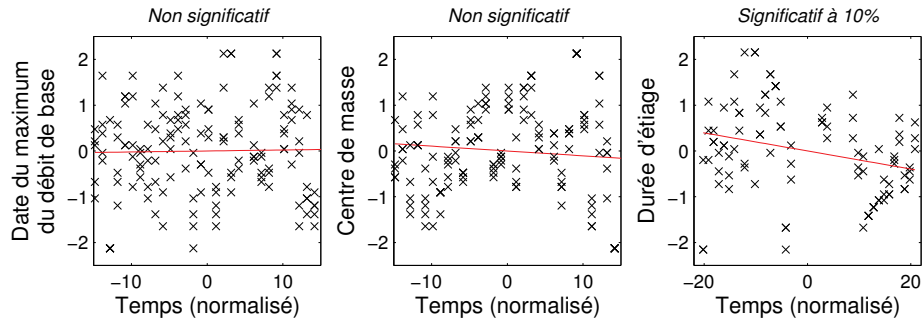


Figure 4-19. Evolution de trois variables dans les Alpes du Sud.

Nous avons évoqué lors de l'analyse locale un regroupement de changements significatifs dans les Alpes du Sud pour la variable « distance au régime inter-annuel ». Afin de préciser ces observations, nous avons évalué au niveau régional l'évolution de la contribution des débits mensuels au module, pour chacun des douze mois de l'année. Cette étude devrait permettre de préciser de quelle manière évolue la répartition des écoulements au cours de l'année. Les résultats sont présentés en Figure 4-20. Il y apparaît que pour la plupart des mois, l'évolution semble mineure à l'échelle régionale. On relève malgré tout une diminution de la contribution des trois mois d'été au module, cette diminution étant significative au mois de juillet.

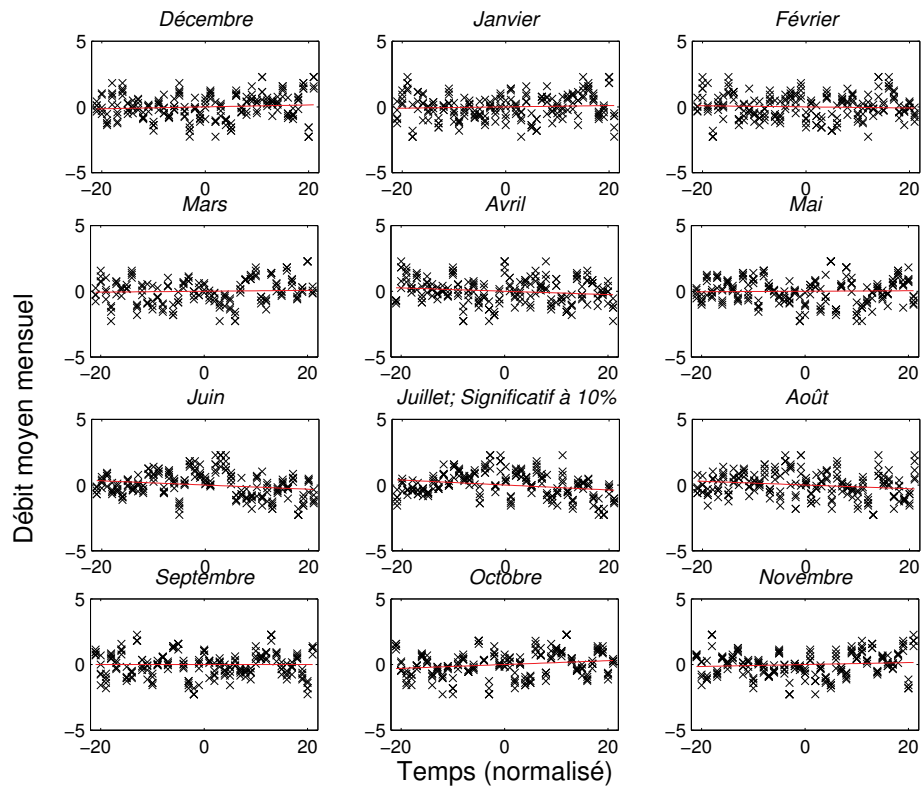


Figure 4-20. Evolution des débits moyens mensuels, exprimés en pourcentage du module annuel, dans les Alpes du Sud.

Glaciaires (Figure 4-21)

Les trois stations à régime glaciaire présentent deux changements sur l'onde de fonte : une précocité accrue de la date du pic de fonte, ainsi qu'une augmentation assez nette du module total de la saison nivale. Bien que ces résultats n'aient été obtenus qu'avec 17 années de données communes, ils mériteraient d'être complétés par des analyses complémentaires, car l'augmentation du module semble cohérente avec une récession des glaciers dans les Alpes Françaises (Vincent, 2002; Vincent et al., 2004).

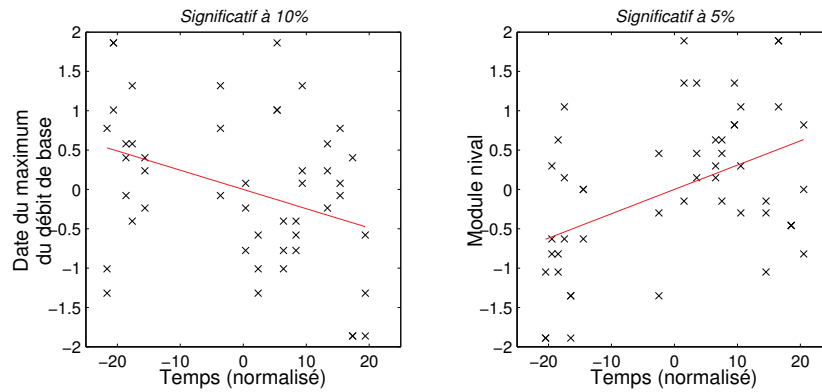


Figure 4-21. Evolution de deux variables pour les régimes glaciaires.

Pyrénées (Figure 4-22)

La composante nivale des crues ne semble pas présenter de changement dans cette région. Par contre, on relève une diminution de l'intensité des crues d'origine pluvial. Les étiages semblent également s'aggraver (déficit de volume et durée en hausse), alors que le module connaît une baisse significative à 5%. La tendance générale semble donc à l'assèchement dans les Pyrénées.

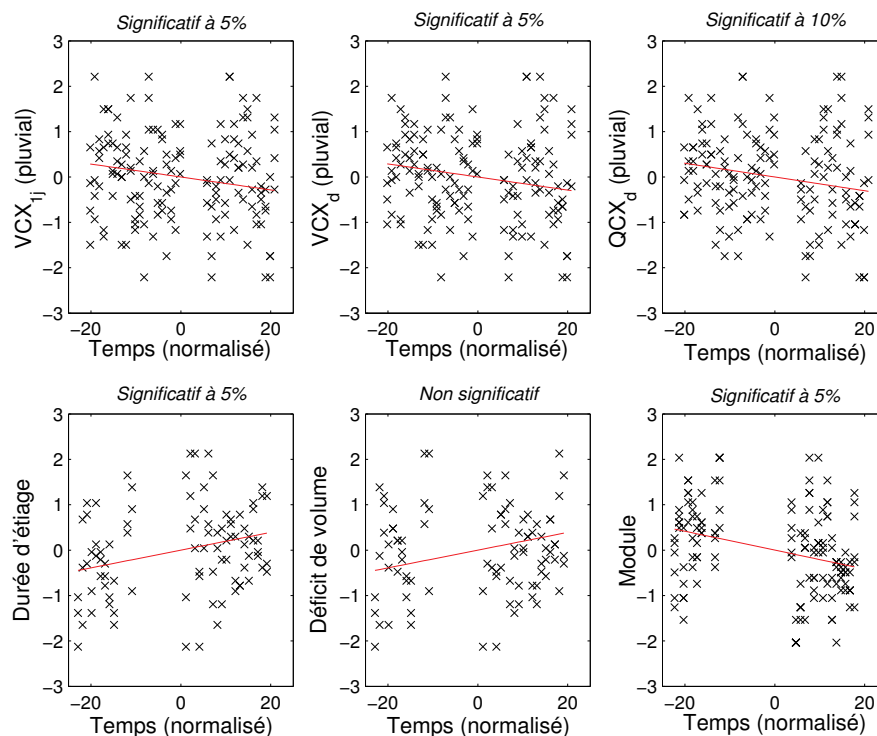


Figure 4-22. Evolution de six variables dans les Pyrénées.

Pays Basque (Figure 4-23)

Les changements observés dans le Pays Basque concernent les étiages, qui tendent à s'aggraver. Ces évolutions peuvent d'ailleurs être rapprochées des tendances relevées sur les Pyrénées voisines. Le gestionnaire de la DIREN Aquitaine nous a confirmé que cette tendance à l'aggravation était ressentie dans la région.

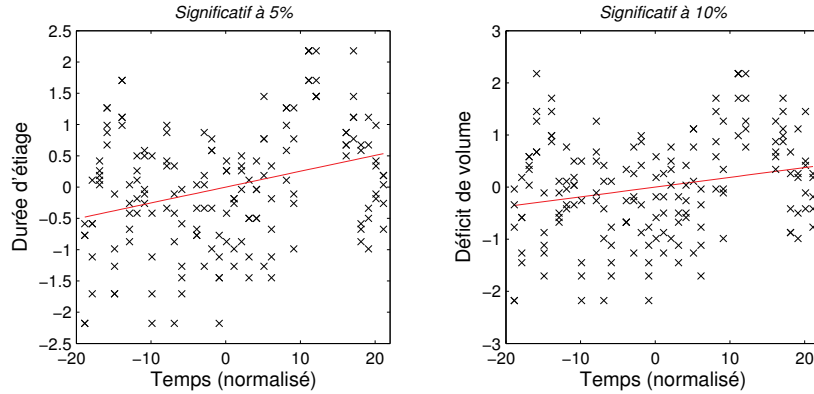


Figure 4-23. Evolution de deux variables dans le Pays Basque.

Continentaux

Les résultats sont présentés sur la région complète (Figure 4-24), ainsi que pour les sous-régions « Ouest » (Figure 4-25) et « Vosges » (Figure 4-26). Bien que la significativité des évolutions diffère suivant l'échelle considérée, tous les changements vont dans le même sens et indiquent une augmentation de l'intensité des crues.

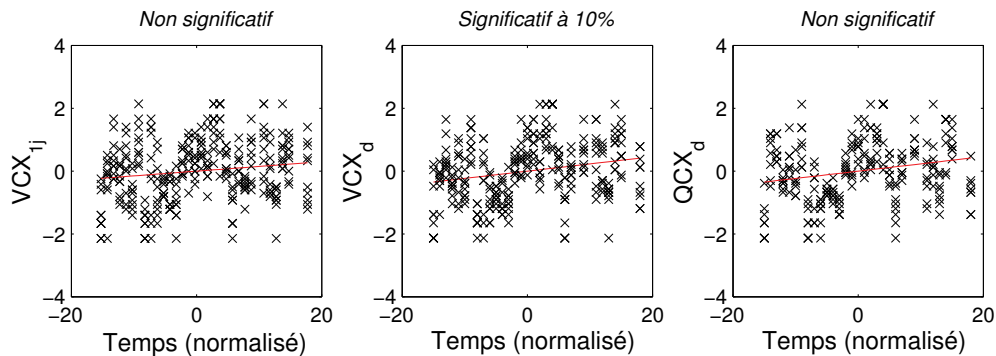


Figure 4-24. Evolution de trois variables en région Continentale.

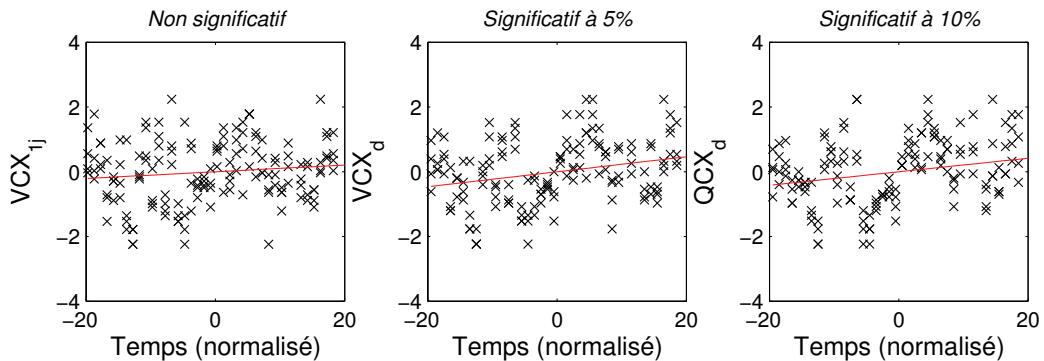


Figure 4-25. Evolution de trois variables dans la sous-région « Ouest » de la zone Continentale.

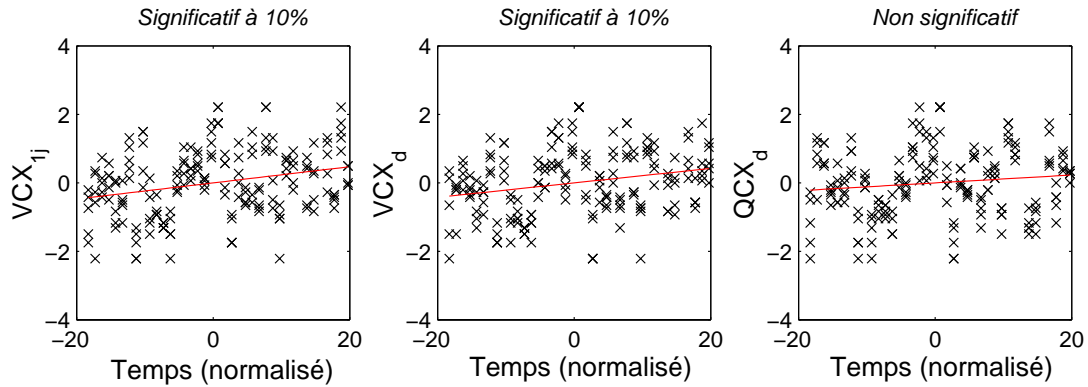


Figure 4-26. Evolution de trois variables dans la sous-région « Vosges » de la zone Continentale.

Océaniques

Les évolutions concernant les hautes eaux sont présentées dans la Figure 4-27. Que ce soit dans la partie Nord ou dans le Bassin Parisien, la tendance est à l'augmentation du risque de crue, avec des maxima annuels en hausse. La fréquence des événements semble également augmenter dans le Bassin Parisien.

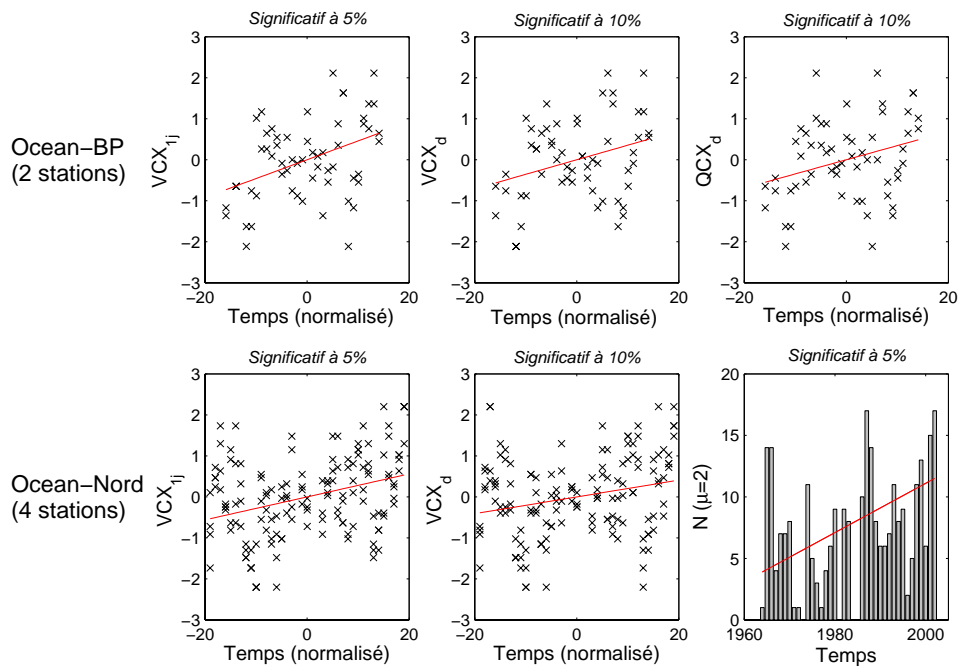


Figure 4-27. Evolution de six variables de crue en zone Océanique. Les variables sont transformées par normal score, à l'exception de la variable « Nombre annuel d'événements sur la région ($N, \mu=2$) ».

En basses eaux, cette tendance au plus humide apparaît également (Figure 4-28), à l'exception de la variable « durée d'étiage » sur les stations du Nord, qui semble en augmentation.

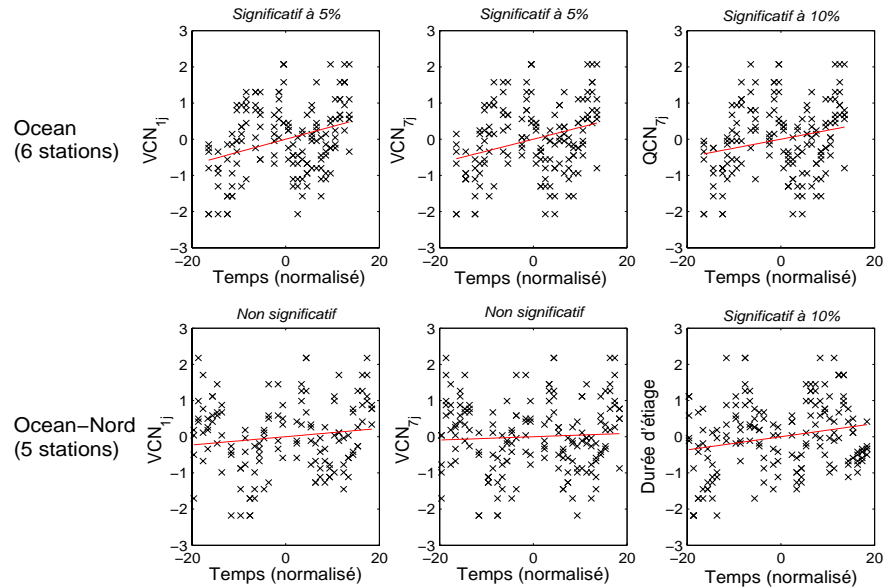


Figure 4-28. Evolution de six variables d'étiage en zone Océanique.

Les évolutions constatées sur cette région doivent être considérées avec précaution : le contrôle par la nappe d'accompagnement est en effet important, ce qui, d'un point de vue statistique, rend douteuse l'hypothèse d'indépendance des observations, y compris en hautes eaux. En guise d'illustration, la Figure 4-29 présente la superposition des hauteurs de nappe, des minima et des maxima annuels de la Lys à Delettes. On remarque que la valeur des minima comme des maxima est fortement liée au niveau de remplissage de la nappe. La variabilité de cette dernière semble assez complexe, avec une alternance de périodes sèches et humides, qui pourrait s'expliquer par l'inertie naturelle de ce type de système hydrologique. En conséquence, nous pensons que l'analyse statistique telle qu'elle a été effectuée n'est pas adaptée à ce type de cours d'eau. Des études complémentaires, avec une prise en compte du niveau de la nappe et des pluies, seraient nécessaires pour étudier le déterminisme climatique de ces évolutions.

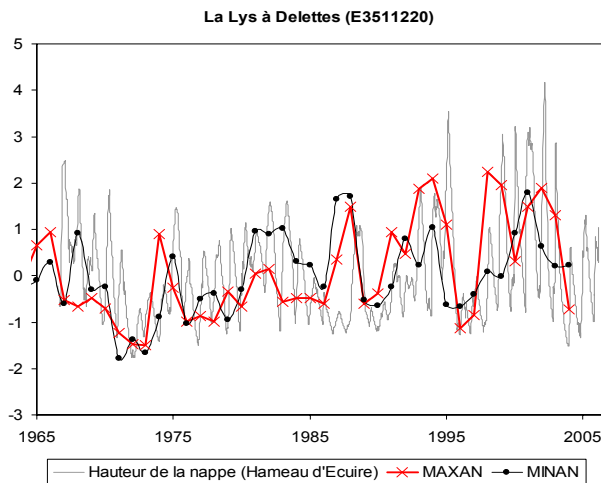


Figure 4-29. Valeurs centrées-réduites des hauteurs mensuelles de nappe, des minima annuels journaliers et des maxima annuels journaliers.

Autres (Figure 4-30)

Trois changements supplémentaires ont également été relevés au niveau régional, mais leur importance semble moindre, notamment car ils ne concernent qu'une unique variable. Dans la région Centre, les cours d'eau à régime pluvial pur présentent une légère diminution des maxima annuels des débits seuils sur d jours (durée caractéristique de crue). Dans le Sud Ouest, on relève une augmentation significative à 5% des maxima annuels journaliers. Enfin, les débits de trois stations du Rhône présentent une augmentation des minima annuels des débits seuils sur 7 jours.

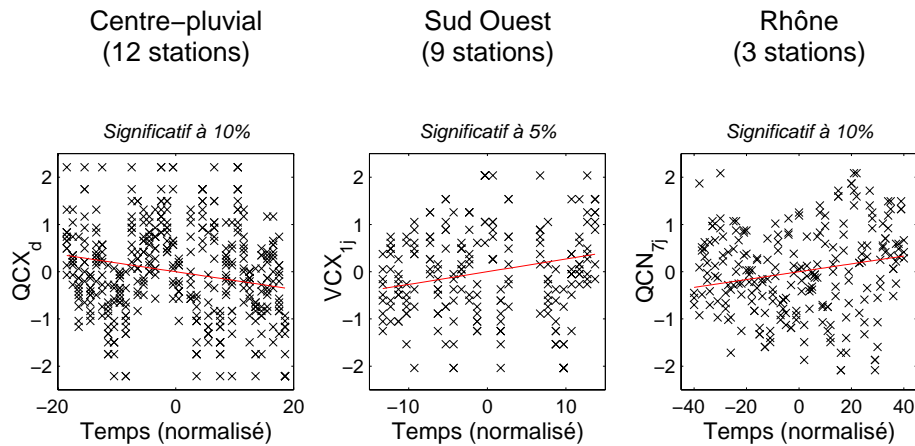


Figure 4-30. Evolution de trois variables dans les régions Centre (pluvial), Sud Ouest et Rhône.

IV. Lien avec des variables explicatives

L'étude régionale a permis de mettre en avant quelques régions présentant des changements cohérents. Ceci incite à penser que ces changements sont bien réels, et ne sont pas dus à des problèmes météorologiques ou des influences directes liées à des aménagements, qui ne devraient pas *a priori* exercer une influence régionale sur les débits. Cependant, une étude uniquement basée sur des données hydrométriques ne suffit pas à affirmer que ces évolutions sont climatiques. En effet, le rôle filtre joué par le bassin versant est une composante importante dans la transformation pluie/débit. Or, en France, rares sont les bassins versants n'ayant pas évolué au cours du vingtième siècle (urbanisation, agriculture, déforestation ou reboisement, etc.). L'impact de ces changements affectant le bassin versant est difficilement quantifiable, en particulier pour les événements extrêmes, mais ne peut être exclu. Il apparaît donc nécessaire d'utiliser des données climatologiques afin de faire la part de ce qui relève de la variabilité climatique (températures, précipitations) de ce qui relève de changements dans l'occupation des sols.

Dans cette section, nous allons tenter d'évaluer la cohérence entre des évolutions détectées régionalement sur les débits et des évolutions détectées sur les pluies ou les températures. Dans un premier temps, nous exploiterons de façon empirique les résultats du projet IMFREX (Impact des changements anthropiques sur la FRéquence des phénomènes EXtrêmes de vent, de température et de précipitations). Malheureusement, la forte non-linéarité de la relation pluie-débit rend ce type de comparaison délicate, particulièrement pour les crues. Nous tenterons donc, dans un second temps, de relier l'évolution observée des débits avec l'évolution de données issues d'un modèle hydrologique, pour lequel les variables d'entrée utilisées sont les précipitations et les températures observées.

IV.1. Comparaison qualitative avec les résultats du projet IMFREX

Les conclusions générales du projet IMFREX ont été présentées au Chapitre 1. Nous détaillons dans cette section quelques résultats pouvant être reliés à l'évolution des débits. Toutes les figures présentées dans cette section sont issues du site Internet du projet (<http://medias.dsi.cnrs.fr/imfrex/web/index>), avec l'aimable autorisation de Michel Déqué.

IV.1.1. Crues pluviales

Les crues pluviales constituent le phénomène le plus difficile à relier empiriquement avec les pluies. En première approximation, on peut supposer que si les précipitations intenses ont tendance à augmenter, il devrait en être de même pour les crues. Malheureusement, la réalité est loin d'être aussi simple. Tout d'abord, l'intensité des pluies génératrices de crues diffère très fortement entre les bassins versants. De plus, l'état hydrique du sol au moment de l'événement pluvieux est très important. Ainsi, une forte pluie pourra n'occasionner qu'une crue modérée si l'infiltration est importante sur le bassin, alors qu'une pluie plus faible pourra avoir des effets plus importants sur un sol saturé. En plus de ces difficultés physiques, nous nous heurtons également à une difficulté pratique, qui est la faible couverture spatiale des stations pluviométriques journalières étudiées dans le projet IMFREX. En effet, en l'absence de procédure d'homogénéisation de données fiable au pas de temps journalier, les acteurs du projet ont privilégié la création de séries quotidiennes de référence (SQR). Ces séries répondent à un certain nombre de critères garantissant leur qualité et leur homogénéité, mais n'ont pas été corrigées.

La Figure 4-31 représente les évolutions détectées sur les maxima annuels de précipitations, et la Figure 4-32 celles concernant le demi-décile supérieur et le nombre de

jours de fortes pluies. Dans tous les cas, le principal point commun avec nos résultats est la difficulté à observer une cohérence spatiale dans ces évolutions. De plus, la plupart des changements ne sont pas significatifs. Enfin, la plupart des zones dans lesquelles nous suspectons une évolution sur les crues (Centre, Pyrénées, Sud Ouest, Continentale (Nord-Est) et Océanique) ne contiennent presque aucune SQR. Seul le Bassin Parisien possède une densité de poste qui pourrait nous permettre d'effectuer la comparaison. Ainsi, il semblerait que le nombre de jours de pluie et les maxima annuels de cumuls sur 10 jours soient en augmentation, ce qui pourrait être cohérent avec l'aggravation des crues que nous avons détectée. Cependant, le contrôle des écoulements par la nappe d'accompagnement ajoute un degré de complexité supplémentaire, qui semble rendre difficile toute conclusion empirique sur le lien entre ces évolutions.

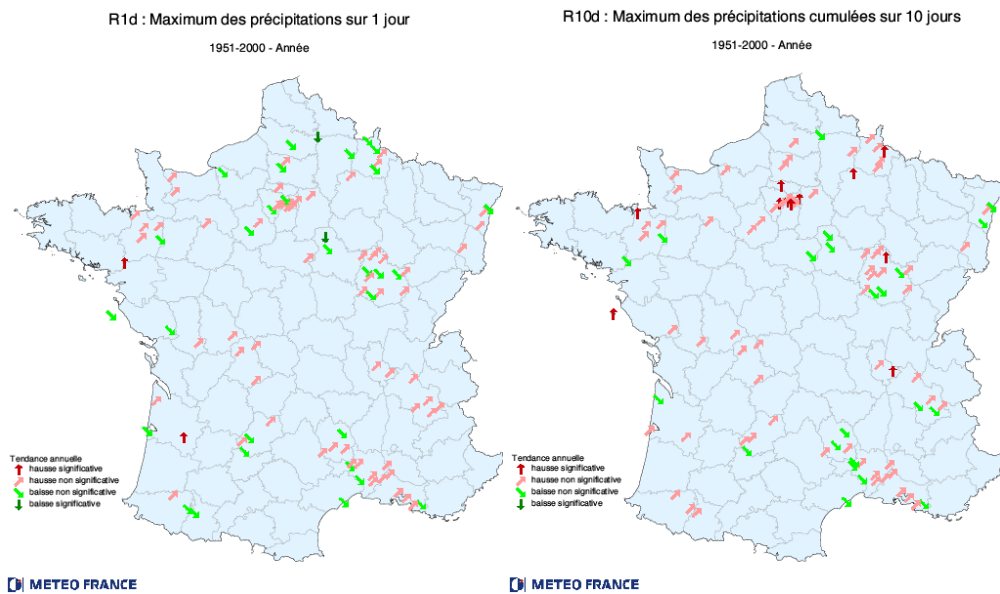


Figure 4-31. Evolution des maxima annuels des pluies journalières et cumulées sur 10 jours pour les séries quotidiennes de référence.

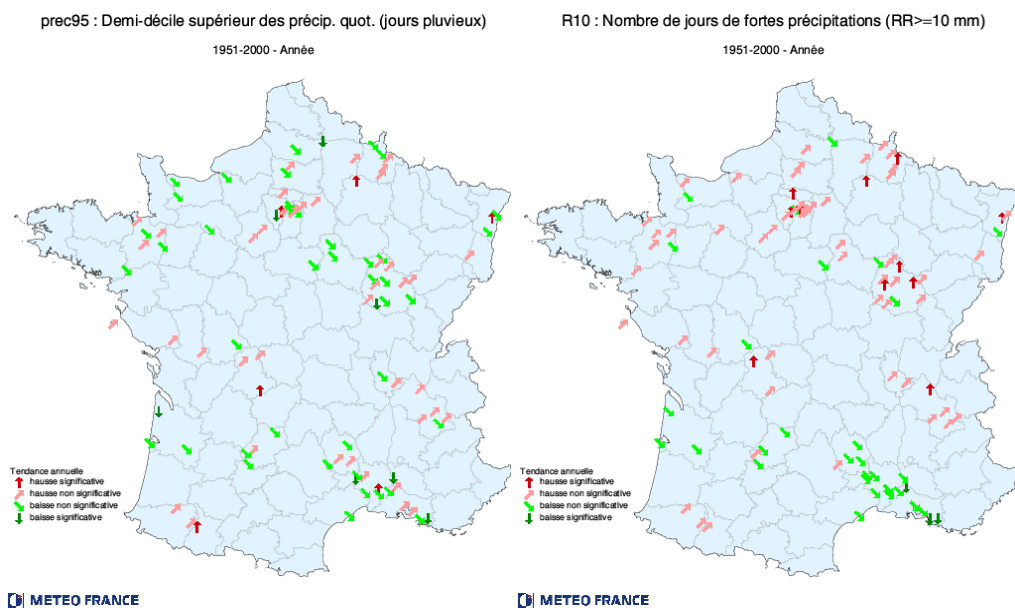


Figure 4-32. Evolution du demi-décile supérieur et de nombre de jours de fortes précipitations pour les séries quotidiennes de référence.

IV.1.2. Crues nivales

Les crues nivales sont *a priori* plus appropriées à une comparaison empirique : en effet, elles dépendent essentiellement de variables climatologiques moyennes (cumuls hivernaux de précipitations et températures moyennes printanières) qui sont plus simples à analyser que les extrêmes, et pour lesquelles une meilleure couverture spatiale est disponible. Cependant, la relation entre la fonte nivale et ces variables explicatives n'est pas triviale. Par exemple, des températures printanières et estivales en hausse associées à des pluies hivernales stationnaires devraient avoir pour effet de décaler l'onde de fonte vers plus de précocité. Mais si une partie du bassin est recouverte de glaciers, que l'on peut considérer comme des stocks « infinis », alors les écoulements de fonte pourront être observés plus tardivement dans la saison. En conséquence, la durée de la fonte sera plus importante, résultant en une onde plus étendue, mais dont le centre de masse sera inchangé.

Ainsi, concernant les changements détectés sur les Alpes du Nord, l'augmentation du module des rivières glaciaires pourrait être associée à une augmentation de la fonte des glaciers, elle-même liée à la hausse des températures (Figure 4-33). Ce réchauffement pourrait également expliquer la précocité accrue de l'onde de fonte pour les stations nivales. Cependant, il peut paraître étonnant que cette précocité ne s'observe pas sur les Alpes du Sud, d'autant que le réchauffement semble y être plus marqué. Cependant, la couverture spatiale des pluviomètres est très mauvaise pour cette région (Figure 4-33), ce qui ne permet pas de dresser une quelconque conclusion, la hausse du stock de neige pouvant éventuellement compenser celle des températures.

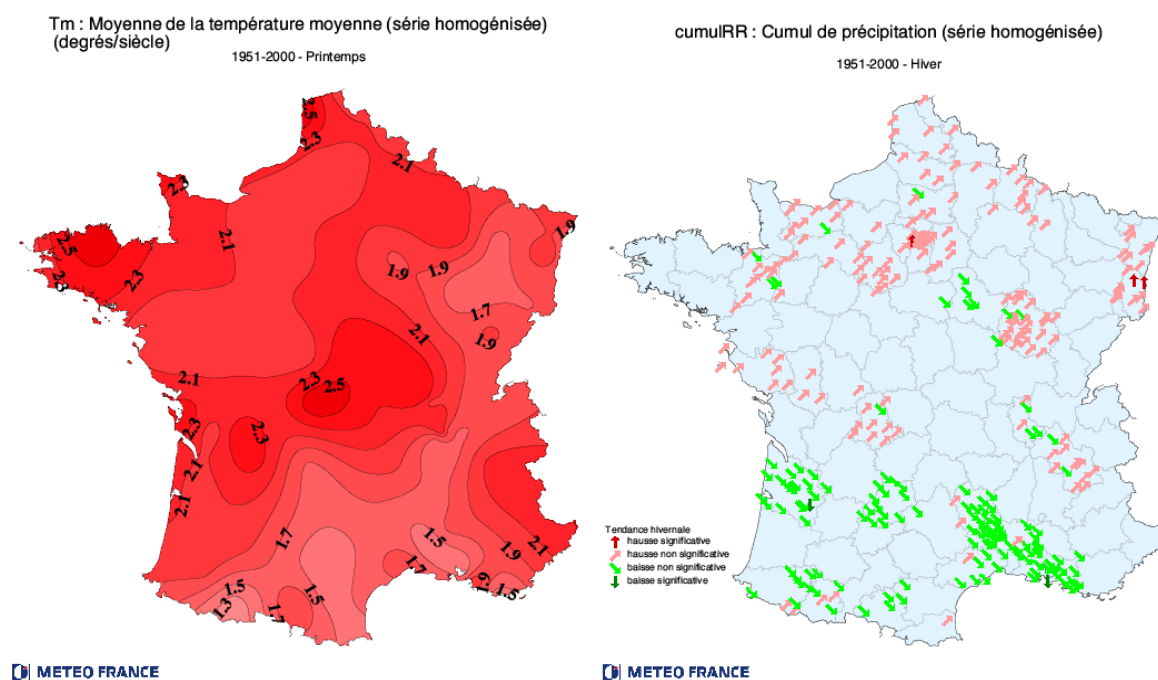


Figure 4-33. Evolution des températures printanières moyennes et des cumuls hivernaux de précipitations pour les séries mensuelles homogénéisées.

IV.1.3. Etiages

Comme pour les crues nivales, les étiages sont contrôlés au moins en partie par des variables moyennes (long déficit de précipitations, périodes chaudes), ce qui devrait rendre la comparaison empirique plus pertinente que pour les crues pluviales.

Les stations des Alpes sont caractérisées par des étiages hivernaux, liés à l'absence de précipitations liquides. Dans ce type de région montagneuse, l'accompagnement de la nappe ne contribue que faiblement aux écoulements de surface, de sorte que la sévérité d'un étiage dépendra principalement de la durée durant laquelle les précipitations tombent sous forme de neige, et donc des températures. Une hausse des températures devrait en effet avoir pour conséquence une diminution de cette durée, et donc des étiages moins marqués. Il existe donc une certaine cohérence entre les évolutions vers des étiages moins sévères que nous avons détectés sur les Alpes, et le réchauffement des températures hivernales moyennes (Figure 4-34, gauche). Des indices plus raffinés peuvent être calculés, mais nécessitent des données de température au pas de temps journalier, qui n'existent que sous forme de SQR, et dont la couverture spatiale est très éparse (Figure 4-34, droite).

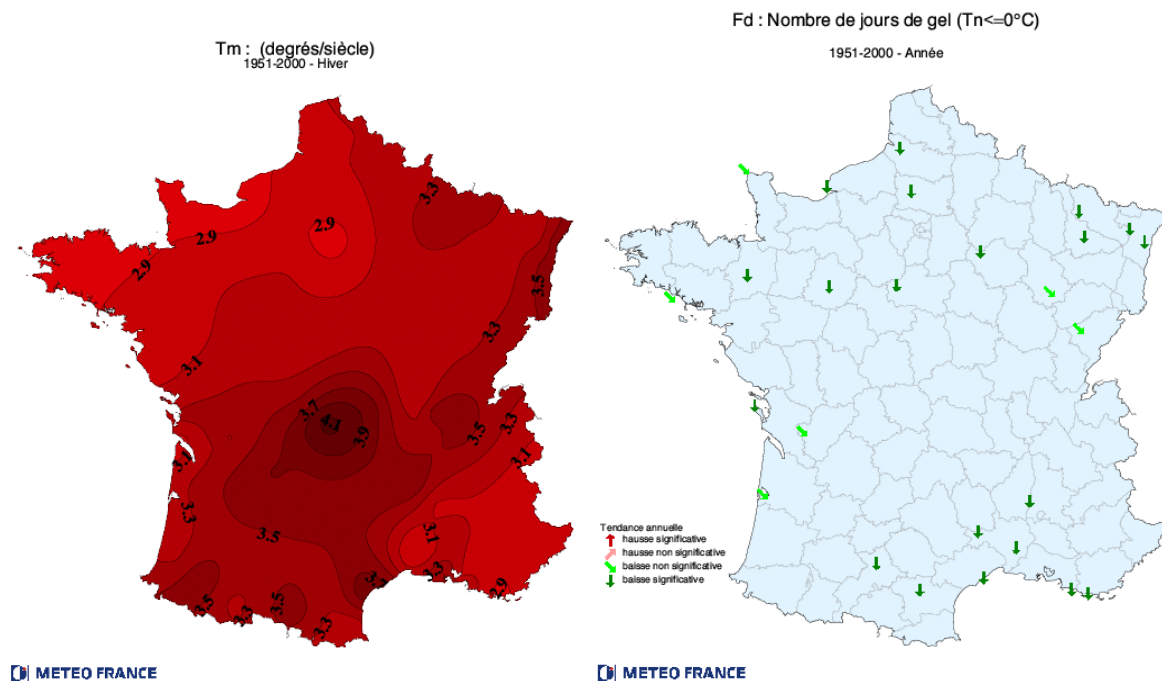


Figure 4-34. Evolution des températures hivernales moyennes (séries mensuelles homogénéisées) et du nombre annuel de jours de gel (SQR).

Les stations des Pyrénées sont plus problématiques à analyser. En effet, les bassins versants étant situés à des altitudes plus modérées, la part des écoulements liés à la composante nivale devient négligeable relativement tôt dans l'année. En conséquence, les étiages peuvent résulter de deux phénomènes distincts : la sécheresse estivale, ou le stockage des précipitations sous forme de neige. La saisonnalité des étiages est ainsi moins marquée sur ces stations, et la saison des basses eaux s'étend de la fin de l'été jusqu'à la fin de l'hiver.

On peut ainsi remarquer que la hausse des températures estivales (Figure 4-35, gauche) pourrait entraîner des étiages plus sévères en fin d'été, mais que comme dans le cas des Alpes, la hausse des températures hivernales devrait par contre résulter en des étiages moins sévères. Il semble donc difficile de tirer une quelconque conclusion sur ces stations. Remarquons également que cette aggravation des étiages peut paraître surprenante alors que les Pyrénées constituent la seule région où les précipitations estivales ne semblent pas baisser.

Les stations du Pays Basque possèdent quant à elles une période d'étiage située entre la fin de l'été et la fin de l'automne. La hausse des températures estivales pourrait donc expliquer l'aggravation des étiages détectée, d'autant que les précipitations ne semblent pas cette fois-ci à la hausse. Il peut cependant sembler étonnant que cette évolution des étiages ne soit pas constatée dans d'autres régions, qui connaissent un réchauffement estival bien plus marqué.

Pour conclure, l'évolution des stations Océaniques ne peut être comparée aux évolutions des températures et des précipitations, le rôle prépondérant de la nappe rendant cette analyse empirique trop complexe.

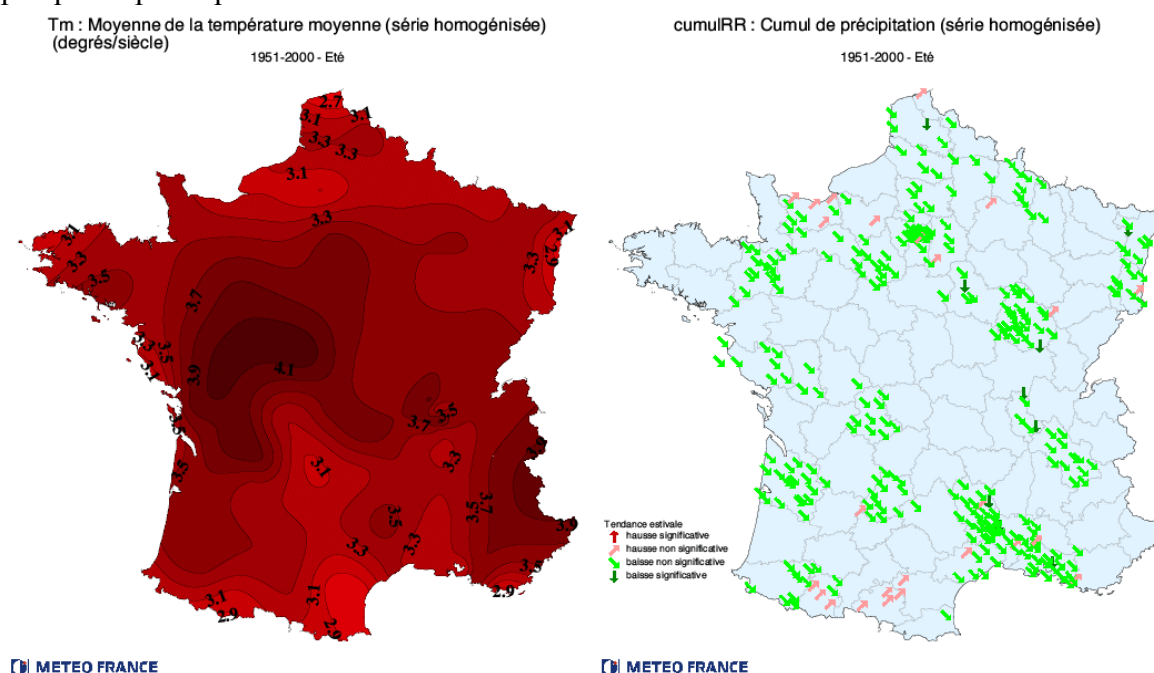


Figure 4-35. Evolution des températures estivales moyennes et des cumuls estivaux de précipitations pour les séries mensuelles homogénéisées.

IV.1.4. Régime

L'unique évolution régionale observée en module concernait les stations des Pyrénées, qui présentaient une baisse des écoulements annuels. Cette observation n'est *a priori* pas cohérente avec l'évolution observée des cumuls annuels de précipitations (Figure 4-36), qui semblent plutôt à la hausse, bien que les changements ne soient pas significatifs. Il est ainsi envisageable que la relation pluie-débit ait évolué, pour des raisons qui restent à déterminer.

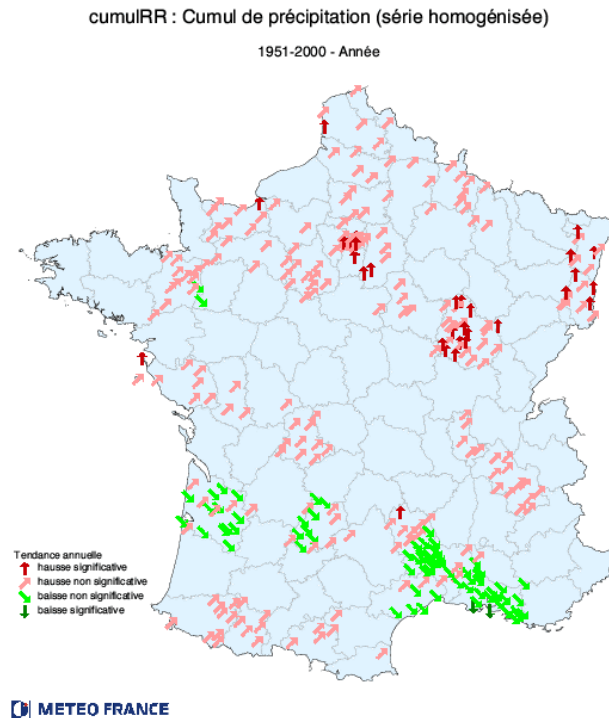


Figure 4-36. Evolution des cumuls annuels de précipitations pour les séries mensuelles homogénéisées.

IV.2. Lien avec les pluies et les températures *via* la modélisation hydrologique

L'analyse empirique n'a pas permis de mettre en évidence des relations simples entre les évolutions détectées sur les débits et les variables explicatives. Ceci n'est pas étonnant, étant donné la complexité des relations pluie-température-débit. Afin de vérifier si ces évolutions sont bien compatibles avec le comportement des pluies et des températures, la seule solution est d'utiliser la modélisation hydrologique, afin de comparer l'évolution observée des débits avec celle obtenue par une modélisation hydrologique. Nous détaillons dans cette section deux exemples d'une telle analyse. Le premier concerne l'évolution des crues pluviales constatée dans la région Continentale, le second celle des crues nivales dans les Alpes.

IV.2.1. Crues pluviales du Nord-Est

L'analyse sur la région Continentale, composée de dix stations, faisait apparaître une tendance possible à la hausse des maxima annuels (cf. section III.2.6.). Nous allons à présent exploiter les données pluviométriques concernant quatre stations de la sous-région « Vosges », dont les principales caractéristiques sont résumées dans le Tableau 4-16.

Code HYDRO	Rivière / Station	Superficie du bassin versant (km ²)	Module (m ³ .s ⁻¹)	Débit journalier décennal (m ³ .s ⁻¹)
A4200630	La Moselle / Saint-Nabord	621	24	360
A4250640	La Moselle / Epinal	1220	39	470
A5431010	Le Madon / Pulligny	940	10	220
A5730610	La Moselle / Toul	3350	63	790

Tableau 4-16. Caractéristiques hydrologiques des quatre bassins étudiés.

Modèle et données

Le modèle pluie-débit utilisé est le modèle GR4J (Figure 4-37), dans la version proposée par Perrin (2000). Il s'agit d'un modèle global conceptuel. Il considère tout d'abord le bassin versant comme une entité homogène, sans prendre en compte la distribution spatiale des processus. De plus, il n'utilise pas les équations de la physique pour modéliser les transferts au sein du bassin, mais décrit la transformation pluie-débit comme la résultante de la vidange et du remplissage de deux réservoirs (production et routage) connectés entre eux. Les données d'entrée sont les pluies journalières, ainsi que l'évapotranspiration potentielle (ETP) estimée au pas de temps décadaire. Le débit est obtenu, au pas de temps journalier, en transformant ces données *via* le modèle, cette transformation dépendant de quatre paramètres. Ces paramètres ne sont pas des grandeurs physiques mesurables, mais sont calés, à partir d'une période d'observations conjointes de pluies et de débits, afin que les débits simulés soient le plus conforme possible aux débits observés. Le critère d'adéquation utilisé ici est le critère de Nash (1970).

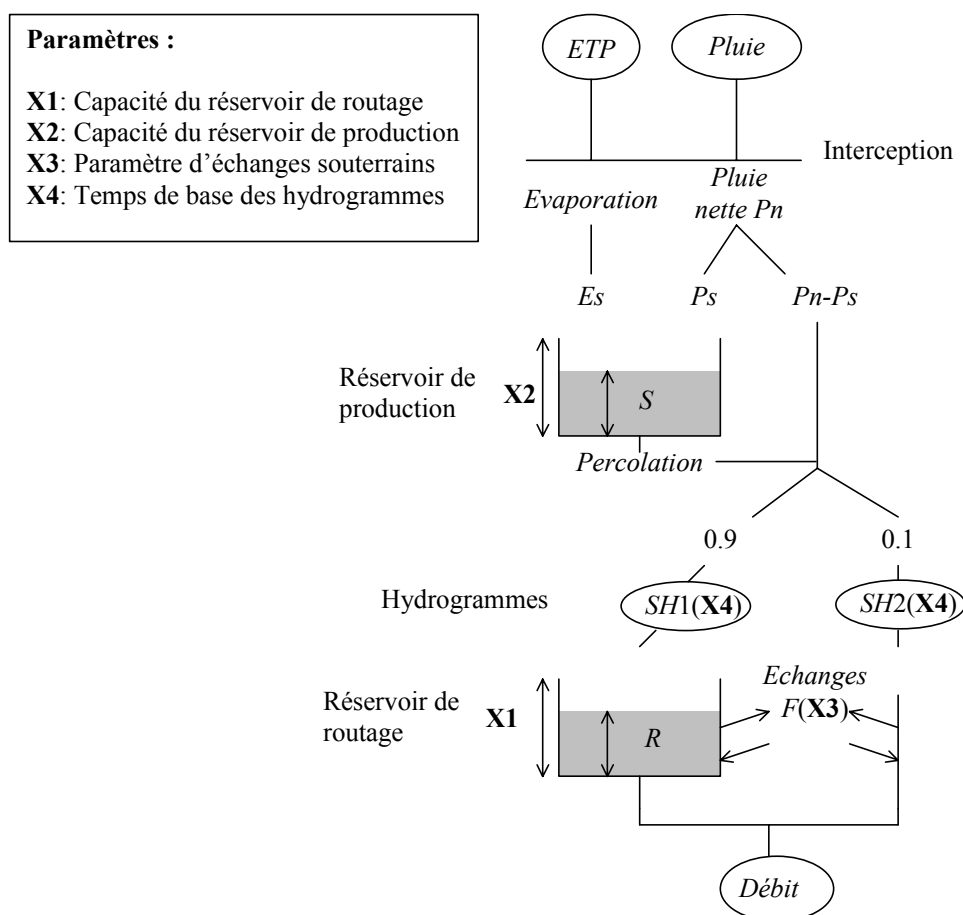


Figure 4-37. Schéma de la structure du modèle GR4J.

Les données de précipitations utilisées sont décrites dans la thèse de Pierre Javelle (2001). A partir des pluies journalières ponctuelles, des pluies de bassin sont calculées pour chacune des quatre stations hydrométriques analysées. Elles sont obtenues par simple moyenne arithmétique des pluies observées sur les postes pluviométriques du bassin. Les données d'ETP correspondent toutes à la valeur calculée à Nancy, à partir de données de température et d'humidité relative (source : Météo France). Les données d'entrée sont finalement disponibles entre les années 1968 et 1998.

Comparaison des évolutions observées et simulées

Avant de s'intéresser à l'évolution temporelle des débits simulés par le modèle, nous avons effectué une première analyse sur les précipitations. Les variables évaluées sont les maxima annuels des pluies journalières (RX_{1j}) et des cumuls sur trois jours (RX_{3j}), le nombre annuel de jours de fortes pluies (supérieures à 10 mm, R_{10mm}), et la pluie annuelle (R_m). Chaque variable a été analysée à l'aide du test LR-Reg. Les résultats, présentés dans la Figure 4-38, montrent que seul le nombre de jours de fortes pluies présente une tendance à la hausse significative (à 10%). De plus, la tendance estimée sur les maxima annuels de pluie journalière est quasiment nulle, alors que les maxima annuels de débits journaliers présentaient une tendance à la hausse significative. Cette observation justifie l'utilisation d'un modèle pour comparer ces évolutions, et illustre la complexité de la relation pluie-débit, y compris du point de vue de la stationnarité.

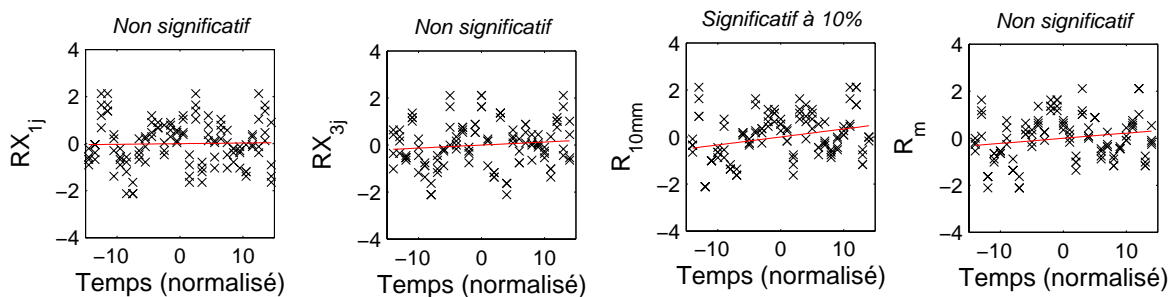


Figure 4-38. Evolution de quatre variables descriptives des précipitations relatives aux quatre bassins versants étudiés.

Le test LR-Reg a ensuite été appliqué aux maxima annuels de débits journaliers, extraits à partir des observations et des simulations sur la période 1968-1998. Afin d'évaluer l'influence de la période de calage choisie, trois périodes distinctes ont été considérées : 1968-1979, 1978-1989 et 1988-1998. La première année de calage est dédiée à l'initialisation de l'état de saturation du bassin, de sorte que le critère de calage est calculé sur des périodes disjointes. Le contrôle des débits simulés a également été effectué sur les périodes complémentaires. Les résultats, détaillés dans le Tableau 4-17, montrent que les débits simulés recréent convenablement les observations, à la fois en calage et en contrôle.

Période de calage		Station Hydrométrique			
		A4200630	A4250640	A5431010	A5730610
1968-1979	Calage	85.3	85.5	78.6	87.9
	Contrôle	87.4	87	83	89.1
1978-1989	Calage	87.3	87.5	86.4	89.7
	Contrôle	87.4	87.1	83.5	89.9
1988-1998	Calage	91.5	91.2	87.1	93.2
	Contrôle	86.9	86.4	82.7	89.1

Tableau 4-17. Critères de Nash (%) en calage et en contrôle pour trois périodes de calage.

Les résultats des tests de stationnarité sont détaillés dans la Figure 4-39. Les observations (Figure 4-39a) présentent sur cette période une tendance à la hausse significative à 5%. Sur cette même période, les débits simulés (Figure 4-39b, c, et d) conduisent également à une tendance à la hausse, significative à 10% pour deux périodes de calage et non significative pour une troisième. Ce résultat nous laisse penser que cette évolution observée sur les débits est liée, au moins en partie, à une évolution affectant les précipitations. Cela ne suffit pas à incriminer le changement climatique (il peut s'agir de variabilité naturelle), mais le déterminisme climatique de cette tendance semble plus vraisemblable que la présence de biais météorologiques.

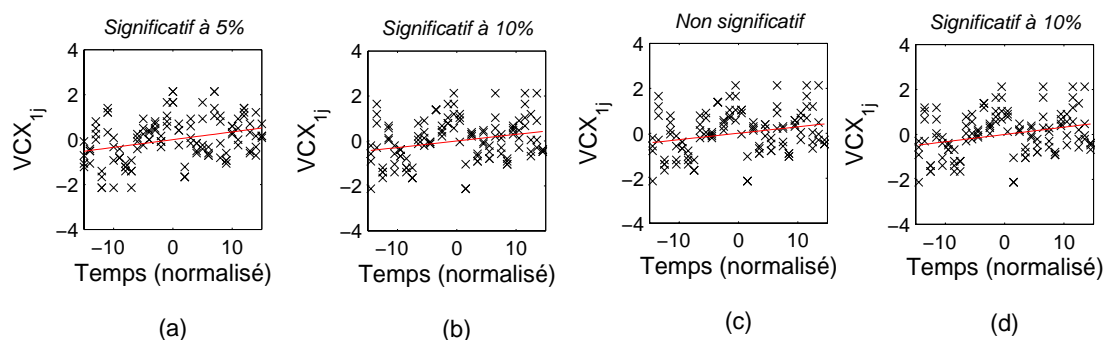


Figure 4-39. Evolution des maxima annuels de débits journaliers pour les quatre bassins versants étudiés. (a) : observations, (b) : simulations GR4J avec calage sur 1968-1979, (c) : simulations GR4J avec calage sur 1978-1989, (d) : simulations GR4J avec calage sur 1988-1998.

Stationnarité de la relation pluie/débit

Une évolution des débits observée sur un bassin versant peut être liée à une évolution des variables climatiques, mais également à un changement dans les processus de génération de ces débits. Par exemple, d'importantes modifications de l'occupation du sol dans le bassin versant peuvent avoir une incidence sur les écoulements. Afin d'évaluer ce type de phénomène, Andréassian *et al.* (2003) ont proposé un test basé sur la modélisation hydrologique. Le principe de cette méthode repose sur des simulations croisées (Figure 4-40). La période d'étude est tout d'abord découpée en p sous-périodes. Sur chacune de ces sous-périodes, les paramètres du modèle sont calés, et des débits sont simulés à partir des pluies R_1, \dots, R_p observées. On obtient donc une matrice dont l'élément figurant en ligne i et en colonne j est lié aux débits simulés à partir des pluies de la période i par le modèle calé sur la période j . Si l'on considère cette matrice par colonne, les débits simulés sont relatifs à un calage donné, et à l'ensemble des pluies observées. Les éventuelles évolutions seront donc d'origine pluviométrique. C'est cette évaluation qui a été réalisée dans la section précédente, avec trois sous-périodes de calage. Si la matrice est considérée par ligne, alors les débits simulés seront relatifs à une pluie donnée, pour différents jeux de paramètres calés. Une éventuelle évolution sera donc liée à une modification des paramètres du modèle, et donc de la relation pluie/débit.

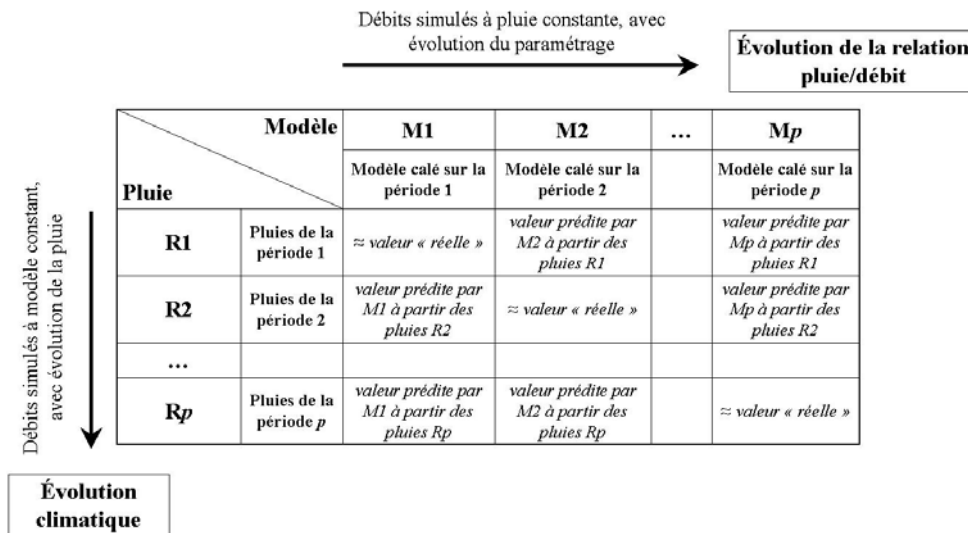


Figure 4-40. Principe des simulations croisées (d'après Andréassian et al. (2001)).

La significativité des évolutions détectées peut alors être évaluée par une technique de permutation. Tout d'abord, pour chaque cellule de la matrice, il faut définir une variable hydrologique à évaluer. Les écoulements en crue peuvent par exemple être décrits par le volume total écoulé au dessus d'un seuil de haut débit pour une sous-période donnée. Si une tendance affecte la relation pluie/débit en crue, les valeurs obtenues dans chaque ligne devraient avoir tendance à croître ou à décroître. Andréassian *et al.* (2003) proposent donc d'utiliser la statistique de test suivante :

$$S = \sum_{i=1}^p \left[\sum_{j=1}^{i-1} (q_{ii} - q_{ij}) + \sum_{j=i+1}^p (q_{ij} - q_{ii}) \right] \quad (4-2)$$

où q_{ij} correspond à l'élément de la ligne i et de la colonne j . Une valeur critique de la statistique peut alors être estimée en recréant un grand nombre de matrices par permutations des colonnes de la matrice originale. Sous l'hypothèse H_0 d'absence d'évolution dans la relation pluie/débit, la valeur observée de la statistique devrait être comprise entre les quantiles de probabilité $\alpha/2$ et $1-\alpha/2$ de la distribution des statistiques obtenues par permutation (α étant le risque de première espèce).

Cette procédure a été appliquée pour chacune des quatre stations hydrométriques étudiées. Le seuil de haut débit a été choisi égal au quantile à 95% de la courbe des débits classés. Six sous-périodes de cinq années ont été utilisées, et la significativité du changement a été évaluée à partir des 720 (=6!) permutations possibles. Les résultats, illustrés par la Figure 4-41, montrent que l'hypothèse de stationnarité de la relation pluie/débit ne peut être rejetée sur trois stations hydrométriques. Seule la troisième station, le Madon à Pulligny, présente une évolution significative à 10%.

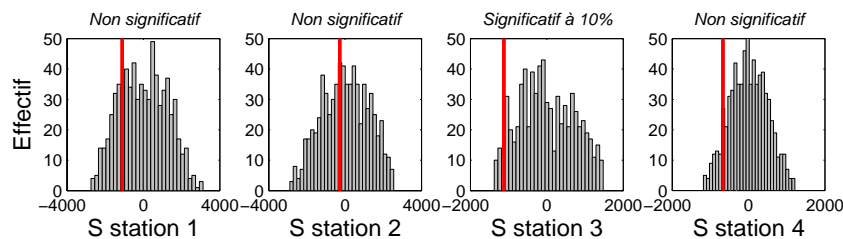


Figure 4-41. Comparaison des statistiques d'Andréassian et al. (2003) simulées (histogrammes) et observées (trait vertical) pour les quatre stations étudiées.

La matrice des simulations croisées de cette station est présentée dans le Tableau 4-18. Le recodage visuel proposé par Andréassian et al. (2003) a été appliqué. Il consiste à remplacer les valeurs numériques par le signe « + » si la valeur d'une cellule est rangée dans le sens croissant par rapport à la valeur diagonale, et par le signe « - » sinon. En observant ces matrices ligne par ligne, on remarque que la tendance serait à la décroissance des volumes écoulés en crue pour une pluie identique, ce qui est contraire à l'aggravation observée sur les débits journaliers. Des mécanismes de compensation pouvant intervenir, il est difficile de conclure de manière définitive quant à la cause de cette aggravation. A notre connaissance, aucun changement majeur dans l'occupation du sol de ce bassin versant ne peut expliquer cette évolution de la relation pluie/débit. En conclusion, nous pensons que les changements observés restent en grande partie liés à l'évolution de la pluviométrie, et que des analyses complémentaires seraient nécessaires pour décider si l'évolution de la transformation pluie/débit du Madon à Pulligny est réelle ou est simplement due à un effet d'échantillonnage.

	M1	M2	M3	M4	M5	M6	M1	M2	M3	M4	M5	M6
P1	100.0	87.9	89.9	92.9	80.1	78.5	0	-	-	-	-	-
P2	114.2	100.0	100.5	104.3	99.0	92.0	-	0	+	+	-	-
P3	112.2	98.8	100.0	104.1	99.1	91.0	-	+	0	+	-	-
P4	110.0	94.5	95.5	100.0	92.7	87.0	-	+	+	0	-	-
P5	116.0	103.4	103.6	104.5	100.0	92.9	-	-	-	-	0	-
P6	127.7	113.1	111.5	115.5	104.1	100.0	-	-	-	-	-	0

Tableau 4-18. Matrices des simulations croisées pour le Madon à Pulligny. Les valeurs correspondent aux volumes écoulés au-dessus d'un seuil de haut débit. Pour chaque ligne de la matrice de gauche, ces volumes sont exprimés en pourcentage par rapport à la valeur diagonale. La matrice de droite correspond au recodage des signes.

IV.2.2. Crues nivales des Alpes

Le second cas d'étude concerne la composante nivale des stations hydrométriques situées dans les Alpes. Les données journalières de pluies et de températures nous ont été fournies par EDF-DTG Grenoble pour trois stations, dont les caractéristiques sont résumées dans le Tableau 4-19. Bien que deux de ces stations soient situées dans les Alpes du Nord, et la troisième dans les Alpes du Sud, nous avons considéré qu'elles pouvaient être étudiées conjointement à l'échelle régionale.

Code HYDRO	Rivière / Station	Superficie du bassin versant (km ²)	Module (m ³ .s ⁻¹)	Débit journalier décennal (m ³ .s ⁻¹)
W2714010	La Romanche / Chambon amont	220	7.5	48
V0222010	L'Arve / Arthaz	1664	75	440
X0500010	La Durance / Serre-Ponçon	3580	79	560

Tableau 4-19. Propriétés hydrologiques des trois stations étudiées.

Modèle et données

Le modèle utilisé est nommé MORDOR, et a été développé par EDF (Garcon, 1999; Paquet, 2004). Il s'agit encore une fois d'un modèle conceptuel global à réservoirs, mais sa complexité est plus grande que GR4J, puisqu'il permet notamment de prendre en compte les écoulements liés à la neige (Figure 4-42). En contrepartie, il est nécessaire de fournir des données de précipitations et de températures en entrée du modèle, et le nombre de paramètres à caler (22) est bien plus important.

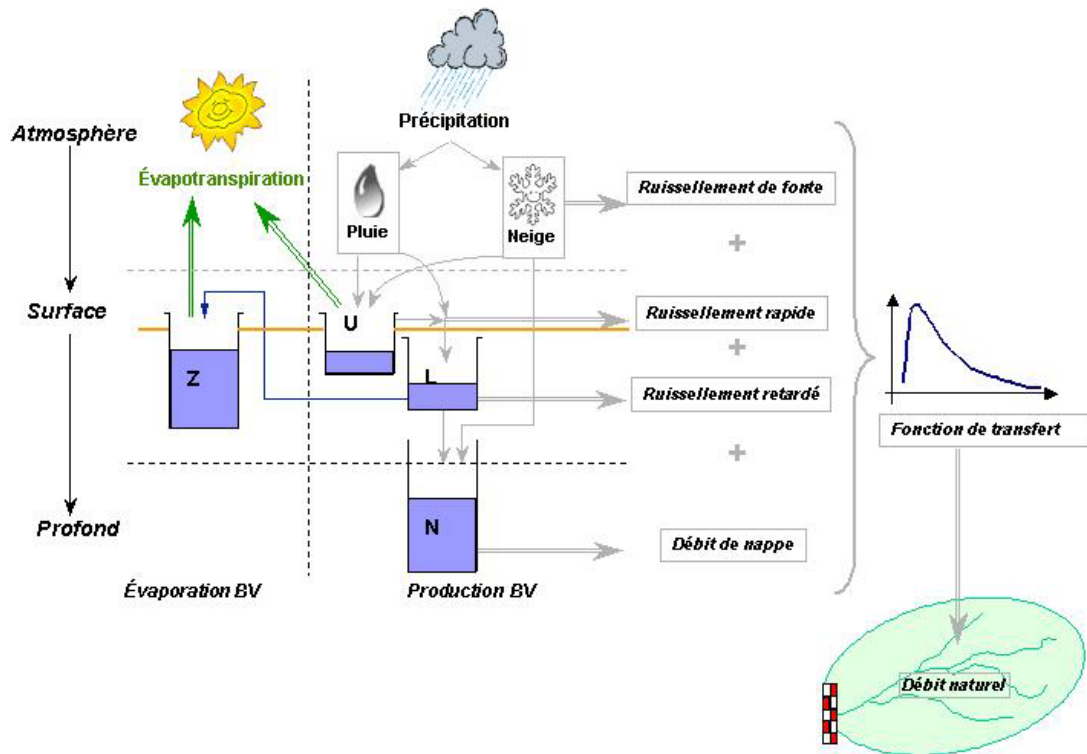


Figure 4-42. Schéma de fonctionnement du modèle MORDOR.

Les précipitations de bassin sont des moyennes pondérées des précipitations ponctuelles mesurées sur différents postes (ici, entre 2 et 6 postes sont utilisés par bassin). Les pondérations dépendent du pouvoir explicatif des pluies ponctuelles, mesuré à partir d'une régression multiple entre les débits simulés avec chacune des pluies ponctuelles et les débits mesurés. Les températures sont elles obtenues, sur ces trois bassins, à partir d'un unique point de mesure. Bien que le modèle ne soit pas distribué, la modélisation du stockage des précipitations sous forme de neige impose de tenir compte de la variabilité des températures dans de tels bassins d'altitude. Ainsi, pour une température donnée au point de mesure, un pourcentage de précipitations stocké sous forme de neige est calculé, en utilisant la courbe hypsométrique du bassin, et un abattement de la température mesurée en fonction de l'altitude.

Le calage des paramètres a été effectué par Emmanuel Paquet et Joël Gailhard (EDF-DTG Grenoble), en utilisant un critère mixte prenant en compte l'adéquation des débits journaliers observés et simulés, ainsi que l'adéquation entre les courbes de débits classés observées et simulées. Etant donné le nombre de paramètres à caler, la période de calage est égale à la période d'étude. Il ne sera donc pas possible d'évaluer l'impact de la période de calage sur les résultats. L'ensemble des données utilisées pour le calage ainsi que l'adéquation des débits simulés aux observations est résumé dans le Tableau 4-20.

Station (Code HYDRO)	Poste de température	Nombre de postes pluviométriques	Période de calage	Critère de Nash (%)
W2714010	Chazelet	2	1951-2002	88.1
V0222010	Chamonix	6	1962-2003	83.4
X0500010	Embrun	6	1961-2002	91.6

Tableau 4-20. Description du calage du modèle hydrologique MORDOR sur les trois stations étudiées.

Stationnarité des variables climatologiques

Comme nous l'avons fait dans le cas des crues pluviales du Nord-Est, la première étape consiste à évaluer la stationnarité des données d'entrée du modèle. La Figure 4-43 présente les résultats du test LR-Reg appliqué à différentes variables, extraites des températures journalières sur une période commune 1961-2005. Ces variables sont les températures moyennes annuelles (T_m) et saisonnières (T_{hiver} , $T_{printemps}$, $T_{été}$, $T_{automne}$), les minima (T_N) et maxima (T_X) annuels des températures journalières, ainsi que le nombre annuel de jour de gel T_{gel} (température journalière négative) et le nombre annuel de jours chauds T_{chaud} . Pour cette dernière variable, étant donné les différences d'altitude entre les trois bassins, nous avons considéré qu'un jour chaud commençait à partir d'une température de 16°C pour l'Arve et la Romanche, et à partir de 20°C pour la Durance, ces valeurs correspondant approximativement au dernier décile de la distribution des températures journalières. Toutes ces variables présentent une tendance significative au réchauffement, la plupart de ces évolutions étant hautement significatives à 1%. Les variables les moins sensibles sont les températures automnales et les maxima annuels, avec des tendances significatives à 5%.

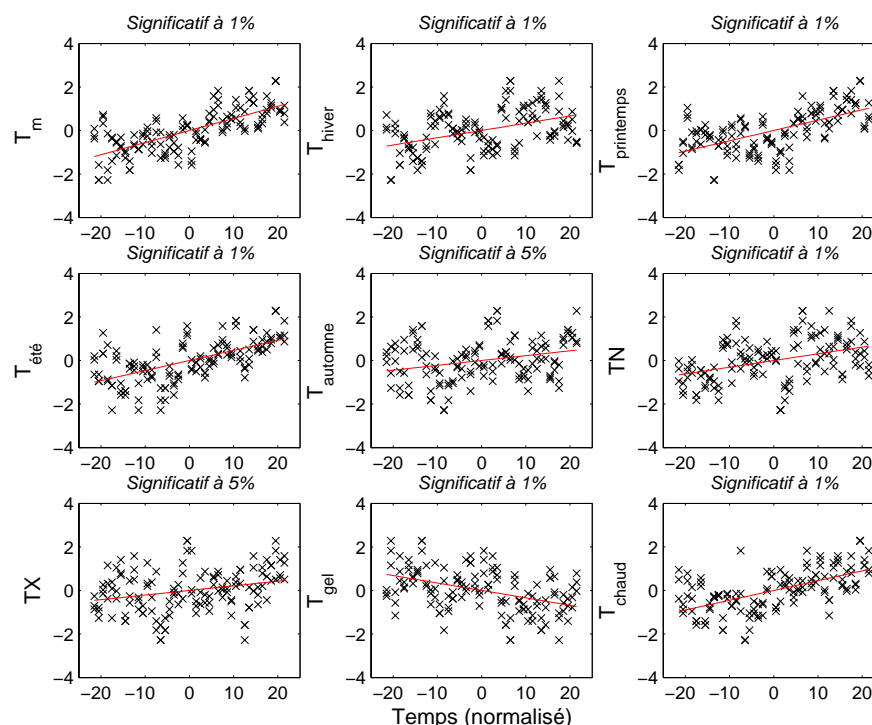


Figure 4-43. Evolutions observées sur les variables de températures relatives aux trois bassins versants étudiés.

Une étude similaire a été réalisée sur les variables pluviométriques. Les quatre premières variables sont identiques à celles évaluées à la section IV.2.1. Nous avons également évalué les moyennes saisonnières, qui influent d'avantage sur les écoulements liés à la neige que les précipitations extrêmes. Les résultats de cette analyse sont illustrés dans la Figure 4-44, et montrent qu'aucune tendance ne peut être détectée à l'échelle régionale sur la pluviométrie des bassins.

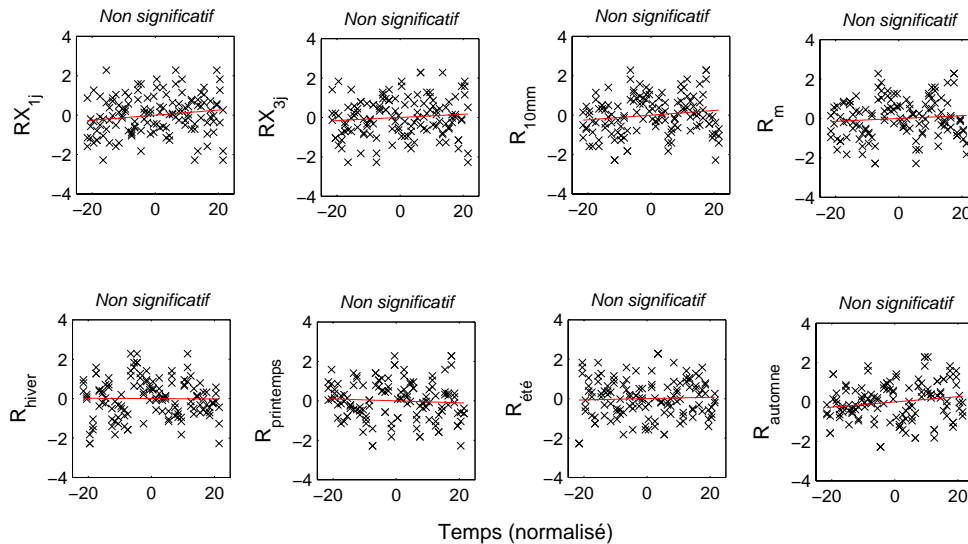


Figure 4-44. Evolutions observées sur les variables de précipitations relatives aux trois bassins versants étudiés.

Stationnarité des débits simulés

Les évolutions affectant les crues nivales ont été comparées sur les observations et sur les débits simulés. Trois variables qui présentaient un changement régionalement cohérent sur les Alpes du Nord ont été utilisées : la date du QCX_{10j} , la date du maximum du débit de base et le centre de masse. Ces trois variables évoluaient dans le sens d'une précocité accrue de l'onde de fonte. La période de données communes aux trois stations s'étend entre 1961 et 2002 pour les observations, et entre 1961 et 2005 pour les données simulées. Les résultats sont présentés dans la Figure 4-45. Il apparaît que les résultats sont cohérents entre observations et simulations, même si la significativité des tendances peut différer. Il semble donc que cette précocité constitue bien la réponse du bassin versant au réchauffement observé sur les séries de températures.

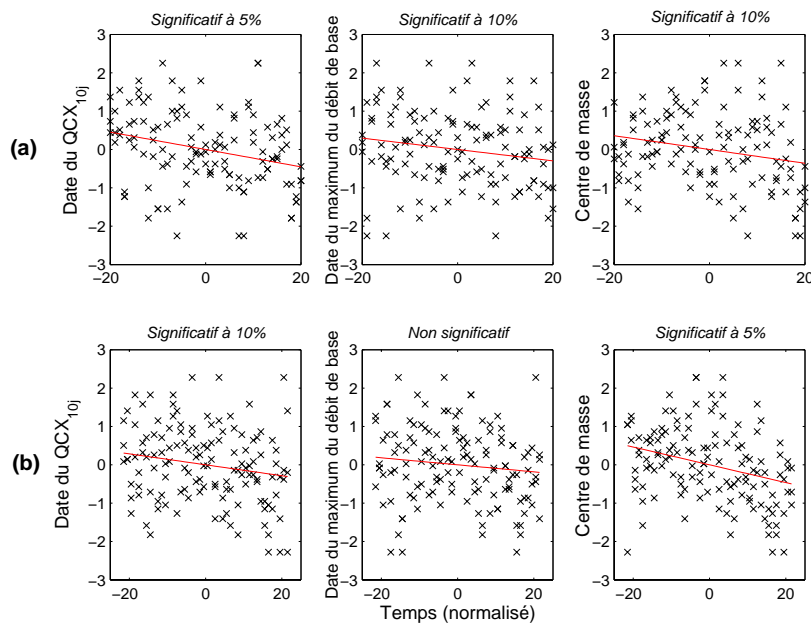


Figure 4-45. Evolutions des variables descriptives des crues nivales : (a) débits journaliers observés, (b) débits journaliers simulés par MORDOR.

Stationnarité de la fonte nivale simulée

Un des principaux avantages du modèle MORDOR est de fournir une estimation de la contribution des différents réservoirs à l'écoulement total. Plus précisément, l'évolution journalière du stock neigeux et des écoulements liés à la fonte peut être analysée. Nous nous affranchissons ainsi des difficultés que nous avons éprouvées lors de la création des variables nivales pour différencier les écoulements liés à des épisodes pluvieux de ceux liés à la fonte. L'utilisation de données de températures et de précipitations permet ainsi d'estimer plus précisément ces différentes contributions. A titre d'illustration, la Figure 4-46 décrit l'évolution du stock neigeux modélisée par MORDOR, alors que la Figure 4-47 décrit les contributions des différents réservoirs à l'écoulement total.

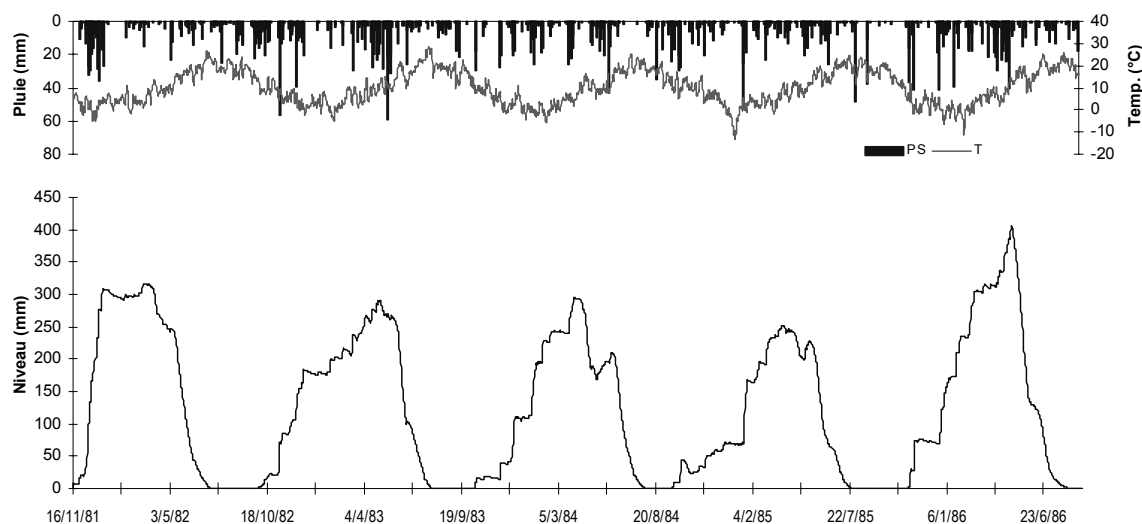


Figure 4-46. Evolution inter-annuelle du stock de neige pour la Durance à Serre-Ponçon. PS représente la pluie de bassin, T la température.

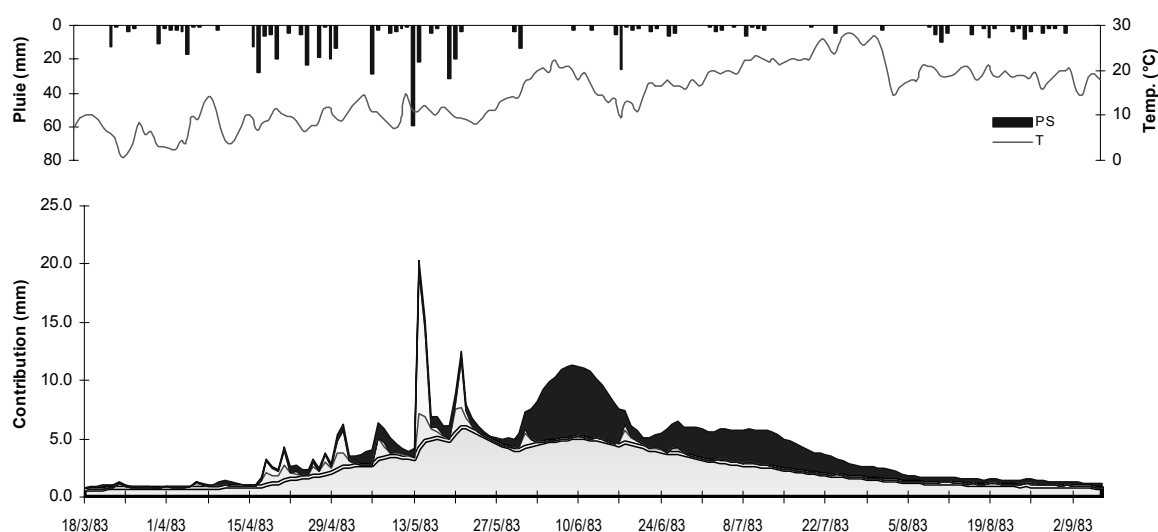


Figure 4-47. Contribution des différents réservoirs à l'écoulement total de la Durance à Serre-Ponçon. L'écoulement lié à la fonte est représenté en noir.

Trois variables ont été extraites à partir du stock neigeux. La date du maximum annuel fournit une estimation du début de la période de fonte, puisque à cette date, la fonte devient prépondérante par rapport au stockage. La valeur du maximum annuel fournit également une estimation de la plus forte quantité de neige couvrant le bassin au cours d'une année. Enfin, les différences entre deux hauteurs successives du stock sont égales aux quantités de neige précipitée (si la différence est positive) ou fondue (si la différence est négative) quotidiennement. La somme des incréments positifs fournit donc une estimation de la quantité de neige précipitée annuellement.

Cinq variables ont également été extraites des écoulements issus du réservoir « neige ». Le module de fonte est égal à la moyenne des écoulements nivaux. Les dates auxquelles 10%, 50% et 90% du volume nival ont été écoulés ont également été calculées. Ces dates peuvent être interprétées comme le début ($F_{10\%}$), le centre de masse ($F_{50\%}$) et la fin ($F_{90\%}$) des écoulements liés à la neige. Enfin, la durée de la période de fonte peut être estimée comme la différence entre $F_{90\%}$ et $F_{10\%}$.

Le test LR-Reg a été appliqué à chacune de ces variables, sur la période 1961-2005. Les résultats sont présentés dans la Figure 4-48. Aucune des trois variables relatives au stock neigeux ne présente de tendance significative. Ce résultat peut sembler inattendu étant donné l'importance du changement sur les températures. Plus particulièrement, avec des précipitations hivernales stationnaires, on pourrait s'attendre à ce que la part de précipitations tombant sous forme de neige diminue fortement. Cette diminution ne semble pourtant pas significative à l'échelle régionale. Précisons cependant que cette quantité neigeuse n'est pas mesurée, mais est estimée par le modèle sur la base des chroniques de précipitations et de températures. Inversement, les variables relatives aux écoulements nivaux présentent des tendances cohérentes avec le réchauffement observé : le début et le centre de masse de la fonte évoluent vers une plus grande précocité, alors que la fin de la fonte est plus tardive. En conséquence, la durée totale de la période de fonte tend à augmenter. Le retard observé sur la variable $F_{90\%}$ est lié à la présence de glaciers sur ces bassins versants, que le modèle considère comme des réservoirs infinis. Il est probable que des bassins d'altitude sans couverture glaciaire auraient plutôt eu tendance à écouler plus rapidement le stock de neige, entraînant un comportement différent de la fin de la fonte.

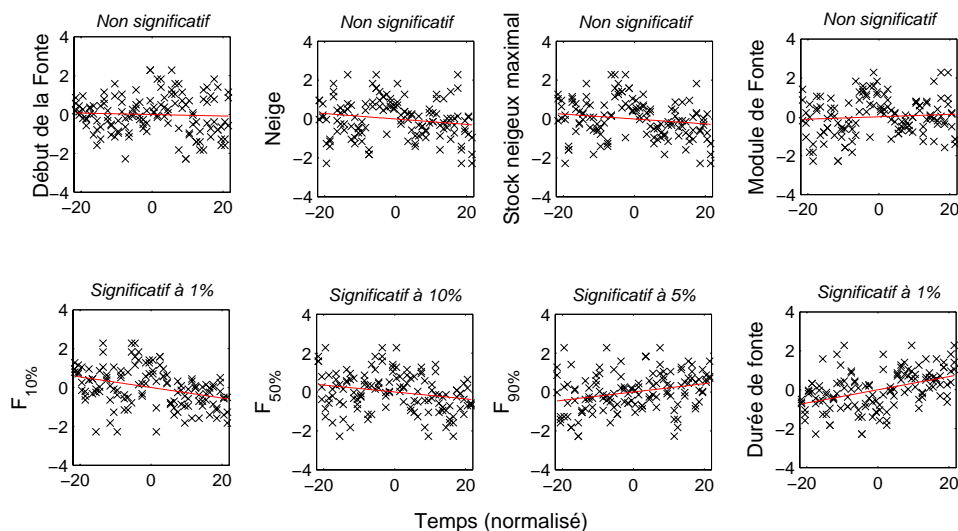


Figure 4-48. Evolutions sur les écoulements liés à la neige, d'après la segmentation effectuée par MORDOR.

SYNTHESE DU CHAPITRE 4

Ce chapitre a été consacré à l'analyse de la stationnarité des extrêmes hydrologiques en France. Dans un premier temps, l'analyse de l'échantillon complet de stations a révélé de nombreux changements dans les séries analysées. Une étude critique des résultats, menée en collaboration avec les gestionnaires de données, a permis de relier une partie de ces changements à des causes non climatiques, principalement des problèmes météorologiques ou des influences directes. Un nouvel échantillon épuré a donc été créé en retirant de l'analyse les stations pour lesquelles les causes de ruptures ont pu être établies.

Les résultats sur l'échantillon restreint ne montrent pas de tendance généralisée à l'échelle de la France, ni en régime, ni en basses eaux, ni en hautes eaux. Ainsi, il semblerait que l'aggravation des conséquences des inondations en France soit bien due à une augmentation de la vulnérabilité, et non de l'aléa. Certains changements localement significatifs demeurent, mais sont en nombre restreint et ne présentent guère de cohérence spatiale.

A l'échelle régionale, nous avons dans un premier temps évalué la significativité régionale à l'échelle de la France. Cette évaluation pourrait être légèrement biaisée à cause notamment de la présence de lacunes, qui impliquent de ne travailler que sur un sous-échantillon. Les résultats montrent que la significativité régionale n'est atteinte pour aucune des variables analysées, qui correspondent aux variables échantillonnées de manière annuelle.

La recherche de changements régionalement cohérents a été effectuée à une échelle plus petite, celle des régions hydro-climatiques définies au Chapitre 2. Certaines évolutions ont cette fois pu être mises en évidence. Parmi les résultats les plus robustes, on peut citer :

- ✓ Des étiages moins sévères dans les Alpes. De plus, dans les Alpes du Nord, l'onde de fonte nivale semble plus précoce.
- ✓ Une augmentation du module nival des stations glaciaires, accompagnée d'une précocité accrue de la date du maximum du débit de base.
- ✓ Une aggravation des étiages dans les Pyrénées et le Pays Basque. Les crues pluviales semblent également en diminution dans les Pyrénées.
- ✓ Une augmentation des maxima annuels dans la région continentale du Nord-Est.
- ✓ Une aggravation des crues en région océanique (Nord de la France), apparemment liée au remplissage des nappes d'accompagnement.

Ces évolutions ont finalement été mises en relation avec l'évolution des pluies et des températures. Dans un premier temps, nous avons procédé de manière empirique, en comparant nos résultats à ceux du projet IMFREX. Cette comparaison s'est cependant révélée difficile, principalement à cause de la forte non linéarité des relations pluies-températures-débits, mais aussi à cause d'un manque de couverture spatiale pour les résultats relatifs aux variables extrêmes. Dans un second temps, nous avons donc utilisé la modélisation hydrologique sur deux cas particuliers. Le modèle GR4J a ainsi permis de montrer que l'évolution observée des débits semblait cohérente avec le comportement des pluies pour la région continentale (Nord-Est), la transformation pluie-débit n'ayant apparemment pas subi de fortes modifications. Le modèle MORDOR a lui été utilisé pour l'analyse des crues nivales des Alpes, et a permis de mettre en évidence un lien apparent entre l'augmentation des températures et une modification de la saisonnalité des écoulements liés à la fonte.

CHAPITRE 5. LA GESTION DU RISQUE EN CONTEXTE NON STATIONNAIRE

« Le principe de la génération de petites quantités d'improbabilité finie par simple raccordement des circuits logiques d'un cerveau submagnétique Patastis 51 avec un conspirateur vectoriel mis en suspension dans un puissant émetteur de mouvement brownien (mettons, par exemple, une bonne tasse de thé brûlant), ce principe était certes parfaitement maîtrisé. ». Douglas Adams, le Guide du Routard Galactique.

La prédétermination des risques hydro-météorologiques utilise, entre autres méthodes, l'analyse statistique des événements passés. Le calcul de quantiles est donc basé sur l'hypothèse sous-jacente de stationnarité : ce qui s'est passé hier est représentatif de ce qui se passe aujourd'hui, voire de ce qui se passera demain. Si l'on accepte que les phénomènes physiques étudiés puissent évoluer avec le temps, alors cette hypothèse n'a plus lieu d'être, et il convient de mettre à jour les méthodes utilisées dans l'analyse fréquentielle. Cette mise à jour n'est pas sans poser quelques problèmes.

Le premier obstacle est d'ordre théorique. La théorie des valeurs extrêmes constitue en effet le cadre probabiliste pour le calcul des quantiles. Or, cette théorie suppose que les variables dont on extrait les caractéristiques extrêmes sont indépendantes et identiquement distribuées, ce qui n'est pas le cas de la plupart des variables étudiées (pluies, débits, températures au pas de temps journalier par exemple). En l'absence de dépendance à longue portée, il est possible de démontrer un théorème des valeurs extrêmes pour des données autocorrélées. Par contre, il n'existe pas de théorie aussi générale pour des variables non stationnaires : la théorie ne garantit donc pas le caractère universel des lois des valeurs

extrêmes pour décrire les extrema de variables non stationnaires. En l'absence d'un cadre totalement rigoureux, il convient donc d'adopter une démarche pragmatique, comme le suggère Coles (2001), en utilisant les lois des valeurs extrêmes comme modèle général, et en supposant que ce sont les paramètres de ces distributions qui évoluent avec le temps.

Dans ce contexte, nous commencerons par discuter de la notion de période retour dans un cadre évolutif. Issu de la théorie du renouvellement (Cox, 1966), ce concept est abondamment utilisé en gestion des risques pour graduer le caractère exceptionnel d'un événement. Il est censé être plus « parlant » pour des gestionnaires ou des décideurs que la notion abstraite de probabilité. Nous verrons cependant que dans le cadre non stationnaire, son interprétation peut prêter à confusion, et qu'elle n'est de plus pas forcément définie.

Du point de vue pratique, il est possible d'utiliser les modèles présentés au Chapitre 3 dans le cadre des tests LR : la théorie de la vraisemblance permet en effet d'estimer sans difficulté des modèles avec covariables, et le test permet de choisir le modèle le plus adapté aux données. Ce type d'approche a été proposé récemment par plusieurs auteurs (Strupczewski et Kaczmarek, 2001; Strupczewski *et al.*, 2001; Cunderlik et Burn, 2003). Cependant, choisir un unique modèle implique de prendre le risque que la prédétermination soit erronée si le modèle choisi n'est pas le bon –si tant est qu'un « vrai » modèle existe... De plus, à notre connaissance, l'effet de nombreux facteurs (déboisement, changement d'occupation des sols, changement climatique) sur les régimes hydrologiques extrêmes reste largement méconnu. La stationnarité des séries devrait donc être considérée comme une incertitude supplémentaire à inclure dans l'analyse. L'analyse Bayésienne constitue un outil adapté à cet objectif, et a été développée dans le cadre de séries Gaussiennes de débits par Perrault *et al.* (2000; 2000a; 2000b; 2000c). Nous présenterons une généralisation de ce travail adaptée au contexte des valeurs extrêmes.

L'analyse fréquentielle régionale doit également être adaptée pour prendre en compte d'éventuelles évolutions temporelles. De plus, dans l'optique d'un changement climatique, on s'attend à une certaine cohérence des évolutions à l'intérieur de régions homogènes du point de vue hydro-climatique. Nous présenterons dans une troisième section les modèles que nous avons développés pour décrire les extrêmes hydrologiques au niveau régional dans un contexte non stationnaire. Un certain nombre de limitations apparaissent à l'échelle régionale, essentiellement liées à la difficulté de décrire et prendre en compte les dépendances spatiales entre valeurs extrêmes. Nous proposerons quelques pistes pour tenter de prendre en compte ce phénomène.

I. La notion de période de retour dans un contexte non stationnaire

Dans le domaine de la gestion des risques, une période de retour s'interprète généralement de la manière suivante : « une valeur x est de période de retour T si le temps séparant deux dépassement successifs de x vaut en moyenne T ». Dans le cas stationnaire, le terme « en moyenne » peut être interprété temporellement, ce qui conduit à la traduction « une valeur de période de retour T ans est dépassée en moyenne tous les T ans ». On voit bien que cette définition n'est plus adéquate si le phénomène étudié évolue avec le temps, puisque la période de retour associée à une valeur va elle aussi dépendre du temps ! Le terme « en moyenne » doit en fait être interprété, à un instant donné, comme une espérance mathématique. Mais le lien existant entre période de retour et probabilité au non-dépassement n'est alors plus évident à décrire formellement. En fait, si l'on se réfère à la théorie du renouvellement, la période de retour est définie comme l'espérance de la durée de vie résiduelle d'un composant. En d'autres termes, la période d'un retour d'un événement est l'espérance du temps d'attente jusqu'au prochain dépassement de cet événement. Nous allons détailler cette définition dans le cas de données obtenues par les méthodes d'échantillonnage MAXAN et SUP-SEUIL. Dans un souci de clarté, nous allons nous placer dans le cadre de l'analyse des crues. Cependant, toutes les considérations de ce paragraphe restent valable pour une autre variable d'étude (vent, pluie, etc.).

La théorie du renouvellement constitue un cadre très riche qui permet de considérer des processus de défaillance évoluant dans le temps, pour prendre notamment en compte le vieillissement des composants. Nous ne prétendons pas dans cette section dresser un état des lieux exhaustif de cette théorie dans le cas non stationnaire, mais seulement mettre en avant les ambiguïtés pouvant exister si les définitions utilisées sont de simples extrapolations du cas stationnaire. Précisons également que la période de retour n'est pas l'unique façon de quantifier les risques hydro-météorologiques. Ainsi, aux Pays-Bas, il est d'usage d'utiliser la probabilité d'être confronté à un événement extrême au moins une fois au cours d'une vie humaine (Mosselman, 2006), ce qui présente certains avantages en terme de communication et de culture du risque.

I.1. Données MAXAN

Soit x une valeur de débit, et X_t la variable aléatoire « débit maximum de l'année t ». Le temps étant ici discrétisé en années, la période de retour relative à x est en général définie comme l'espérance du temps d'attente jusqu'à une année qui contient un événement supérieur à x . Cette définition est donc légèrement différente de celle issue de la théorie du renouvellement (Lang, 1995). Notons $\theta^x(t)$ le temps d'attente, à partir du temps t , jusqu'à une année contenant un événement supérieur à x . Dans un premier temps, nous allons rappeler quelques résultats dans le cas stationnaire. La loi de cette variable aléatoire peut être déduite des relations suivantes :

$$\begin{aligned} Pr(\theta^x(t) = k) &= Pr(X_{t+1} \leq x, \dots, X_{t+k-1} \leq x, X_{t+k} > x) \\ &= \left(\prod_{i=1}^{k-1} F_{t+i}(x) \right) (1 - F_{t+k}(x)) \end{aligned} \quad (5-1)$$

Sous l'hypothèse de stationnarité, la loi de X_t ne dépend pas du temps, et la loi du temps d'attente est donc une loi géométrique de paramètre $F(x)$:

$$Pr(\theta^x(t) = k) = F^{k-1}(x)(1 - F(x)) \quad (5-2)$$

La période de retour associée à x est donc l'espérance de cette loi géométrique. Elle peut se calculer de la manière suivante (si $F(x) < 1$):

$$\begin{aligned} T(x) &= E(\theta^x(t)) = \sum_{k=1}^{+\infty} k F^{k-1}(x)(1 - F(x)) \\ &= (1 - F(x)) \sum_{k=1}^{+\infty} k F^{k-1}(x) \\ &= (1 - F(x)) \sum_{k=1}^{+\infty} (F^k(x))' \\ &= (1 - F(x)) \left(\sum_{k=1}^{+\infty} F^k(x) \right)' \\ &= (1 - F(x)) \left(\frac{1}{1 - F(x)} \right)' \\ &= (1 - F(x)) \frac{1}{(1 - F(x))^2} = \frac{1}{1 - F(x)} \end{aligned} \quad (5-3)$$

On retrouve bien dans ce cas la formule classique reliant la probabilité au non dépassement à la période de retour : un événement de probabilité au non dépassement 0.9 a une période de retour 10 ans.

Dans le cas non stationnaire, la formule (5-1) reste valide mais ne se simplifie plus, et la période de retour s'écrit de la manière suivante :

$$\begin{aligned} T(x) &= E(\theta^x(t)) = \sum_{k=1}^{+\infty} k \left(\prod_{i=1}^{k-1} F_{t+i}(x) \right) (1 - F_{t+k}(x)) \\ &= \sum_{k=1}^{+\infty} k \left(\prod_{i=1}^{k-1} F_{t+i}(x) \right) - \sum_{k=1}^{+\infty} k \left(\prod_{i=1}^k F_{t+i}(x) \right) \\ &= \sum_{k=1}^{+\infty} k \left(\prod_{i=1}^{k-1} F_{t+i}(x) \right) - \sum_{k=2}^{+\infty} (k-1) \left(\prod_{i=1}^{k-1} F_{t+i}(x) \right) \\ &= 1 + \sum_{k=2}^{+\infty} \left(\prod_{i=1}^{k-1} F_{t+i}(x) \right) \end{aligned} \quad (5-4)$$

La série apparaissant dans cette dernière expression n'est malheureusement pas assurée de converger : l'existence de la période de retour dépendra de l'évolution temporelle de la fonction de répartition. Intuitivement, si $F_t(x)$ tend trop rapidement vers 1 lorsque t tend vers l'infini, alors la décroissance vers zéro du produit $\prod F_{t+i}(x)$ pourrait être trop lente pour assurer la convergence de la série. De manière plus hydrologique, cela signifie que si les crues connaissent une diminution trop rapide, alors le temps d'attente jusqu'au dépassement d'une forte valeur peut avoir une espérance infinie. Il apparaît donc que la période de retour associée à une valeur x peut ne pas exister, alors même que la probabilité au non-dépassement de cette valeur en un temps donné est, elle, parfaitement définie. En conséquence, il semble préférable de conserver une échelle probabiliste pour qualifier le caractère exceptionnel d'événements dans ce contexte évolutif. Eventuellement, la formule (5-3) peut être appliquée pour associer

une période de retour à une probabilité en un temps t , ce qui supposerait qu'il n'y a plus d'évolution de la distribution des données après t . Une telle période de retour devrait donc s'interpréter de la manière suivante : « *En l'absence de toute évolution ultérieure*, le temps d'attente jusqu'au prochain dépassement de x a une espérance égale à T ».

I.2. Données SUP-SEUIL

Dans le cas de données supérieures à un seuil, le problème se complique puisque deux processus doivent être pris en compte : le processus d'occurrence des dépassements du seuil u , et les intensités de ces dépassements. Comme précédemment, $\theta^x(t)$ désigne le temps d'attente à partir de t jusqu'au prochain dépassement de x ($x > u$). $N_t^x(d)$ désigne le nombre de dépassements de x entre t et $t+d$, et X_t la valeur du débit supérieur au seuil. On s'intéresse donc encore une fois à la loi du temps d'attente. Cette loi est ici continue, et on peut écrire la relation :

$$\begin{aligned} Pr(\theta^x(t) \leq \delta) &= 1 - Pr(\theta^x(t) > \delta) \\ &= 1 - Pr(N_t^x(\delta) = 0) \end{aligned} \quad (5-5)$$

Une hypothèse classique consiste à supposer que le processus de dépassement du seuil u peut être décrit par une loi de Poisson. Sous l'hypothèse de stationnarité et d'indépendance entre les processus d'occurrence et d'intensité, on peut alors écrire :

$$\begin{aligned} Pr(N_t^x(\delta) = 0) &= \sum_{k=0}^{+\infty} Pr(N^u(\delta) = k) (Pr(X \leq x | X > u))^k \\ &= \sum_{k=0}^{+\infty} \frac{(\mu\delta)^k}{k!} \exp(-\mu\delta) (Pr(X \leq x | X > u))^k \\ &= \exp\{-\mu\delta(1-F(x))\} \end{aligned} \quad (5-6)$$

On en déduit alors aisément la fonction de répartition et la densité du temps d'attente :

$$\begin{aligned} G(\delta) &= Pr(\theta^x \leq \delta) = 1 - \exp\{-\mu\delta(1-F(x))\} \\ g(\delta) &= \mu(1-F(x)) \exp\{-\mu\delta(1-F(x))\} \end{aligned} \quad (5-7)$$

L'espérance du temps d'attente s'obtient alors par une intégration par partie :

$$\begin{aligned} T(x) &= E(\theta^x) = \int_0^{\infty} \delta g(\delta) d\delta \\ &= \int_0^{\infty} \delta \mu(1-F(x)) \exp\{-\mu\delta(1-F(x))\} d\delta \\ &= \left[-\delta \exp\{-\mu\delta(1-F(x))\} \right]_0^{\infty} + \int_0^{\infty} \exp\{-\mu\delta(1-F(x))\} d\delta \\ &= \left[-\frac{\exp\{-\mu\delta(1-F(x))\}}{\mu(1-F(x))} \right]_0^{\infty} \\ &= \frac{1}{\mu(1-F(x))} \end{aligned} \quad (5-8)$$

On retrouve bien dans le cas stationnaire la formule reliant la période de retour à la probabilité au non dépassement associée à l'intensité des valeurs supérieures au seuil.

Dans le cas non stationnaire, la première ligne de l'équation (5-6) ne peut plus s'écrire si simplement, puisque la probabilité pour qu'une valeur supérieure au seuil soit inférieure à x va dépendre de la date d'occurrence de cette valeur. Il faut donc conditionner par rapport à toutes les combinaisons possibles de ces k occurrences entre t et $t+\delta$. Une difficulté supplémentaire est liée au fait que les données, au pas de temps journalier par exemple, ne sont pas indépendantes : toutes les combinaisons ne sont donc pas équiprobables, et il faut tenir compte de la présence de clusters de valeurs élevées. Or, dans un contexte évolutif, si les intensités des dépassements ne sont pas stationnaires, les propriétés des clusters, leur taille notamment, n'ont pas de raison de l'être, ce qui compliquerait toute tentative de modélisation de ces clusters.

Devant ces difficultés supplémentaires pour estimer en toute rigueur une période de retour en contexte non stationnaire avec des données SUP-SEUIL, plusieurs solutions peuvent être envisagées. La première consiste, pour une valeur x fixée, à extraire des données un échantillon de temps d'attente entre deux dépassements de x . Cet échantillon peut être modélisé, par exemple, par une loi exponentielle non stationnaire, dont l'espérance, et donc la période de retour relative à x , peut être calculée en tout instant. Le désavantage de cette méthode est qu'elle ne donne que la période de retour relative à x , et qu'elle ne permet pas d'extrapoler au-delà des observations, puisque les intensités ne sont pas modélisées. La seconde solution consiste à adopter une autre définition pour la période de retour. Par exemple, comme nous l'avons fait pour les valeurs MAXAN, il est possible d'utiliser les formules du cas stationnaire, en y intégrant les dépendances au temps, ce qui implique que l'on suppose l'absence de toute évolution ultérieure. Une autre solution, reprise par Parey *et al.* (2006), consiste à définir le quantile de période de retour T ans comme la valeur dont le nombre de dépassements entre t et $t+T$ a une espérance égale à un. Cette définition permet d'effectuer les calculs dans le cas non stationnaire, mais ne coïncide pas forcément avec la définition classique de la période de retour.

Enfin, la dernière solution consiste à utiliser les probabilités plutôt que la période de retour. On peut par exemple chercher à calculer la probabilité p de non dépassement d'une valeur au cours d'une année donnée. Si on fait l'approximation qu'au cours d'une année, les évolutions temporelles sont négligeables (ce qui est acceptable puisque l'on modélise ici des évolutions à long terme), alors l'équation (5-6) peut être utilisée et on a :

$$\begin{aligned}
 p &= Pr(N_t^x(1 \text{ an}) = 0) \\
 &= \sum_{k=0}^{+\infty} Pr(N_t^u(1) = k) \left(Pr(X_t \leq x | X_t > u) \right)^k \\
 &= \sum_{k=0}^{+\infty} \frac{(\mu_t)^k}{k!} \exp(-\mu_t) \left(Pr(X_t \leq x | X_t > u) \right)^k \\
 &= \exp \left\{ -\mu_t \left(1 - F_t(x) \right) \right\}
 \end{aligned} \tag{5-9}$$

Autrement dit, étant donné une probabilité au non dépassement p pour une année donnée t , le quantile correspondant se calcule comme le quantile de probabilité $1+\log(p)/\mu_t$ de la loi des intensités des dépassements du seuil,

$$F_t(x) = 1 + \log(p) / \mu_t \tag{5-10}$$

Un lien peut être établi entre cette probabilité annuelle de non dépassement et une période de retour calculée à partir de données SUP-SEUIL en adaptant naïvement les formules du cas

stationnaire. Par exemple, considérons en une date t donnée une période de retour T . L'adaptation de la formule (5-8) pour prendre en compte la dépendance au temps peut se faire de la manière suivante :

$$T(x) = \frac{1}{\mu_t (1 - F_t(x))} \quad (5-11)$$

Autrement dit, le quantile associé à la période de retour T est le quantile de probabilité $1 - 1/(\mu_t T)$ de la loi des intensités des dépassements du seuil au temps t ,

$$F_t(x) = 1 - 1/(\mu_t T) \quad (5-12)$$

Quelle est alors la probabilité de dépasser ce quantile au cours de l'année t ? En injectant la formule (5-12) dans la formule (5-9), on obtient :

$$p = \exp\{-\mu_t (1 - F_t(x))\} = \exp(-1/T) \quad (5-13)$$

Dans le cas non stationnaire, le quantile calculé naïvement associé à la période de retour T peut être considéré comme le quantile ayant une probabilité $1 - \exp(-1/T)$ d'être dépassé au cours de l'année considérée. Ainsi, le quantile « décennal » ne correspond pas au quantile de probabilité annuelle au non dépassement 0.9. Cependant, cette différence est minime, comme le montre la Figure 5-1 : la probabilité annuelle au non dépassement associée à une période de retour T se rapproche rapidement de la formule classique $p = 1 - 1/T$ lorsque T augmente. En fait, cette différence existe aussi dans le cas stationnaire, et est liée à la méthode d'échantillonnage adoptée, qui implique dans le cas de données MAXAN une discrétisation du temps (Lang, 1995). Plus précisément, il est facile de montrer, à partir de la formule (5-13), que la période de retour T_X associée à des données MAXAN est liée à la période de retour T_S associée à des données SUP-SEUIL par la relation $T_X = T_S + 1/2 + O(1/T_S)$. En conséquence, pour des périodes de retour importantes (supérieures à 10 ans), le quantile de période de retour T calculé naïvement en une année t pourra s'interpréter classiquement comme la valeur ayant une probabilité $1/T$ d'être dépassée au cours de l'année considérée.

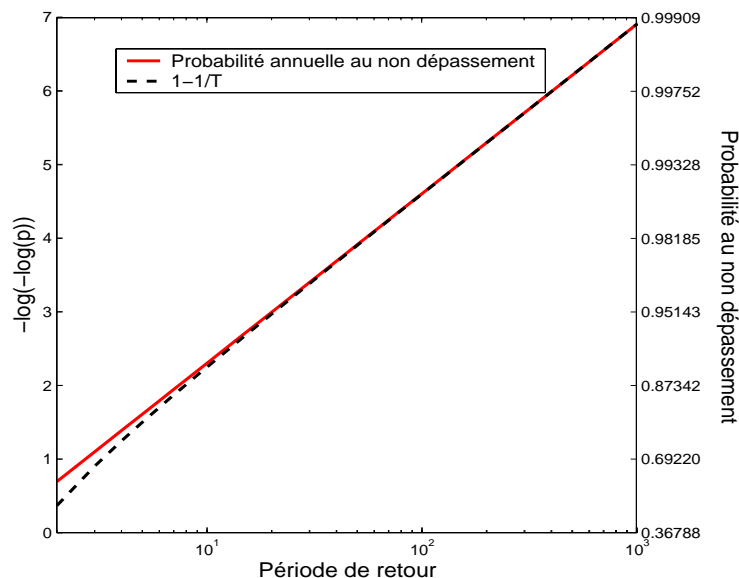


Figure 5-1. Lien entre probabilité au non dépassement et période de retour.

II. Echelle locale : l'analyse Bayésienne

Cette section décrit une méthode Bayésienne pour l'estimation de quantiles dans un contexte non stationnaire. Ce travail a fait l'objet d'une publication dans *Stochastic Environmental Research and Risk Assessment* (Renard *et al.*, 2006b), présentée en Annexe 1. Tous les détails techniques figurant dans cet article, nous nous contenterons de présenter les principales étapes nécessaires à la réalisation d'une telle analyse.

II.1. Principe général

II.1.1. La stationnarité comme une incertitude

Il est courant lors d'une analyse statistique de constater que plusieurs modèles sont d'une efficacité comparable pour décrire un jeu de données. Cependant, ces mêmes modèles peuvent donner des résultats bien différents en extrapolation (*i.e.* prévision de la variable d'intérêt avec des valeurs non observées des covariables).

Dans le cadre de la modélisation d'éventuelles tendances, un tel problème peut être illustré par l'exemple suivant. La série à analyser est constituée des valeurs maximales annuelles de débits journaliers sur la Moselle à Toul (3350 km²). Ces données sont supposées issues d'une loi GEV, dont le paramètre de position β évolue avec le temps de la manière suivante :

$$\begin{aligned} \beta(t) &= \beta_0 && M_0 \text{ (stationnaire)} \\ \beta(t) &= \beta_0 (1 + \beta_1 t) && M_1 \text{ (tendance linéaire)} \\ \beta(t) &= \beta_0 \exp(\beta_1 t) && M_2 \text{ (tendance exponentielle)} \end{aligned} \quad (5-14)$$

L'estimation de ces tendances est illustrée par la Figure 5-2. On y observe que les modèles à tendance linéaire et à tendance exponentielle fournissent une description presque équivalente de l'évolution de la moyenne des observations au cours du temps. Pourtant, dès que l'on cherche à extrapoler ces résultats vers le futur, ces deux modèles divergent fortement, non seulement en valeur moyenne, mais également au niveau des incertitudes d'estimation (Figure 5-3). Dans ces conditions, en l'absence d'arguments exogènes aux données pour privilégier une des deux descriptions, quel modèle choisir ?

Ce type de problème a conduit à l'élaboration des techniques de mélange de modèles : plutôt que de choisir un modèle, qui serait « meilleur » que les autres, tous les modèles sont pris en considération, et pondérés suivant leur aptitude à décrire les données. L'incertitude liée au choix du modèle est donc prise en compte, au moins partiellement (une collection de modèles pouvant difficilement être exhaustive). Un exemple de cette approche dans le domaine de l'hydrologie est fourni par Sharma et Lall (2004). Le cadre Bayésien est apparu être particulièrement adapté à cette approche, car il permet de définir une loi de probabilité sur l'ensemble des modèles, facilitant ainsi le calcul des pondérations, dans un cadre probabiliste unifié. Hoeting *et al.* (1999) proposent une revue de ces méthodes Bayésiennes de mélange de modèles (*Bayesian Model Averaging*). L'approche que nous allons développer dans les paragraphes suivants s'apparente à cette famille de méthodes.

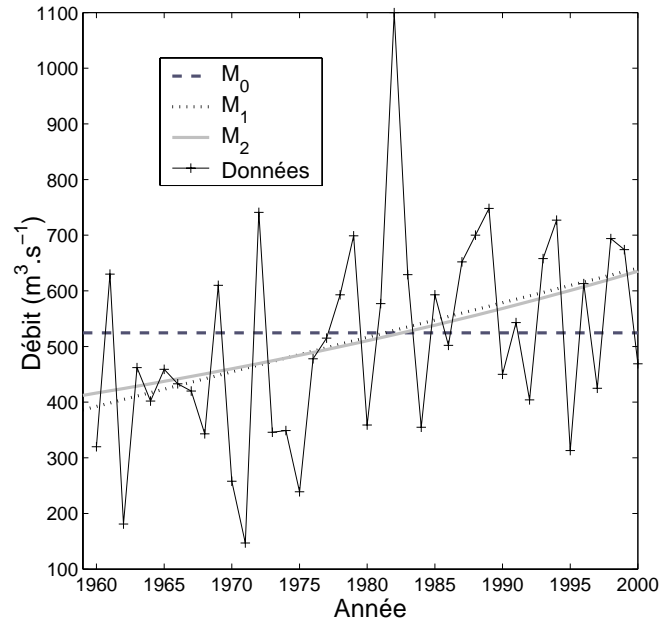


Figure 5-2. Moyenne des observations suivant les modèles M_0 (stationnaire), M_1 (tendance linéaire) et M_2 (tendance exponentielle).

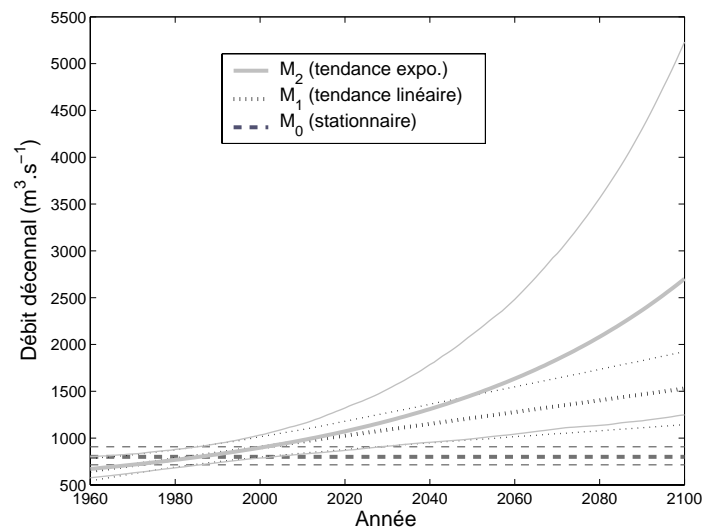


Figure 5-3. Illustration de la divergence des trois modèles en extrapolation. Les traits fins représentent les intervalles de confiance à 90%.

II.1.2. A priori et a posteriori

Quatre étapes sont nécessaires pour mettre en place une analyse Bayésienne (Coles, 1999), comme l'illustre la Figure 5-4 :

- ✓ Choix de la distribution des observations : comme dans le cadre de l'analyse fréquentielle classique, nous supposons que les données suivent une loi paramétrée par un vecteur θ . Cependant, θ n'est pas considéré ici comme un vecteur constant, dont nous souhaitons estimer la « vraie » valeur, mais comme une variable aléatoire, dont nous cherchons à connaître les caractéristiques. Sous l'hypothèse d'indépendance des observations, il est alors possible de calculer la vraisemblance du vecteur d'observations X , que nous noterons $p(X|\theta)$.
- ✓ Spécification de la distribution *a priori* des paramètres : cette étape doit traduire les connaissances exogènes aux données que nous souhaitons intégrer sous la forme d'une distribution $\pi(\theta)$. Cette distribution est paramétrée par un vecteur h , appelé vecteur des hyperparamètres. Les connaissances *a priori* doivent donc permettre d'affecter une valeur à chacun des hyperparamètres, sans utiliser les données.
- ✓ Application du théorème de Bayes : la distribution *a posteriori* des paramètres, $p(\theta|X)$, est calculée par la formule suivante :

$$p(\theta|X) = \frac{\pi(\theta)p(X|\theta)}{\int \pi(\theta)p(X|\theta)d\theta} \tag{5-15}$$

- ✓ Utilisation de la distribution *a posteriori* : l'inférence ne se limite pas ici à une valeur estimée suivant un critère prédéfini, mais à la distribution complète des quantités à estimer (paramètres du modèle ou quantiles, par exemple). Il est également possible, dans le cas où différentes distributions ont été supposées pour les données, d'évaluer la pertinence de ces modèles *via* leur probabilité *a posteriori*.

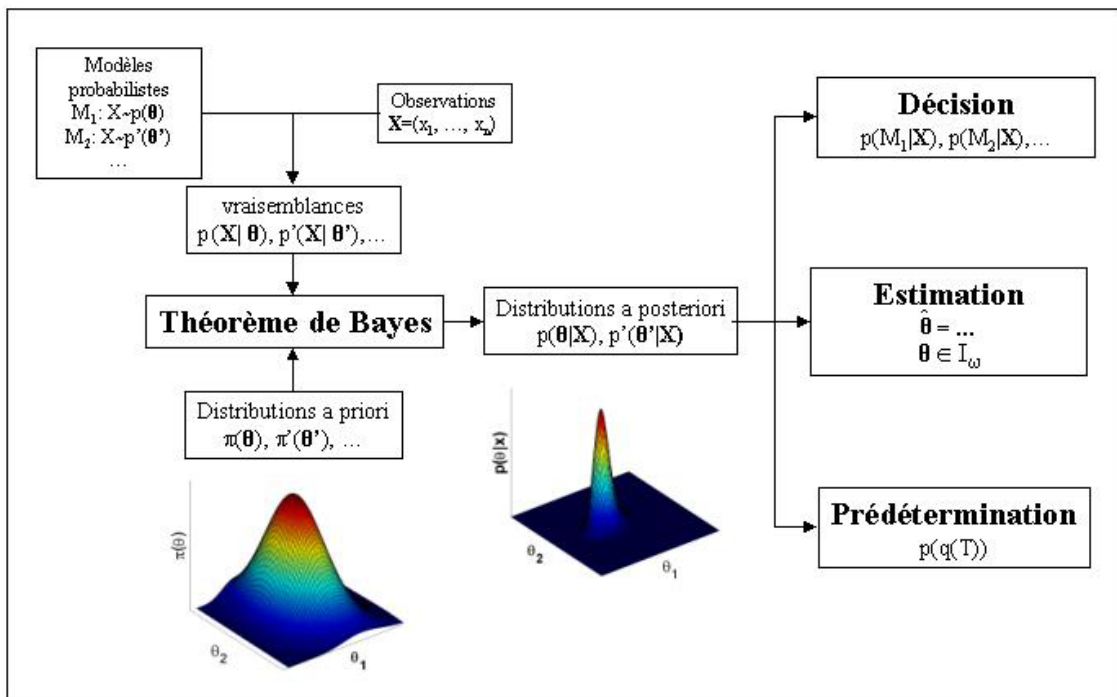


Figure 5-4. Principe de l'analyse Bayésienne.

L'approche Bayésienne permet donc, dans un cadre probabiliste unifié, de prendre en considération des connaissances *a priori* pour calculer la distribution des paramètres d'intérêt, donnant ainsi accès à une inférence plus complète que l'inférence ponctuelle classique via, par exemple, la méthode du maximum de vraisemblance. Cependant, ces avantages apparaissent aux yeux de certains statisticiens comme autant d'inconvénients, puisque les conclusions de l'analyse deviennent dépendantes des spécifications *a priori*. Une part de subjectivité serait ainsi introduite dans la démarche statistique. De manière plus pragmatique, une des principales difficultés est liée au caractère multidimensionnel du modèle : ainsi, la probabilité *a posteriori* et ses lois marginales ne sont en général pas calculables explicitement. Il est donc nécessaire de recourir à des méthodes de simulation stochastique pour effectuer l'inférence. Notons enfin que des intervalles de confiance Bayésiens ne peuvent pas s'interpréter de manière fréquentiste (intervalle qui contiendrait la vraie valeur du paramètre dans 90% des cas si l'expérience était reconduite de manière indépendante un grand nombre de fois). Il s'agit d'intervalles qui contiennent 90% de la densité *a posteriori*, cette dernière densité s'interprétant comme la mise à jour des connaissances *a priori* après observation des données. On parlera donc plutôt d'intervalles de crédibilité. De même, le terme d'« analyse fréquentielle » pour décrire le calcul de quantiles est inadéquat dans un contexte Bayésien. Nous ferons cependant l'abus de langage, puisque le but d'une telle analyse reste ici la prédétermination d'événements extrêmes.

L'ouvrage de Berger (1985) est en général considéré comme une référence dans la littérature hydrologique. Citons également Gelman *et al.* (1995), auquel nous nous sommes principalement référés lors de la mise en place des modèles présentés plus loin.

II.1.3. Difficultés numériques

Dans la plupart des situations, l'analyse Bayésienne ne fournit pas de formules explicites de la distribution *a posteriori* des paramètres d'intérêt. Une des rares exceptions concerne les distributions appartenant à la famille exponentielle, la loi Gaussienne notamment. Ceci explique pourquoi jusqu'à récemment, la plupart des études utilisant le formalisme Bayésien étaient relatives à des variables Gaussiennes (Chernoff et Zacks, 1964; Gardner, 1969; Lee et Heghinian, 1977; Buishand, 1982; Bernier, 1994). Au début des années 90, Gelfand et Smith (1990) rappellent l'intérêt des méthodes de Monte Carlo par Chaînes de Markov (MCMC). Ce type de méthode consiste à simuler des données multivariées à partir de densités de probabilité, ou éventuellement de fonctions proportionnelles à des densités (Metropolis et Ulam, 1949; Metropolis *et al.*, 1953; Hastings, 1970; Geman et Geman, 1984). L'application à l'analyse Bayésienne est alors apparue évidente : puisque la loi *a posteriori* est une fonction multidimensionnelle complexe, en général seulement connue à une constante multiplicative près (le dénominateur de l'équation (5-15) étant en général incalculable), la solution est de générer un grand nombre de données issues de la loi *a posteriori*, et de baser l'inférence sur cet échantillon simulé. Par exemple, les distributions marginales peuvent être estimées *via* les histogrammes des échantillons marginaux simulés, et les moyennes ou les médianes de ces échantillons peuvent être utilisées pour obtenir des estimations ponctuelles. La redécouverte de ces outils, associée à l'essor des moyens de calcul informatiques, a été le principal facteur expliquant l'explosion du nombre d'études utilisant le formalisme Bayésien au cours des années 90. Aujourd'hui, cette approche est devenue courante dans de très nombreux domaines d'application, y compris en hydrologie (Coles et Tawn, 1996; Khodja *et al.*, 1998; Lu et Berliner, 1999; Perreault *et al.*, 1999; Perreault *et al.*, 2000a, b; Perreault *et al.*, 2000c; Thyer et Kuczera, 2000; Wang, 2001; Thyer *et al.*, 2002; Coles et Pericchi, 2003; Coles *et al.*, 2003; Parent et Bernier, 2003; Perreault et Fortin, 2003; Thyer et Kuczera, 2003a, b; Tapsoba *et al.*, 2004; Cooley *et al.*, 2005; Moyeed et Clarke, 2005; Reis et Stedinger, 2005; Renard *et al.*, 2006b).

Malgré leur indéniable efficacité, les méthodes MCMC ne sont pas triviales à prendre en main, tant du point de vue théorique que du point de vue pratique. Par souci de clarté, nous ne développerons pas ces outils dans ce chapitre. Cependant, un travail important a été réalisé sur cet aspect (Garreta, 2005), et a fait l'objet d'un article actuellement soumis, présenté en Annexe 2 (Renard *et al.*, 2006a).

II.2. Exemple de mise en œuvre pour l'analyse des crues

Le cadre de travail présenté dans ce chapitre a été développé dans l'optique de la prédétermination des crues. Il se généralise cependant sans difficultés à l'analyse des valeurs extrêmes de nombreuses autres variables (précipitations, températures, vent, vagues, etc.). Tous les détails techniques figurent dans l'article présenté en Annexe 1.

II.2.1. Les modèles utilisés

Nous avons utilisé 12 modèles (Tableau 5-1), représentant deux types d'échantillonnages (MAXAN ou SUPSEUIL), deux types de distributions (2 ou 3 paramètres) et trois modèles de changement. Les modèles M_0 sont des modèles stationnaires classiques. Dans les modèles M_1 , un changement abrupt à une date inconnue est supposé sur les paramètres de la distribution (hors paramètre de forme). Enfin, les modèles M_2 sont caractérisés par la présence d'une tendance linéaire sur les paramètres (hors paramètre de forme). Les distributions utilisées sont issues de la théorie des valeurs extrêmes, et devraient donc s'appliquer dans de très nombreuses situations. Le choix d'un type de changement est plus arbitraire, et de nombreux autres modèles sont également possibles : tendances exponentielles ou polynomiales, ruptures multiples, etc. De même, le temps est utilisé ici comme covariable, mais d'autres descripteurs peuvent être employés. Par exemple, dans le cas d'un bassin soumis à d'importants prélèvements pour l'irrigation, il serait possible de lier l'évolution des paramètres à celle du volume de prélèvement. La méthodologie présentée ici s'adapte sans difficulté majeure à ces situations. Comme dans toute modélisation statistique, il convient cependant de rester parcimonieux sur la complexité du modèle, des difficultés d'estimation pouvant apparaître avec de trop nombreux paramètres, surtout dans ce contexte de valeurs extrêmes.

Echantillonnage	Distribution	Modèle		
		Stationnaire	Rupture	Tendance
SUP-SEUIL	Exponentielle	M_0^{exp} $\theta(t)=\theta$	M_1^{exp} $\lambda(t) = \begin{cases} \lambda_1 & \text{si } t \leq \tau \\ \lambda_2 & \text{si } t > \tau \end{cases}$	M_2^{exp} $\lambda(t) = \lambda_0 (1 + \lambda_1 t)$
	GPD	M_0^{gp} $\theta(t)=\theta$	M_1^{gp} $(\lambda(t), \xi(t)) = \begin{cases} (\lambda_1, \xi) & \text{si } t \leq \tau \\ (\lambda_2, \xi) & \text{si } t > \tau \end{cases}$	M_2^{gp} $\lambda(t) = \lambda_0 (1 + \lambda_1 t)$ $\xi(t) = \xi$
MAXAN	Gumbel	M_0^{gu} $\theta(t)=\theta$	M_1^{gu} $(\alpha(t), \beta(t)) = \begin{cases} (\alpha_1, \beta_1) & \text{si } t \leq \tau \\ (\alpha_2, \beta_2) & \text{si } t > \tau \end{cases}$	M_2^{gu} $\alpha(t) = \alpha_0 (1 + \alpha_1 t)$ $\beta(t) = \beta_0 (1 + \beta_1 t)$
	GEV	M_0^{gev} $\theta(t)=\theta$	M_1^{gev} $\begin{pmatrix} \alpha(t) \\ \beta(t) \\ \xi(t) \end{pmatrix}^T = \begin{cases} (\alpha_1, \beta_1, \xi) & \text{si } t \leq \tau \\ (\alpha_2, \beta_2, \xi) & \text{si } t > \tau \end{cases}$	M_2^{gev} $\alpha(t) = \alpha_0 (1 + \alpha_1 t)$ $\beta(t) = \beta_0 (1 + \beta_1 t)$ $\xi(t) = \xi$

Tableau 5-1. Modèles utilisés.

II.2.2. Spécification des *a priori*

Ces spécifications doivent être effectuées sans utiliser les données, et doivent refléter l'ensemble des connaissances disponibles avant l'analyse. De telles connaissances existent souvent, mais la difficulté consiste à les exprimer sous la forme d'une distribution multivariée de paramètres. Nous avons suivi la démarche proposée par Coles et Tawn (1996), qui consiste à effectuer les spécifications *a priori* sur des quantiles plutôt que directement sur les paramètres des distributions. En effet, le problème se ramène ainsi à l'estimation de quantiles sur un site non jaugeé (puisque les données ne doivent pas être utilisées à cette étape). La spécification se fait alors sur des différences de quantiles, car la relation d'ordre existant entre deux quantiles de probabilités $p_1 < p_2$ induit une structure de dépendance forte. L'hypothèse d'indépendance entre les différences de quantiles permet par contre de spécifier simplement une loi *a priori* multivariée. Cette hypothèse est émise par simplicité, mais ne devrait pas influencer fortement sur les résultats finaux (Coles et Tawn, 1996).

De nombreuses méthodes sont alors envisageables pour déterminer la distribution *a priori* : transfert d'information régionale, utilisation d'informations historiques, régressions multiples, modèles physiques, connaissances subjectives d'un expert, etc. Une fois la distribution *a priori* spécifiée sur ces quantiles, qui sont des fonctions des paramètres, un changement de variable permet de déduire la loi *a priori* des paramètres (Figure 5-5).

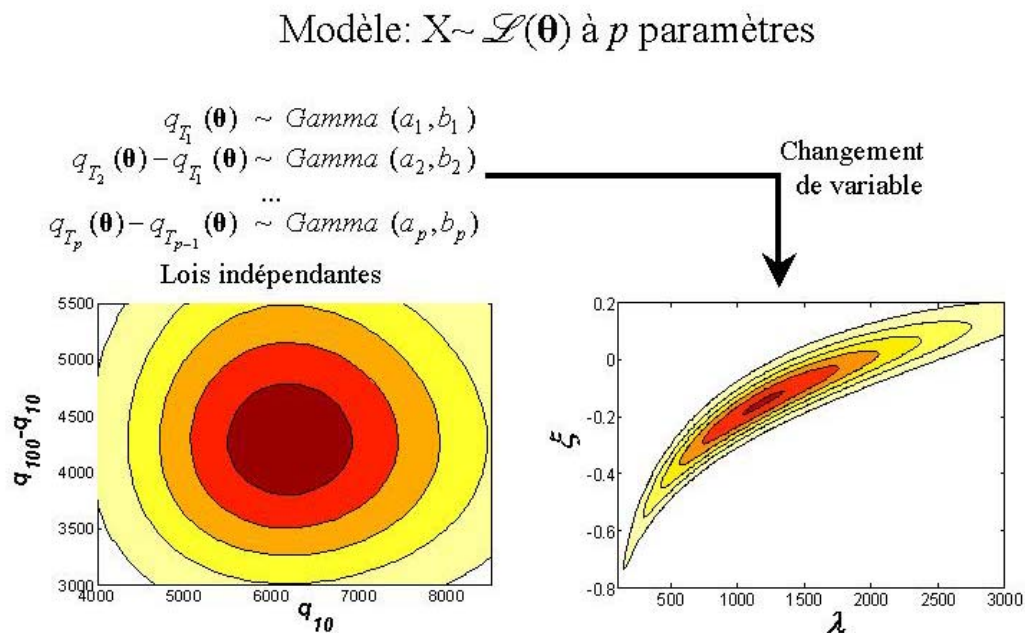


Figure 5-5. Spécifications *a priori* (exemple sur les paramètres d'échelle et de forme d'une loi de Pareto généralisée)

Signalons qu'en l'absence de toute connaissance, il est possible d'utiliser des lois *a priori* non informatives ou impropres. Cependant, l'utilisation de telles lois est problématique dans le contexte du choix de modèle (cf. section suivante). Nous supposons donc dans cette partie que les lois *a priori* sont propres, et qu'elles reflètent donc une réelle connaissance, aussi vague soit-elle.

II.2.3. Utilisation de la distribution *a posteriori*

La loi *a posteriori* des paramètres s’obtient simplement en appliquant le théorème de Bayes (Figure 5-6).

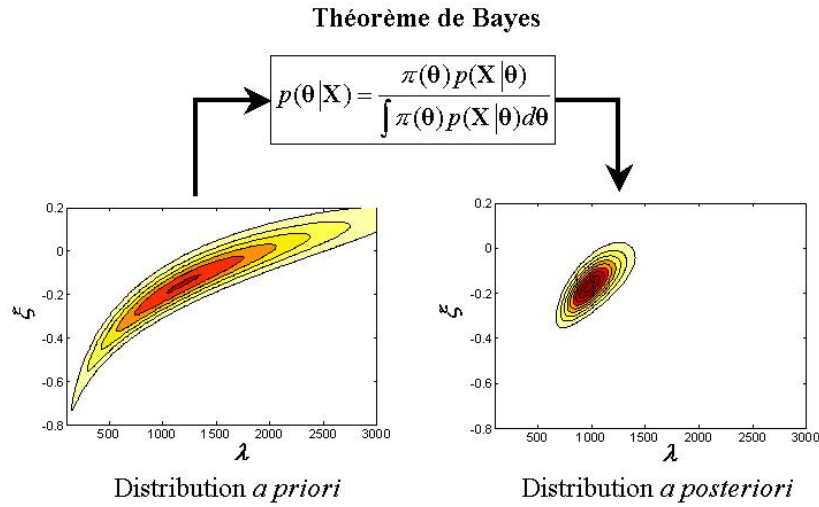


Figure 5-6. Application du théorème de Bayes.

Lois marginales

La première préoccupation lors d’une analyse probabiliste est en général d’estimer les paramètres du modèle. Nous avons vu que cette estimation était fournie par la distribution *a posteriori*. Cependant, cette distribution étant multidimensionnelle, il peut s’avérer plus avantageux de considérer les distributions marginales *a posteriori* de chaque paramètre (Figure 5-7). Ceci réclame d’intégrer la loi jointe, et donc d’utiliser une méthode MCMC lorsque cette intégration est de dimension trop élevée. Signalons que la loi exponentielle, parfois utilisée dans l’analyse des valeurs extrêmes pour décrire des dépassements de seuil ou des intervalles inter-occurrences, permet d’obtenir des formules explicites pour les lois marginales *a posteriori*, à condition d’utiliser des distributions *a priori* particulières, nommées distributions conjugués (cf. Annexe 1, section 6).

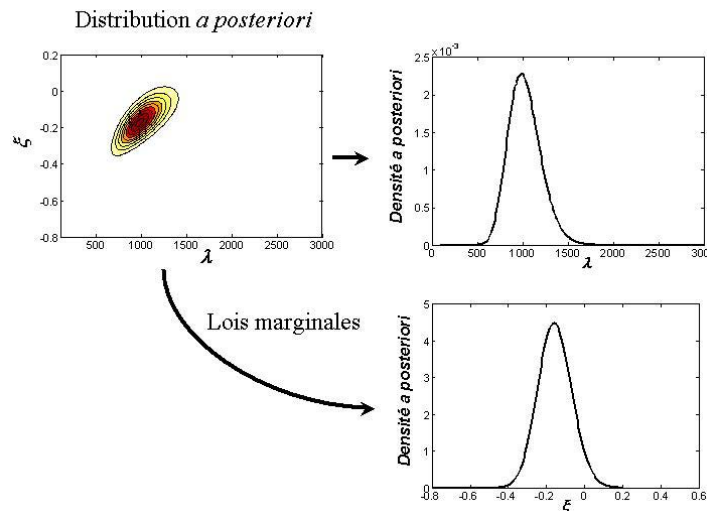


Figure 5-7. Lois marginales à partir de la distribution jointe *a posteriori*.

Coles et Powell (1996) ont montré que l’estimation des paramètres pouvait être plus précise par la méthode Bayésienne que par la méthode du maximum de vraisemblance, à la condition que les spécifications *a priori* soient suffisamment pertinentes.

Choix de modèle

La seconde application concerne la prise de décision par comparaison de modèles (Perreault, 2000; Perreault *et al.*, 2000b). Commençons par définir, dans le cadre d'un modèle M , la loi marginale des observations :

$$p(\mathbf{X} | M) = \int p(\mathbf{X}, \boldsymbol{\theta} | M) d\boldsymbol{\theta} = \int p(\mathbf{X} | \boldsymbol{\theta}, M) \pi(\boldsymbol{\theta} | M) d\boldsymbol{\theta} \quad (5-16)$$

Il s'agit en fait du dénominateur du théorème de Bayes, qu'il est possible de calculer même en dimension élevée grâce à une approche proposée par Chib (1995), encore une fois basée sur des méthodes MCMC. Cette dernière quantité est à la base de la démarche décisionnelle que nous allons adopter. Supposons en effet que nous considérons une collection de modèles $\Omega = \{M_1, \dots, M_q\} = \Omega_1 \cup \Omega_1^c$, où Ω_1 représente par exemple l'ensemble des modèles stationnaires, et Ω_1^c son complémentaire. Supposons également que nous ayons muni l'espace Ω d'une loi de probabilité *a priori*. L'application du théorème de Bayes permet alors de déterminer les probabilités *a posteriori* de chaque modèle, par la formule :

$$p(M_j | \mathbf{X}) = \frac{p(\mathbf{X} | M_j) \pi(M_j)}{\sum_{i=1}^q p(\mathbf{X} | M_i) \pi(M_i)} \quad (5-17)$$

Le facteur de Bayes (Kass et Raftery, 1995) entre les modèles M_i et M_j est alors défini par :

$$B_{i,j} = \left(\frac{p(M_i | \mathbf{X})}{p(M_j | \mathbf{X})} \right) / \left(\frac{\pi(M_i)}{\pi(M_j)} \right) = \frac{p(\mathbf{X} | M_i)}{p(\mathbf{X} | M_j)} \quad (5-18)$$

Plus généralement, nous pouvons également définir un facteur de Bayes composite entre les hypothèses Ω_1 et Ω_1^c :

$$B_{\Omega_1, \Omega_1^c} = \left(\frac{\sum_{M \in \Omega_1} p(M | \mathbf{X})}{\sum_{M \in \Omega_1^c} p(M | \mathbf{X})} \right) / \left(\frac{\sum_{M \in \Omega_1} \pi(M)}{\sum_{M \in \Omega_1^c} \pi(M)} \right) \quad (5-19)$$

Nous disposons ainsi d'une mesure de la pertinence d'une hypothèse par rapport à son alternative. Le facteur de Bayes fait d'ailleurs penser à un rapport de vraisemblance classique (test de déviance). Cependant, les termes du rapport ne sont pas calculés en remplaçant les paramètres par leurs estimateurs du maximum de vraisemblance, mais en intégrant suivant la loi *a priori* de ces paramètres. Plus généralement, cette démarche ne peut être considérée sous le même angle qu'un test statistique classique. La valeur d'un facteur de Bayes dépend en effet des spécifications *a priori*, en plus de la fluctuation d'échantillonnage. En ce sens, le terme de « risque » interprété comme un pourcentage limite d'erreur sur un grand nombre de rééchantillonnages est inapproprié. Les Bayésiens se basent plutôt sur la notion de « coûts ». Si c_I désigne le coût de l'erreur « décider la non-stationnarité à tort » et c_{II} le coût complémentaire « décider la stationnarité à tort » (par défaut, souvent égal à c_I), le critère de prise de décision sera alors :

$$\text{Choisir } \Omega_1 \text{ si } B_{\Omega_1, \Omega_1^c} > c_I / c_{II} \quad (5-20)$$

S'il est impossible dans ce cadre de contrôler les risques, au sens de la statistique classique, liés à la prise de décision, l'avantage de cette méthode est de prendre en compte à la fois les connaissances *a priori* et les incertitudes liées au choix d'un modèle. De plus, la notion de coût est souvent adaptée à la prise de décision dans un contexte opérationnel (cf. les exemples fournis par Perreault (2000) ou Bernier *et al.* (2000)). Une des limitations de cette approche est qu'elle est très sensible aux *a priori* spécifiés, à la fois pour les paramètres et pour les

modèles. De plus, l'utilisation d'*a priori* impropres est à proscrire, car la loi marginale des observations (équation (5-16)) n'est pas forcément définie. En conséquence, l'utilisation de cette méthode de choix de modèle *via* les facteurs de Bayes peut s'avérer problématique si l'information exogène aux données n'est que très peu disponible.

Analyse fréquentielle

La possibilité de calculer la probabilité *a posteriori* des différents modèles considérés permet d'effectuer un mélange pondéré de ces différents modèles. La technique utilisée est similaire à celle proposée par Perreault *et al.* (2000b). La loi *a posteriori* des quantiles $q_p(t)$, où p désigne la probabilité de non-dépassement et t le temps, peut être estimée par la procédure de simulation suivante :

Répéter N fois:

- ✓ Tirage au sort d'un modèle M^* , avec les probabilités $p(M_1|\mathbf{X}), \dots, p(M_k|\mathbf{X})$.
- ✓ Echantillonnage d'un vecteur de paramètres θ^* à partir de $p(\theta|\mathbf{X}, M^*)$.
- ✓ Calcul de $q_p(t)$, le quantile de probabilité p calculé dans le modèle M^* avec les paramètres θ^* au temps t .

L'échantillon ainsi constitué permet d'estimer la loi des quantiles, qui prend en compte l'incertitude d'échantillonnage mais également l'incertitude liée au choix d'un modèle, et donc celle liée à la stationnarité de la série. Cependant, les limitations signalées précédemment restent valables pour l'analyse fréquentielle, puisque le calcul des probabilités *a posteriori* des différents modèles est très dépendant des spécifications *a priori*, à la fois sur les paramètres et sur les modèles.

II.2.4. Cas d'étude

La méthode a été appliquée aux débits journaliers de la Drôme à Luc en Diois (194 km²), disponibles entre 1907 et 2003. L'échantillon de valeurs supérieures au seuil de 23 m³.s⁻¹ est constitué de 93 crues (Figure 5-8), ce qui, compte tenu des lacunes, conduit à environ une valeur par an en moyenne. Les dépassements de seuil ont été analysés en utilisant les lois exponentielles et de Pareto généralisée, tandis que les intervalles inter-occurrences ont été modélisés avec une loi exponentielle, comme décrit en section II.2.1.

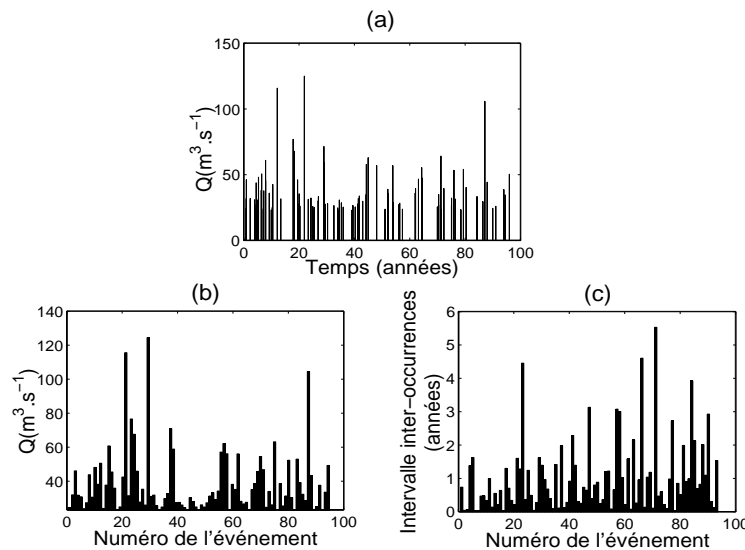


Figure 5-8. Données de crue pour la Drôme à Luc en Diois. (a) événements supérieurs au seuil, (b) dépassements du seuil, (c) intervalles inter-occurrences.

Distributions a priori

Dans un premier temps, des probabilités *a priori* doivent être affectées à chaque modèle considéré. Afin de ne privilégier ni l'hypothèse de stationnarité ni celle de changement, la distribution *a priori* des modèles est la suivante :

$$\text{Dépassements de seuil : } \begin{cases} \pi(M_0^{\text{exp}}) = \pi(M_0^{\text{gp}}) = 0.25 \\ \pi(M_1^{\text{exp}}) = \pi(M_2^{\text{exp}}) = \pi(M_1^{\text{gp}}) = \pi(M_2^{\text{gp}}) = 0.125 \end{cases} \quad (5-21)$$

$$\text{Inter-occurrences : } \begin{cases} \pi(M_0^{\text{exp}}) = 0.5 \\ \pi(M_1^{\text{exp}}) = \pi(M_2^{\text{exp}}) = 0.25 \end{cases} \quad (5-22)$$

Les connaissances *a priori* ont ensuite été exprimées sous forme de quantiles, en utilisant la méthodologie QdF-Bassin de référence (Prudhomme, 1995; Galéa et Prudhomme, 1997). Dans un premier temps, nous avons calculé un estimateur du débit décennal instantané grâce à la formule CRUPEDIX (CTGREF *et al.*, 1980-1982) :

$$q_{0.9} = S^{0.8} \left(\frac{r_{0.9}}{80} \right)^2 R \quad (5-23)$$

où S représente la superficie du bassin versant, $r_{0.9}$ est le quantile décennal de pluie journalière (quantile moyen sur le bassin), et R est un index régional. La durée caractéristique de crue D a ensuite été évaluée par la formule SOCOSE (CTGREF *et al.*, 1980-1982):

$$\log(D) = -0.69 + 0.32 \log(S) + 2.2 \left(\frac{\bar{r}}{r_{0.9} \cdot \bar{t}} \right)^{1/2} \quad (5-24)$$

où \bar{r} est la pluie annuelle moyenne et \bar{t} la température annuelle moyenne ramenée au niveau de la mer. Tous les paramètres nécessaires à l'évaluation de ces régressions sont cartographiés au niveau de la France. Nous avons ensuite utilisé la méthode QdF pour transférer l'information disponible sur le quantile décennal instantané vers la durée journalière et les périodes de retour 10 et 100 ans, de la manière suivante (Prudhomme, 1995):

$$\begin{aligned} Q(T, d) &= (A \log(T) + B) q_{0.9} \text{ pour } T \leq 20 \text{ ans} \\ Q(T, d) &= Q(10, d) + \left(C \log \left(1 + \frac{A T - 10}{C} \right) \right) q_{0.9} \text{ pour } T > 20 \text{ ans} \end{aligned} \quad (5-25)$$

Dans cette formule, les paramètres A , B et C dépendent du rapport entre la durée cible (ici, la journée) et la durée caractéristique de crue D . Ils sont obtenus par comparaison du bassin d'étude à trois bassins de référence, représentatifs des grands types de genèse des crues. Le bassin de référence relatif à la Drôme à Luc en Diois avait déjà été choisi par Prudhomme (1995).

L'utilisation de ces différentes méthodes régionales fournit une estimation locale des quantiles décennaux et centennaux, au pas de temps journalier, $Q(10, 1j)$ et $Q(100, 1j)$. Comme indiqué à la section II.2.2. nous raisonnons ensuite sur des quantiles modifiés $q_1 = Q(10, 1j)$ et $q_2 = Q(100, 1j) - Q(10, 1j)$, supposés indépendants. Afin d'ajuster une loi Gamma à deux paramètres sur chacun de ces quantiles modifiés, il est nécessaire d'avoir une indication concernant la variabilité de ces estimateurs. Un intervalle de confiance à 90% pour l'estimateur fourni par la méthode CRUPEDIX est de la forme $[q/2 ; 2q]$. L'incertitude finale dans la chaîne de traitements effectués pour le calcul des quantiles modifiés est plus difficile à évaluer. Nous adoptons donc le même type d'intervalle de confiance, ce qui nous permet de calculer les hyperparamètres a et b de chaque loi Gamma pour lesquels la moyenne *a priori*

est égale à q_i et le quantile à 95% est égal à $2q_i$. Dans le cas des lois non stationnaires, la même approche est utilisée. En l'absence de connaissance *a priori* quant à l'existence d'un changement, les quantiles avant et après rupture sont supposés identiques (modèle de rupture), et les quantiles en deux années arbitraires t_1 et t_2 sont également supposés égaux (modèle à tendance).

Pour le modèle de dépassement exponentiel, la spécification d'*a priori* conjugués a été réalisée, afin de s'affranchir des méthodes MCMC. Pour ce faire, les hyperparamètres a et b de chaque loi Gamma inverse ont été calculés afin de respecter les conditions décrites précédemment (moyenne *a priori* égale à q_i et quantile à 95% égal à $2q_i$). La loi Gamma inverse portant directement sur les paramètres, la relation $q_p = u - \lambda \log(1-p)$ reliant les quantiles au paramètre de la loi exponentielle a été utilisée. Les hyperparamètres sont donc finalement choisis de sorte que la moyenne *a priori* soit égale à $(u - q_1) / \log(1-p_1)$ et le quantile à 95% soit égal à $(u - 2q_1) / \log(1-p_1)$.

Concernant la série des inter-occurrences (loi exponentielle), l'hyperparamètre a a été fixé à la même valeur que celle calculée précédemment pour les dépassements exponentiels, ce qui conduit au même coefficient de variation. L'hyperparamètre b a été calculé afin que la moyenne de la loi Gamma inverse soit égale à un, conduisant à un événement par an en moyenne.

Enfin, dans le cas des modèles exponentiels non stationnaires, à la fois pour les dépassements et les inter-occurrences, nous utilisons une loi *a priori* uniforme pour la date de rupture, et une loi Gaussienne $N(0 ; 0.05)$ pour la tendance.

L'ensemble des spécifications *a priori* est résumé dans le Tableau 5-2.

	Dépassements de seuil			Inter-occurrences	Paramètres de changement pour les modèles exponentiels	
	Exponentielle	Pareto Généralisée		Exponentielle	Date de rupture	Tendance
Paramètres	$\hat{\lambda}$ 16.025	\hat{q}_1 $60 \text{ m}^3 \cdot \text{s}^{-1}$	\hat{q}_2 $42 \text{ m}^3 \cdot \text{s}^{-1}$	$\hat{\lambda}$ 1	$\hat{\tau}$?	$\hat{\lambda}_1$?
Hyper-paramètres	$IG(\lambda a, b)$ $a=2.5$ $b=24$	$G(\tilde{q}_1 a, b)$ $a=3.6$ $b=3.6/60$	$G(\tilde{q}_2 a, b)$ $a=3.6$ $b=3.6/42$	$IG(\lambda a, b)$ $a=2.5$ $b=1.5$	uniforme sur $[1, \dots, n-1]$	$N(\lambda_1 m, s)$ $m=0$ $s=0.05$

Tableau 5-2. Spécifications *a priori*.

Résultats

Les probabilités *a posteriori* des différents modèles sont résumées dans le Tableau 5-3. Les intervalles inter-occurrences semblent non stationnaires, l'hypothèse la plus probable étant celle d'une tendance linéaire sur le paramètre. A l'inverse, il est plus difficile de trancher quant à la stationnarité des dépassements de seuil, avec un facteur de Bayes proche de un. Cependant, nous suspectons l'estimation de la probabilité du modèle M_1^{gp} de n'être que peu précise. En effet, aucune date de rupture précise n'apparaît *a posteriori* (Figure 5-9). Il se peut que la méthode proposée par Chib (1995) soit moins efficace dans le cas où la distribution ne présente pas de mode bien défini.

	M_0^{exp}	M_1^{exp}	M_2^{exp}	M_0^{gp}	M_1^{gp}	M_2^{gp}	Non-stationnaire vs. stationnaire	Pareto généralisée vs. exponentielle
Inter-occurrences	0.16	0.21	0.63	/	/	/	5.4	/
Dépassements	0.15	0.04	0.01	0.28	0.5	0.02	1.3	4.1

Tableau 5-3. Probabilités *a posteriori* des modèles utilisés, et facteurs de Bayes relatifs aux différentes hypothèses.

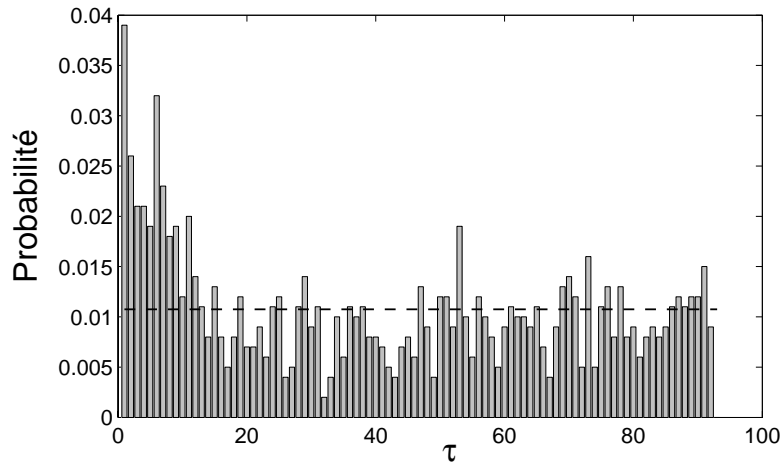


Figure 5-9. Distributions *a priori* (ligne pointillée) et *a posteriori* (barres) de la date de rupture sur les dépassements de seuil, dans le cadre d'une distribution de Pareto généralisée.

En guise d'illustration, l'estimation des différents modèles utilisés pour décrire les inter-occurrences est décrite dans la Figure 5-10. Dans le cas d'une rupture, on observe qu'une date se dégage assez nettement, aux alentours de la vingtième valeur (soit vers 1917). Les paramètres avant et après rupture correspondent à peu près à 1.5 événements par an (avant rupture) et 0.9 événements par an (après rupture), ce qui témoigne de la diminution en fréquence des crues supérieures à $23 \text{ m}^3 \cdot \text{s}^{-1}$. Le modèle à tendance, le plus probable *a posteriori*, conduit à une tendance de l'ordre de 0.018, ce qui peut se traduire par le fait que la fréquence des crues supérieures à $23 \text{ m}^3 \cdot \text{s}^{-1}$ est passée d'environ 2 à 0.8 événements par an entre 1907 et 2003.

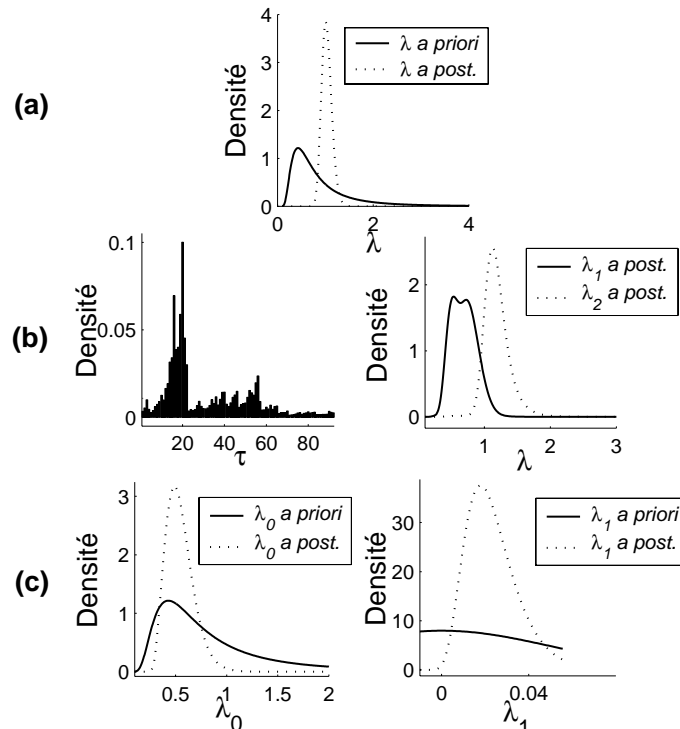


Figure 5-10. Distributions des paramètres des modèles utilisés pour décrire les inter-occurrences. (a) modèle stationnaire, (b) modèle de rupture, (c) modèle à tendance.

L'ensemble de ces résultats peut finalement être utilisé pour réaliser l'analyse fréquentielle non stationnaire des crues sur ce cours d'eau. La procédure présentée en section II.2.3. a été modifiée afin de prendre également en compte les intervalles inter-occurrences Y , que nous supposons indépendants des dépassements de seuil X :

Répéter N fois:

- ✓ Tirage au sort d'un modèle M^* pour les dépassements, avec probabilités $p(M_1|X), \dots, p(M_k|X)$.
- ✓ Tirage au sort d'un vecteur de paramètres θ^* dans la distribution $p(\theta|X, M^*)$.
- ✓ Tirage au sort d'un modèle M^+ pour les inter-occurrences, avec probabilités $p(M_1|Y), \dots, p(M_k|Y)$.
- ✓ Tirage au sort d'un vecteur de paramètres θ^+ dans la distribution $p(\theta|Y, M^+)$.
- ✓ Calcul de $\mu(t)$ ($=1/\lambda(t)$), le nombre moyen d'événements par an dans le modèle M^+ au temps t , et avec les paramètres θ^+ .
- ✓ Calcul de la probabilité relative à la période de retour T au temps t : $p = 1 - 1/T\mu(t)$.
- ✓ Calcul de $q_p(t)$, le p -quantile dans le modèle M^* avec les paramètres θ^* au temps t .

Dans cette procédure, la relation entre période de retour et probabilité est une simple adaptation des formules du cas stationnaire : cela correspond donc à la définition « naïve » présentée en section I.2. Au delà de la période de retour 10 ans, on pourra considérer qu'une période de retour T correspond à une probabilité annuelle au dépassement $1/T$.

Les résultats de cette analyse fréquentielle sont résumés dans la Figure 5-11 et la Figure 5-12. Dans la première, on peut observer que la courbe des quantiles s'est décalée vers le bas entre 1907 et 2003, ce qui traduit la diminution du risque de crue liée à la baisse de fréquence des dépassements de seuil. Dans la seconde, en se focalisant sur une fréquence particulière, on peut observer l'évolution d'un quantile au cours du temps. A titre de comparaison, une analyse fréquentielle stationnaire classique, avec des inter-occurrences exponentielles et des dépassements de seuil suivant une loi de Pareto généralisée, conduit à une estimation du débit décennal journalier d'environ $60 \text{ m}^3 \cdot \text{s}^{-1}$, valeur qui demeure dans l'intervalle de crédibilité Bayésien au cours du temps.

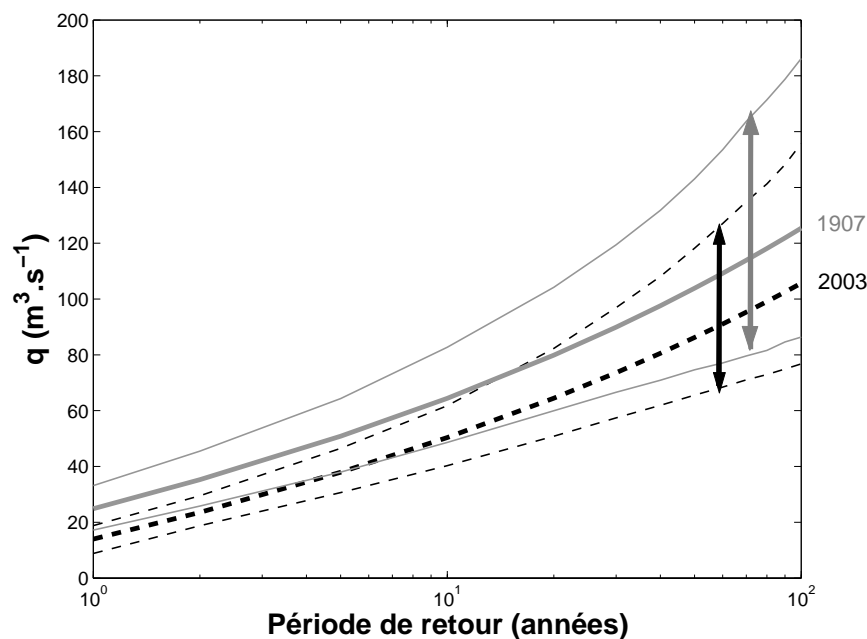


Figure 5-11. Courbe des quantiles et intervalles de crédibilité à 90% aux dates 1907 (gris) et 2003 (noir).

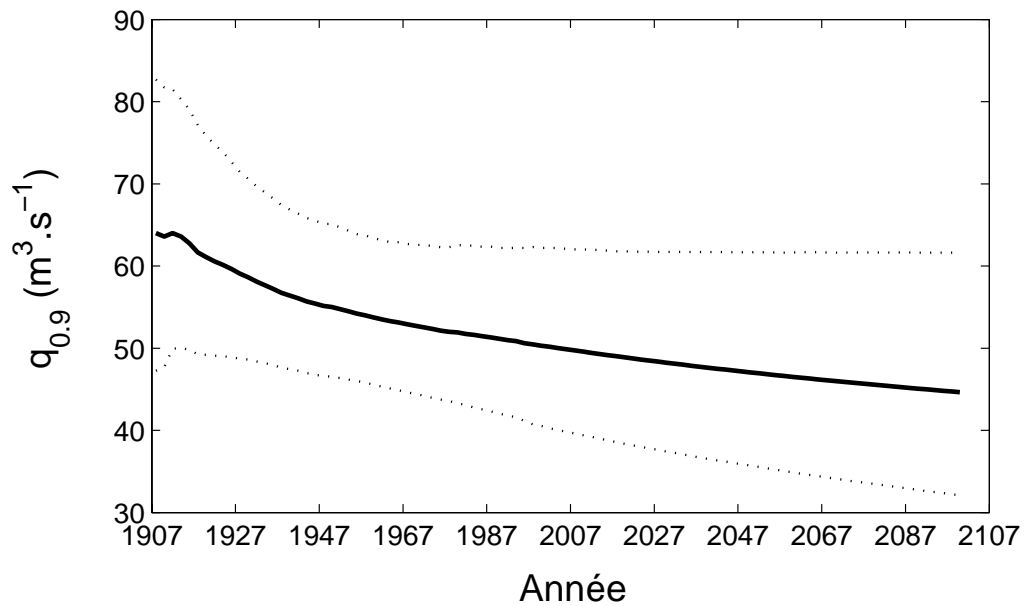


Figure 5-12. Trajectoire de la crue décennale et intervalle de crédibilité à 90%.

Dans cette dernière figure, précisons que l'extrapolation vers le futur de l'évolution des quantiles a avant tout un but illustratif. L'hypothèse sous-jacente est en effet que les tendances observées perdureront à l'identique, ce qui n'est évidemment pas garanti. L'utilisation d'autres covariables que le temps, avec une signification plus climatique, pourrait permettre d'améliorer ces projections. Malheureusement, les liaisons existant entre les crues et les variables climatiques (précipitations, NAO,...) sont en général beaucoup trop complexes pour être simplement décrites par les modèles présentés ici.

III. Analyse régionale non stationnaire

L'analyse régionale est un outil abondamment utilisé par les hydrologues, car elle permet de rendre l'estimation des quantiles plus fiable en élargissant l'échelle spatiale d'analyse. Etant donné la difficulté à décrire des évolutions modérées dans des variables extrêmes qui sont, par nature, soumises à une importante variabilité naturelle, on peut espérer que l'analyse régionale permettra également de mieux estimer des tendances régionales. Malheureusement, la mise œuvre d'une analyse régionale soulève, même dans le cas stationnaire, un certain nombre de difficultés. De nombreuses hypothèses doivent ainsi être faites quant à l'homogénéité de la région considérée, la distribution régionale des données ou la dépendance entre les sites étudiés. L'ajout d'une composante non stationnaire ne simplifie évidemment pas l'analyse. Dans cette section, nous allons décrire, au travers de quelques cas d'étude, des modèles régionaux non stationnaires. Dans un premier temps, nous supposerons que les données sont indépendantes spatialement, puis nous proposerons une première piste pour prendre en compte formellement la dépendance spatiale entre les sites d'étude.

III.1. Comment passer à l'échelle régionale ?

III.1.1. Principe général

Les principes de l'analyse régionale ont été décrits par Dalrymple (1960) *via* la méthode de l'indice de crue. L'auteur suppose que dans une région homogène, les échantillons sont tous issus de la même distribution, à un facteur d'échelle près. Plus précisément, la méthode peut être résumée comme suit :

- ✓ Création d'une région homogène, des points de vue statistique et hydrologique. De nombreuses méthodes peuvent être utilisées à cet effet (GREHYS, 1996b, a; Hosking et Wallis, 1997; Ouarda *et al.*, 2001), mais nous n'étudierons pas cet aspect du problème dans ce manuscrit.
- ✓ Extraction d'un échantillon de valeurs extrêmes (MAXAN ou SUP-SEUIL) en chaque site de la région.
- ✓ Chaque échantillon est divisé par un facteur d'échelle, appelé « indice de crue ». On choisit en général la moyenne ou la médiane des observations.
- ✓ Estimation des paramètres d'une distribution régionale (GEV ou Pareto généralisée en général), en utilisant toutes les valeurs normalisées.

La distribution régionale ainsi obtenue a de nombreuses applications. On peut distinguer trois cas de figure :

- ✓ Estimation en un site non jaugé : l'information régionale peut être transférée sur un site non jaugé. L'indice de crue du site cible devra être estimé à l'aide de régressions multiples sur des facteurs climatiques ou topographiques par exemple (superficie du bassin versant, pente moyenne, pluies de bassin, etc.).
- ✓ Estimation en un site peu jaugé : les données du site cible peuvent cette fois être utilisées pour estimer l'indice de crue.
- ✓ Estimation régionale : la distribution régionale peut être étudiée en tant que telle. Le principal intérêt de l'analyse régionale est alors de réduire l'incertitude d'échantillonnage, en utilisant plus de données pour l'estimation. C'est dans cette optique que nous nous plaçons, puisque nous allons nous intéresser à l'estimation d'une tendance dans un modèle régional.

Bien que la méthode de l'indice de crue ait évolué au cours des années, notamment au niveau des méthodes d'estimation des paramètres, le cadre général reste couramment utilisé. La méthode n'est pourtant pas exempte de limitations. Tout d'abord, la création d'une région homogène à l'intérieure de laquelle les hypothèses de l'indice de crue sont respectées peut s'avérer problématique. De nombreux progrès ont néanmoins été réalisés dans cette problématique, avec une évolution progressive depuis l'utilisation de régions géographiques jusqu'à la création de régions non contiguës, utilisant la notion de voisinage hydrologique (Ouarda *et al.*, 1999). De plus, cette limitation n'est pas propre à la méthode de l'indice de crue : quel que soit le modèle régional utilisé, la pertinence du modèle dans la région devra être vérifiée. Par contre, la normalisation des valeurs des échantillons par la médiane ou la moyenne ne s'appuie sur aucune justification forte : pourquoi ne pas choisir le débit décennal, par exemple ? Quel que soit l'indice de crue retenu, son estimation sera entachée d'incertitudes, qui ne sont généralement pas prises en considération. De plus, dans le cas d'un site faiblement jaugé, la méthode de l'indice de crue ne donne que peu de poids aux données du site, qui sont pourtant les plus informatives. Hosking *et al.* (1985) ainsi que Javelle *et al.* (2002a) ont proposé des méthodes d'estimation utilisant des moments pondérés afin de prendre en compte ce phénomène. Ribatet *et al.* (2006) ont également récemment proposé un modèle Bayésien qui permet de surmonter ces limitations. Enfin, un dernier problème, commun à presque toutes les méthodes régionales, concerne la dépendance entre les sites, qui est en général ignorée. Or, la présence de fortes dépendances peut augmenter la variance des estimateurs, et donc annuler le bénéfice de l'analyse régionale. Ce dernier point sera étudié en détail ultérieurement.

III.1.2. Mise en œuvre dans un cadre non stationnaire

Nous avons abordé la construction d'un modèle régional avec trois objectifs principaux. Tout d'abord, dans l'optique d'éventuels effets du changement climatique, nous souhaitons inclure dans le modèle la notion de cohérence régionale des évolutions. Ensuite, nous souhaitons nous affranchir de l'étape de normalisation des échantillons locaux de l'indice de crue, car cette transformation ne permet pas de prendre en compte rigoureusement les incertitudes d'estimation. Enfin, nous aborderons le problème de la dépendance spatiale entre les sites de la région d'étude.

Un modèle régional non stationnaire a été proposé récemment par Cunderlik et Burn (2003). Il s'agit d'une généralisation de la méthode de l'indice de crue permettant de prendre en compte des non-stationnarités du second ordre, c'est à dire que seuls les deux premiers moments des distributions peuvent évoluer avec le temps. Cependant, cette généralisation comprend une étape de normalisation des échantillons locaux. De plus, les tendances sont estimées localement, et les auteurs s'intéressent à la significativité régionale, et non à la cohérence. Cunderlik et Ouarda (2006) ont eux proposé une généralisation du modèle QdF régional convergent (Javelle *et al.*, 2002b) au cas non stationnaire. Dans cette étude, la tendance est cette fois supposée identique sur toute la région homogène considérée. Cependant, l'étape de normalisation par l'indice de crue demeure, ce qui rend problématique la quantification des incertitudes.

L'approche que nous avons suivie consiste à considérer que la distribution des données en un site comprend des paramètres locaux, propres au site, et des paramètres régionaux, qui sont donc supposés identiques sur toute la région considérée. Quelques modèles possibles ont déjà été proposés au chapitre 3, dans le cadre de la détection de changements cohérents. A titre d'exemple, rappelons la définition du modèle M2 :

$$X_t^{(j)} \square GEV(\alpha_j, \beta_j(1 + \delta t), \xi) \quad (5-26)$$

Dans cette expression, $X_t^{(j)}$ désigne le débit maximum annuel en une année t pour le site j . La dépendance au temps est ici exprimée *via* le paramètre de position, qui s'écrit comme le produit d'une composante locale stationnaire β_j et d'une tendance régionale $(1+\delta t)$. Le paramètre d'échelle est lui supposé local et stationnaire, tandis que le paramètre de forme est supposé régional et stationnaire. Pour l'instant, nous supposons également l'indépendance spatiale des maxima annuels. Précisons que ce modèle est moins contraignant que celui de l'indice de crue, qui impliquerait que le rapport entre les paramètres de position et d'échelle soit constant sur toute la région.

La principale difficulté de ce type de modèle concerne l'estimation des paramètres. Les estimateurs des paramètres locaux et régionaux n'étant pas indépendants, il n'est pas possible d'estimer localement dans un premier temps α_j et β_j , puis les paramètres régionaux dans un second temps. Il faut donc estimer simultanément tous les paramètres, soit, pour une région contenant p sites, $2p+2$ paramètres. Les méthodes MCMC sont en général les plus indiquées pour l'estimation de tels modèles complexes. Nous avons donc encore une fois utilisé le formalisme Bayésien dans le cadre régional. En plus des avantages numériques apportés par les méthodes MCMC, il est possible d'accéder directement à la distribution des paramètres d'intérêt (et notamment de la tendance), sans recourir à des hypothèses asymptotiques. De plus, les avantages mis en avant pour l'analyse locale demeurent : la distribution *a priori* permet d'intégrer des informations exogènes aux données, et il est possible de prendre en compte l'incertitude liée au choix du modèle.

Enfin, signalons l'existence de modèles Bayésiens hiérarchiques, qui permettent d'assouplir les hypothèses faites pour construire le modèle régional (typiquement, l'égalité stricte sur toute la région des paramètres de tendance et de forme). Faute de temps, il n'a pas été possible d'évaluer ce type d'approche pour l'analyse régionale non stationnaire, mais les perspectives seraient certainement très intéressantes. A titre d'exemple, citons le travail de Perreault (2000) qui utilise un modèle hiérarchique pour étudier une rupture régionalement cohérente sur un ensemble de sites, dans le cas Gaussien. Ce modèle permet d'assouplir l'hypothèse que les dates de rupture sont strictement égales sur tous les sites. De même, Cooley (2005) développe un cadre régional stationnaire adapté aux valeurs extrêmes de précipitations. Cette fois-ci, le modèle hiérarchique est utilisé pour assouplir l'hypothèse d'égalité de certains paramètres sur la région, en considérant que ces paramètres doivent plutôt varier de manière continue spatialement, cette variation pouvant de plus s'expliquer par des covariables (l'altitude notamment).

III.2. Cas d'étude

Nous présentons une application du modèle régional présenté ci-dessus dans le cas de données de débits sur six cours d'eau situés dans le centre de la France. Ce cas d'étude est décrit en détail dans l'article présenté en Annexe 2 (Renard *et al.*, 2006a), dans lequel figurent notamment tous les détails techniques nécessaires à la réalisation de l'analyse.

III.2.1. Données et modèles

Six stations hydrométriques situées à l'Ouest du Massif Central ont été utilisées. La Figure 5-13 montre la localisation des stations, et le Tableau 5-4 en décrit les propriétés hydrologiques générales. Les données utilisées sont des débits moyens journaliers. Nous n'avons utilisé que les années pour lesquelles toutes les stations étaient exemptes de lacunes, soit 34 années entre 1961 et 2002, les années 1981, 1982, 1983, 1992, 1996, 1997 et 2001 étant exclues de l'analyse. Les valeurs maximales annuelles ont finalement été extraites pour chacune des six stations.

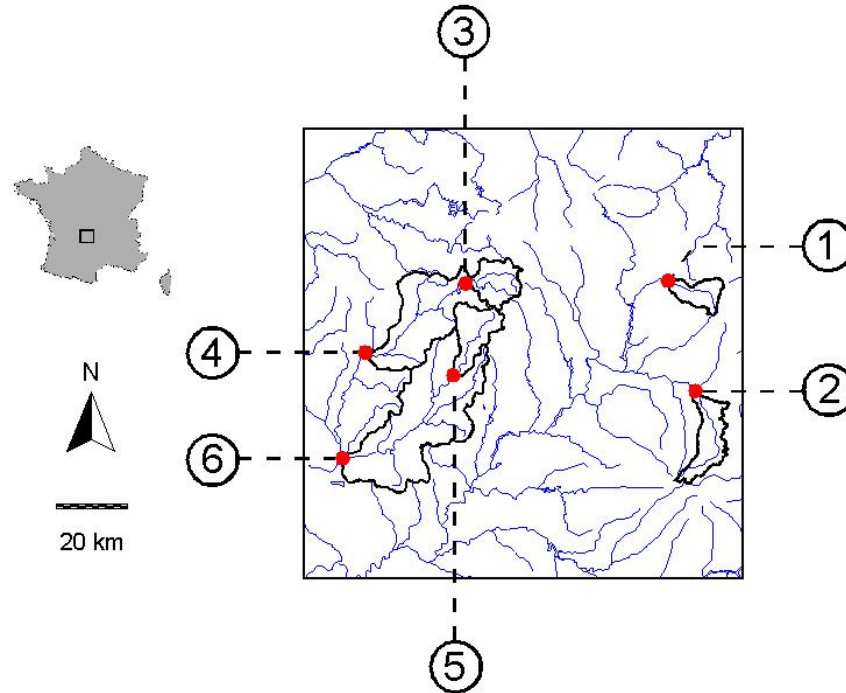


Figure 5-13. Localisation des six stations hydrométriques utilisées. (Figure réalisée par E. Sauquet)

Site	Rivière / Station	Superficie du bassin versant (km ²)	Module (m ³ .s ⁻¹)	Débit journalier décennal (m ³ .s ⁻¹)
1	Dordogne / Saint Sauves d'Auvergne	87	3.6	43
2	Santoire / Condat	172	4.71	74
3	Vézère / Bugeat	143	4.43	47
4	Vézère / Uzerche	601	15.1	136
5	Corrèze / Corrèze	168	5.6	61
6	Corrèze / Brive-la-Gaillarde	947	21.2	268

Tableau 5-4. Propriétés hydrologiques des six stations étudiées.

Le modèle M2 présenté précédemment a été utilisé pour décrire ces données. Un léger changement de paramétrage a été effectué pour prendre en compte la période d'étude. Explicitement, nous avons supposé que les maxima annuels d'un site j en une année t suivaient la loi GEV suivante : $X_t^{(j)} \sim GEV(\alpha_j, \beta_j(1 + \delta(t - 1960)), \xi)$. Sous l'hypothèse d'indépendance spatiale des données, il est alors possible de calculer la vraisemblance des observations :

$$\begin{aligned}
 p(\mathbf{X} | \boldsymbol{\theta}) &= p(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(6)} | (\alpha_1, \beta_1, \dots, \alpha_6, \beta_6, \xi, \delta)) \\
 &= \prod_{i=1}^n p(x_i^{(1)}, \dots, x_i^{(6)} | (\alpha_1, \beta_1, \dots, \alpha_6, \beta_6, \xi, \delta)) \\
 &= \prod_{i=1}^n \prod_{j=1}^6 GEV(x_i^{(j)}; (\alpha_j, \beta_j(1 + \delta(t_i - 1960)), \xi))
 \end{aligned} \tag{5-27}$$

La spécification des *a priori* a été effectuée en deux étapes. Des connaissances empiriques ont tout d'abord été utilisées pour les paramètres régionaux de forme et de tendance. En

hydrologie, le paramètre de forme des distributions de pluies ou de débits extrêmes est généralement compris entre -0.3 et 0.3 . Certaines valeurs sont parfois observées hors de cet intervalle, mais il est en général difficile de juger, étant donné la taille des chroniques, si cela relève d'un effet d'échantillonnage ou d'une réalité physique. Nous avons donc choisi une loi normale de moyenne zéro et d'écart-type 0.3 comme distribution *a priori* du paramètre de forme. Autrement dit, l'intervalle $[-0.6 ; 0.6]$ contient plus de 95% de la densité *a priori*. Cette distribution peut être comparée à l'*a priori* géophysique proposé par Martins et Stedinger (2000), qui est moins variable et a pour support l'intervalle $[-0.5 ; 0.5]$. Nous avons également utilisé une loi normale comme distribution *a priori* du paramètre de tendance. Devant l'absence de certitude quant aux effets du changement climatique sur les crues en France, la moyenne a été fixée à zéro. De plus, nous avons considéré qu'un éventuel impact ne pouvait raisonnablement pas se traduire par un changement de plus de 100% sur le paramètre de position entre 1960 et 2000. En utilisant la notation $\beta_{j,t} = \beta_j(1 + \delta(t-1960))$, la variance *a priori* de la tendance doit donc vérifier la relation suivante :

$$\begin{aligned} & \Pr\left(\left|\frac{\beta_{j,2000} - \beta_{j,1960}}{\beta_{j,1960}}\right| \leq 100\%\right) = 0.95 \\ \Leftrightarrow & \Pr\left(\left|\frac{\beta_j(1 + \delta(2000-1960)) - \beta_j(1 + \delta(1960-1960))}{\beta_j(1 + \delta(1960-1960))}\right| \leq 100\%\right) = 0.95 \quad (5-28) \\ \Leftrightarrow & \Pr(|40\delta| \leq 100\%) = 0.95 \\ \Leftrightarrow & \Pr(|\delta| \leq 1/40) = 0.95 \end{aligned}$$

Ainsi, en fixant l'écart-type de la loi normale *a priori* à $1/80$, le changement en 40 ans sur le paramètre de position appartient à l'intervalle $[-100\% ; 100\%]$ avec une probabilité d'environ 0.95.

Les connaissances empiriques sont plus difficilement utilisables pour les paramètres locaux. En effet, ces paramètres sont essentiellement reliés aux caractéristiques intrinsèques du bassin versant (superficie, géologie, topographie, pluviométrie, etc.). Nous avons donc utilisé les années non communes à toutes les stations pour spécifier les *a priori*. Plus précisément, une loi GEV stationnaire a été ajustée sur chaque site à l'aide de ces données non communes. La variabilité des estimateurs de (α_j, β_j) ainsi obtenus a été évaluée par Bootstrap, car la taille des échantillons ne permettait pas d'utiliser la variance asymptotique fournie par la matrice d'information de Fisher. Ainsi, 10000 répliques Bootstrap ont été effectuées, conduisant à autant d'estimations des paramètres α_j et β_j . Une loi Gamma a été ajustée à cet échantillon d'estimateurs. Les distributions Gamma *a priori* spécifiées sont détaillées dans le Tableau 5-5.

Finalement, nous avons supposé l'indépendance *a priori* de tous les paramètres, ce qui permet d'écrire la distribution *a priori* utilisée pour ce modèle régional :

$$\begin{aligned} \pi(\theta) &= \pi(\alpha_1, \beta_1, \dots, \alpha_6, \beta_6, \xi, \delta) \\ &= N(\xi; 0, 0.3^2) N(\delta; 0, (1/80)^2) \prod_{j=1}^6 Ga(\alpha_j; a_j^{(\alpha)}, b_j^{(\alpha)}) Ga(\beta_j; a_j^{(\beta)}, b_j^{(\beta)}) \end{aligned} \quad (5-29)$$

Site	Effectif disponible pour la spécification <i>a priori</i>	Paramètre de position β_j		Paramètre d'échelle α_j	
		<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>
1	39	20.03	2.46	41.88	2.68
2	20	21.09	1.72	84.11	2.42
3	18	28.39	2.20	65.12	2.26
4	51	66.96	2.10	244.19	3.19
5	26	21.51	1.32	51.40	1.47
6	48	72.73	1.39	346.00	2.22

Tableau 5-5. Spécifications *a priori* pour les paramètres locaux.

III.2.2. Méthodes MCMC

La distribution *a posteriori* des paramètres peut être calculée, à une constante près, comme le produit de la vraisemblance et de l'*a priori*. Les méthodes MCMC sont encore une fois indispensables pour utiliser cette distribution. Etant donné la dimension du problème, une attention particulière a été portée aux différents algorithmes disponibles afin de choisir le plus approprié. Une description détaillée des méthodes envisagées est fournie dans la première partie de l'article figurant en Annexe 2. Pour résumer, suite aux comparaisons empiriques effectuées par Garreta (2005), nous avons privilégié l'utilisation de l'algorithme de Metropolis (Metropolis et Ulam, 1949; Metropolis *et al.*, 1953) et de l'algorithme de Gibbs (Geman et Geman, 1984). Le premier algorithme est relativement simple à implémenter, et possède l'avantage de nécessiter un temps de calcul raisonnable, même en dimension élevée. La fonction de saut utilisée est une loi Gaussienne. Par contre, le choix des paramètres de fonctionnement de l'algorithme (point de départ et variance de la fonction de saut) doit être effectué avec précaution pour assurer la convergence dans un délai raisonnable. Inversement, l'utilisation de l'algorithme de Gibbs est plus délicate, car il ne peut pas en général être utilisé dans sa version initiale. En effet, à chaque itération de l'algorithme, une valeur doit être échantillonnée dans les distributions *a posteriori* conditionnelles, qui ne sont généralement pas connues. Une possibilité est d'inclure une étape de Metropolis pour obtenir une valeur issue de ces distributions, comme l'ont suggéré Gelman *et al.* (1996) et Coles et Tawn (1996). Le désavantage de cette étape est qu'elle rend le temps de calcul très dépendant de la dimension du problème. Par contre, il est possible d'adapter, au fur et à mesure des itérations, les caractéristiques de l'étape de Metropolis, ce qui a pour effet, au moins sur les simulations que nous avons effectuées, de rendre l'algorithme beaucoup moins sensible à d'éventuelles erreurs sur les points de départ ou les variances de saut. Ceci constitue un avantage certain, car il est souvent difficile d'effectuer ces réglages de départ, qui conditionnent grandement le bon fonctionnement des algorithmes. Sur la base de ces considérations, les échantillons issus de la loi *a posteriori* ont été simulés de la manière suivante :

- ✓ Pour chaque paramètre, le point de départ est égal à la moyenne de la distribution *a priori* spécifiée, et la variance de saut est égale à la variance de l'*a priori*.
- ✓ 1000 vecteurs de paramètres sont simulés par l'algorithme de Gibbs, avec une étape de Metropolis comme expliqué précédemment. Les 500 premières itérations sont considérées comme un temps de chauffe. A partir des 500 valeurs restantes, nous calculons la moyenne et la matrice de variance des paramètres. Précisons que les valeurs issues de l'échantillonnage de Gibbs ne seront pas utilisées pour l'inférence. En effet, d'un point de vue théorique, le fait d'adapter les caractéristiques de l'algorithme au cours des itérations peut poser des problèmes de convergence. Sur la base de quelques simulations effectuées, la convergence de l'algorithme n'a pu être mise en défaut. Cependant, étant donné que les propriétés d'ergodicité d'un tel algorithme mixte n'ont, à notre connaissance, jamais été

démontrées, nous préférons ne pas prendre de risque. En conséquence, ces premières itérations doivent être considérées comme un temps de pré-chauffe, qui permet d'estimer une variance de saut et des points de départ adéquats pour les prochaines itérations de l'algorithme de Metropolis.

- ✓ 100 000 vecteurs de paramètres sont finalement simulés par l'algorithme de Metropolis, dont les caractéristiques sont spécifiées en utilisant la moyenne et la matrice de variance calculées précédemment. Seules les 50 000 dernières valeurs sont utilisées pour l'inférence, les 50 000 premières étant considérées comme du temps de chauffe.

Une fois les paramètres simulés, la première étape consiste à vérifier la convergence des algorithmes MCMC. Une revue partielle des méthodes disponibles a été effectuée par El Adlouni *et al.* (2006). L'approche que nous avons utilisée est celle proposée par Gelman *et al.* (1995), qui consiste à utiliser plusieurs simulations effectuées en parallèle avec des points de départ différents, et à vérifier qu'après un certain nombre d'itérations, les différentes chaînes simulées proviennent bien d'une même distribution. Afin de quantifier cette vérification, Gelman *et al.* proposent de calculer la statistique \sqrt{R} , qui estime le rapport entre la variance totale et la variance intra-chaîne des données simulées en parallèle. Si la convergence est atteinte, alors la variance inter-chaîne devrait être proche de zéro, et donc la variance totale devrait être proche de la variance intra-chaîne, ce qui conduit à une statistique proche de 1. Les valeurs de cette statistique, calculée à partir de 10 chaînes simulées en parallèle, sont illustrées pour quelques paramètres dans la Figure 5-14. Dans le cas de l'algorithme de Gibbs, 200 itérations semblent suffire pour atteindre une valeur de la statistique très proche de un, ce qui semble confirmer que les 500 itérations de temps de chauffe sont suffisantes. De même, pour l'algorithme de Metropolis, environ 5000 itérations semblent suffisantes pour atteindre la convergence, ce qui valide amplement le temps de chauffe de 50 000 itérations que nous avons adopté.

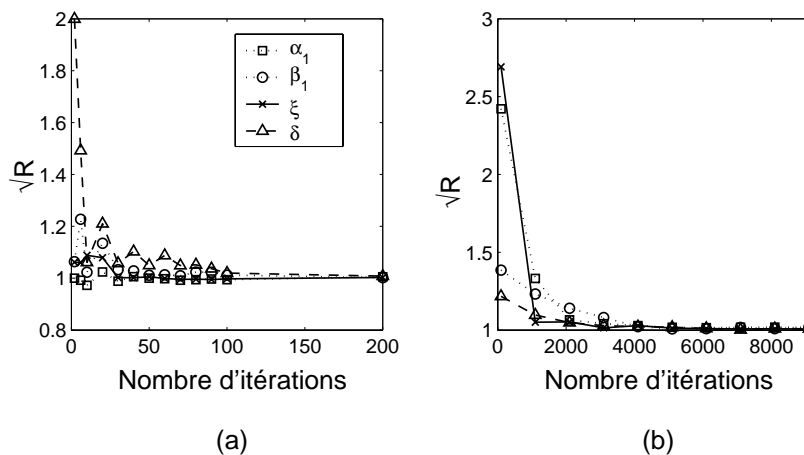


Figure 5-14. Valeur de la statistique $\sqrt{\hat{R}}$ pour différentes itérations : (a) algorithme de Gibbs, (b) algorithme de Metropolis.

III.2.3. Distribution *a posteriori*

Les distributions marginales *a posteriori* sont estimées à l'aide des histogrammes issus des échantillons MCMC. A titre d'exemple, la Figure 5-15 représente les résultats pour quatre paramètres du modèle. Dans les deux panneaux du haut, les paramètres d'échelle et de position du premier site sont représentés. On peut y observer que les distributions *a posteriori* sont cohérentes avec des estimations effectuées localement sur les mêmes données, avec la méthode du maximum de vraisemblance (traits verticaux). Le paramètre d'échelle α_1 semble

légèrement plus grand dans le cadre de l'estimation régionale, ce qui est certainement dû à l'influence de la distribution *a priori* (courbe en traits fin). Inversement, le paramètre de position β_1 est légèrement plus faible avec une estimation régionale qu'avec une estimation locale. Concernant les paramètres régionaux, l'hypothèse d'un paramètre de forme ξ identique sur toute la région peut sembler douteuse à la vue des estimations locales. Cependant, il convient de garder en mémoire que ces estimations ponctuelles sont entachées d'une incertitude importante : des investigations complémentaires sont donc nécessaires pour valider cette hypothèse, et seront décrites ultérieurement. A l'inverse, l'hypothèse d'une tendance commune à toutes les stations semble plus plausible, comme le confirme la Figure 5-16. On peut également remarquer que la variance *a posteriori* de la tendance régionale semble faible comparée à l'étendue des tendances estimées localement.

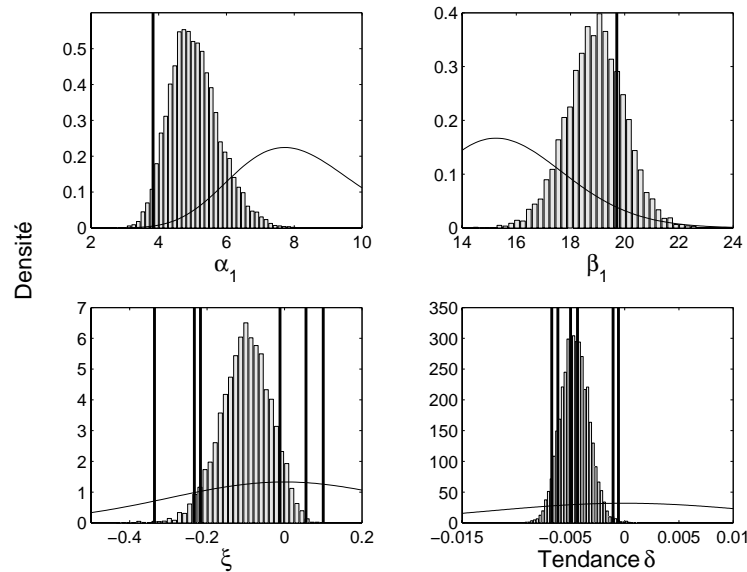


Figure 5-15. Estimation de quatre paramètres du modèle régional. Les histogrammes représentent la distribution *a posteriori*, estimée à partir des échantillons MCMC. La courbe en trait fin représente la distribution *a priori*, et les traits verticaux les estimations locales par maximum de vraisemblance.

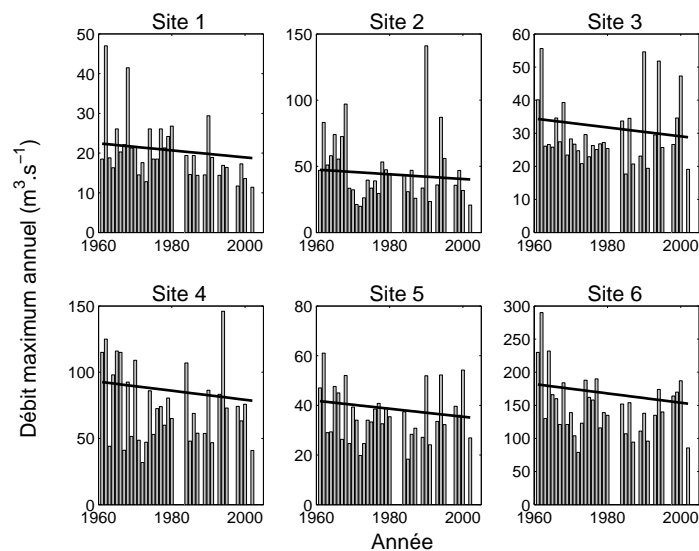


Figure 5-16. Débits maximum annuels pour les stations étudiées. Les droites représentent la tendance régionale sur la moyenne.

Afin de comparer de manière plus précise les estimations locales et régionales, nous avons estimé sur chacun des six sites les paramètres du modèle local suivant :

$$X_t^{(j)} \sim GEV(\alpha_j, \beta_j(1 + \delta_j(t - 1960)), \xi_j) \quad (5-30)$$

Les distributions *a priori* ont été spécifiées de manière analogue à celle utilisée pour le modèle régional, c'est à dire que pour un site j donné, la loi *a priori* est la suivante :

$$\begin{aligned} \pi(\theta_j) &= \pi(\alpha_j, \beta_j, \xi_j, \delta_j) \\ &= N(\xi_j; 0, 0.3^2) N(\delta_j; 0, (1/80)^2) Ga(\alpha_j; a_j^{(\alpha)}, b_j^{(\alpha)}) Ga(\beta_j; a_j^{(\beta)}, b_j^{(\beta)}) \end{aligned} \quad (5-31)$$

Afin d'estimer les lois *a posteriori* pour chaque site, les méthodes MCMC décrites précédemment ont été utilisées à l'identique. Les boîtes à moustaches présentées en Figure 5-17 permettent de comparer les distributions *a posteriori* locales à la distribution *a posteriori* régionale des paramètres de forme et de tendance. Les boîtes s'étendent du premier au troisième quartile, le trait horizontal représentant la médiane. Les moustaches s'étendent elles jusqu'aux quantiles à 5% et 95%. Concernant le paramètre de forme, la grande variabilité des estimations locales est ici visible. Ainsi, bien que les valeurs médianes semblent sensiblement différentes, tous les intervalles à 90% se recouvrent, ce qui signifie que cette différence pourrait n'être qu'un simple effet d'échantillonnage. L'autre information, commune aux deux paramètres, concerne la variance de la distribution *a posteriori* régionale : celle-ci est nettement réduite par rapport aux estimations locales, ce qui montre l'intérêt d'une analyse régionale. Cependant, il convient de rappeler qu'une hypothèse d'indépendance entre les stations a été faite, ce qui pourrait avoir pour effet de sous-estimer les variances régionales.

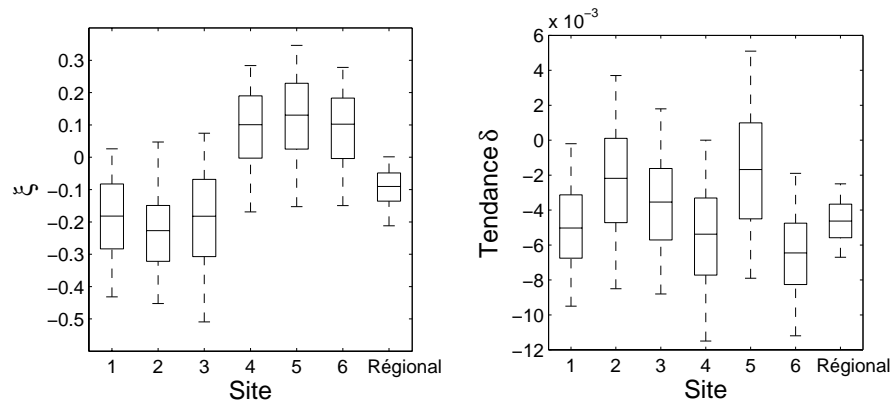


Figure 5-17. Comparaison des estimations locales et régionales pour les paramètres de forme et de tendance.

III.2.4. Validation du modèle

De nombreuses hypothèses ont été formulées afin de construire le modèle régional : il convient de vérifier qu'elles sont en adéquation avec les observations. Dans le cadre Bayésien, la validation du modèle se fait généralement en comparant la distribution empirique des observations avec la distribution prédictive *a posteriori*, qui est définie de la manière suivante :

$$p(\mathbf{y}^* | \mathbf{X}) = \int p(\mathbf{y}^* | \theta) p(\theta | \mathbf{X}) d\theta \quad (5-32)$$

Dans cette expression, \mathbf{y}^* est une valeur observable, dont on cherche à prédire la densité, après observation des données. La comparaison se base sur une statistique T : la valeur réalisée $T(\mathbf{X})$ (i.e. calculée sur les données) est comparée à la distribution prédictive de cette statistique, qui représente l'ensemble des valeurs qui auraient pu être observées d'après le

modèle. Si le modèle n'est pas cohérent avec les données, alors la statistique observée devrait être extrême pour la densité prédictive. Dans la pratique, cette dernière est estimée par simulation : un vecteur de paramètre est tout d'abord échantillonné suivant la loi *a posteriori* (l'échantillon issu des méthodes MCMC peut être utilisé à cet effet). Puis, un échantillon de données, de même effectif que l'échantillon original, est simulé suivant le modèle, avec le vecteur de paramètre précédemment sélectionné. En répétant cette opération un grand nombre de fois, on accède à un échantillon de statistiques $T(\mathbf{y}^{(r)})$ pour $r=1, \dots, N_{rep}$, qui estime la densité prédictive. Une *p*-valeur Bayésienne est alors estimée, en calculant la proportion de valeurs $T(\mathbf{y}^{(r)})$ qui excèdent la statistique observée. On considère généralement que des *p*-valeurs comprises entre 0.05 et 0.95 sont acceptables.

Le choix de la statistique de test est évidemment primordial : il doit permettre de mettre à l'épreuve les hypothèses les plus suspectes qui ont été faites pour construire le modèle. Dans le modèle régional considéré ici, les principales hypothèses sont les suivantes :

- ✓ Les données locales suivent une loi GEV : cette hypothèse est issue de la théorie des valeurs extrêmes, et est en général bien adaptée à la modélisation des maxima annuels. Nous évaluerons les statistiques moyennes, écart-type et asymétries empiriques comme indicateurs du bon ajustement de la loi GEV aux observations.
- ✓ Tous les paramètres de forme sont identiques : l'asymétrie sera un premier indicateur, puisque pour une loi GEV, elle ne dépend que du paramètre de forme. La statistique « valeur maximale » sera également utilisée, car elle est très dépendante de la valeur du paramètre de forme.
- ✓ Toutes les stations sont soumises à une tendance identique : le coefficient de Kendall entre le temps et les maxima annuels de chaque station sera utilisé.
- ✓ Hypothèse d'indépendance spatiale : le coefficient de Kendall entre les maxima annuels de deux sites sera utilisé.

Les résultats de cette étape de validation sont résumés dans le Tableau 5-6. Globalement, il apparaît que le modèle est cohérent avec les observations. L'hypothèse d'un coefficient de forme identique pour tous les sites semble notamment acceptable, si l'on se réfère aux indicateurs que sont l'asymétrie et la valeur maximale. Concernant l'hypothèse d'une tendance régionale, seule la première station semble présenter une tendance plus forte que celle obtenue par le modèle régional. Par contre, l'hypothèse d'indépendance entre les sites est clairement abusive. A titre d'exemple, trois coefficients de corrélation de Kendall avec le premier site présentent des *p*-valeurs inférieures à 0.05, ce qui signifie que les dépendances observées sont plus fortes que celles (nulles) supposées par le modèle. L'impact de cette hypothèse sera discuté plus en détail ultérieurement.

	<i>p</i> -valeur					
	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
Moyenne	0.547	0.202	0.741	0.911	0.749	0.922
Ecart-type	0.400	0.154	0.711	0.936	0.919	0.939
Asymétrie	0.156	0.185	0.346	0.801	0.851	0.542
Maximum	0.288	0.127	0.724	0.914	0.920	0.807
tau de Kendall avec le temps	0.976 *	0.786	0.398	0.622	0.372	0.670
tau de Kendall avec le site 1	/	0.003 *	0.068	0.042 *	0.036 *	0.119

Tableau 5-6. *p*-valeurs Bayésiennes pour différentes statistiques. Les astérisques dénotent les valeurs inférieures à 0.05 ou supérieures à 0.95.

III.2.5. Analyse fréquentielle

La loi *a posteriori* des paramètres ayant été estimée, il est possible de réaliser, comme dans le cas local, l'analyse fréquentielle des crues en prenant en compte l'incertitude liée au choix du modèle. Le modèle régional non stationnaire a donc été mis en concurrence avec le modèle régional stationnaire suivant :

$$X_t^{(j)} \sim GEV(\alpha_j, \beta_j, \xi) \quad (5-33)$$

La probabilité *a posteriori* de chacun de ces modèles a été calculée en utilisant la méthode de Chib. En supposant que les deux modèles sont équiprobables *a priori*, on obtient des probabilités *a posteriori* égales à environ 0.03 pour le modèle stationnaire et 0.97 pour le modèle non stationnaire. La prise en compte d'une évolution temporelle semble donc améliorer très nettement la description de ce jeu de données. Le mélange des modèles a été effectué comme dans le cas local. En guise d'illustration, l'évolution du quantile de probabilité 0.9 au cours du temps est illustrée pour chacune des six stations dans la Figure 5-18.

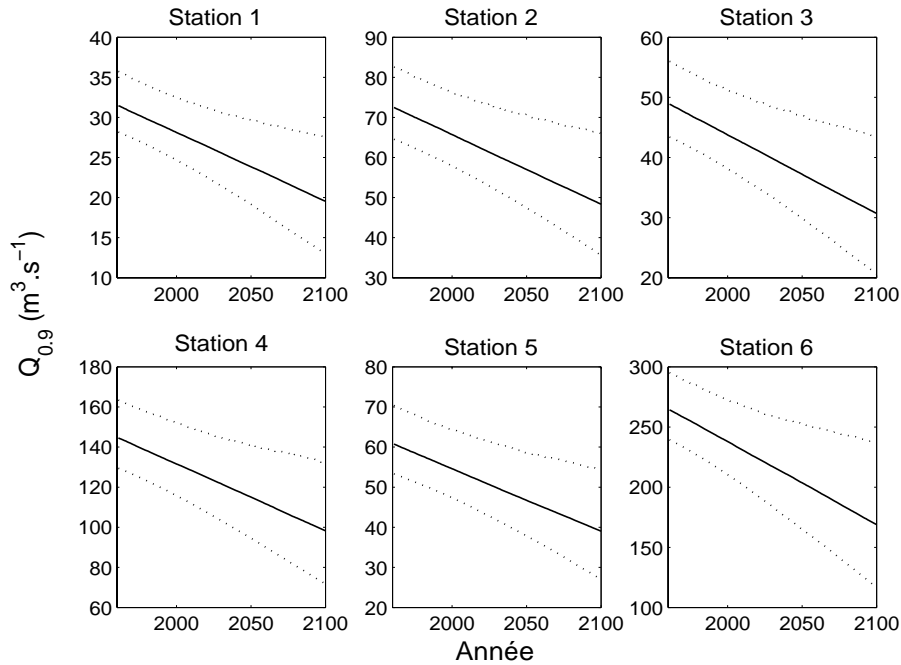


Figure 5-18. Evolution du quantile de probabilité 0.9 au cours du temps.

III.3. Le problème de la dépendance spatiale

Dans le cas d'étude précédent, le modèle utilisé semblait être adapté aux données, à l'exception de l'hypothèse d'indépendance entre les sites, qui s'est avérée nettement erronée. Dans cette section, nous proposons une première piste pour prendre en compte la dépendance spatiale, et évaluons l'impact que peut avoir la non-prise en compte de ce phénomène, dans l'optique de la description d'évolutions régionales notamment.

III.3.1. Quels sont les dangers d'ignorer la dépendance spatiale ?

Le problème de la dépendance spatiale dans le cadre de l'analyse régionale a été notamment étudié par Stedinger (1983), Hosking et Wallis (1988; 1997) et Madsen et Rosbjerg (1997). Ces études tentent de quantifier l'impact de la dépendance sur les estimateurs obtenus dans un cadre régional, mais ne proposent pas de méthode pour la prendre en compte formellement dans le modèle régional. Pour résumer, la conclusion partagée par ces études est que la dépendance augmente la variance des estimateurs, mais

n'induit pas de biais supplémentaire. De plus, la précision des estimateurs, mesurée par l'erreur quadratique moyenne, reste meilleure avec un modèle régional qui ignore la dépendance que dans le cas de l'analyse locale. Cette dernière conclusion reste même valable dans des régions modérément hétérogènes.

Toutefois, ces résultats doivent être considérés avec précaution dans le cadre des modèles régionaux présentés ici. Tout d'abord, les méthodes d'estimation que nous utilisons sont différentes de celles employées par ces auteurs (L-moments). De plus, des simulations Monte Carlo, équivalentes à des simulations *via* une copule Gaussienne, ont été utilisées pour ces études. Il serait intéressant de mettre à l'épreuve ces résultats dans d'autres contextes de simulation, impliquant notamment d'autres structures de dépendance. Enfin, la présence d'un paramètre de tendance est une différence majeure dans le contexte du travail présenté ici. L'impact de la dépendance spatiale pourrait donc être différent dans un cadre non stationnaire. De plus, être capable de quantifier fidèlement la variance d'un tel estimateur est capital : en effet, la significativité du changement dépend directement de cette variance. Dans le cadre Bayésien décrit à la section précédente, il est ainsi possible que la prise en compte des dépendances conduise à une décision différente quant au choix du modèle stationnaire ou non stationnaire. Nous étudions donc dans les parties suivantes comment la prise en compte explicite de la dépendance spatiale *via* une copule gaussienne modifie les résultats de l'analyse Bayésienne régionale.

III.3.2. La copule Gaussienne

Une description de la copule Gaussienne ainsi que de ses applications et limitations est fournie dans l'article figurant en annexe 3 (Renard et Lang, 2006), soumis à *Advances in Water Resources*. Nous nous contentons donc de décrire son utilisation dans le cadre de l'analyse régionale non stationnaire.

L'intérêt de la copule Gaussienne est qu'elle fournit un cadre paramétrique pour modéliser des données multivariées, avec des lois marginales arbitraires. Elle est paramétrée par une matrice Σ décrivant les dépendances deux à deux entre variables. Sa définition analytique est la suivante :

$$c(u_1, \dots, u_d) = \Phi_d \left(\phi^{-1}(u_1), \dots, \phi^{-1}(u_d) \right) \quad (5-34)$$

Dans cette expression, ϕ est la fonction de répartition de la loi normale centrée réduite $N(0, 1)$ et Φ_d est la fonction de répartition de la loi Gaussienne multivariée de dimension d , de moyenne θ et de matrice de variance Σ . En d'autres termes, la distribution multivariée dont la j^{eme} composante suit la loi F_j peut s'écrire :

$$\begin{aligned} F(x_1, \dots, x_d) &= c(F_1(x_1), \dots, F_d(x_d)) \\ &= \Phi_d \left(\phi^{-1}(F_1(x_1)), \dots, \phi^{-1}(F_d(x_d)) \right) \end{aligned} \quad (5-35)$$

La densité multivariée des observations s'écrit alors :

$$f(x_1, \dots, x_d) = f_1(x_1) \times \dots \times f_d(x_d) \times \left\{ \frac{|\Sigma|^{-1/2} \exp \left\{ -\frac{\left[\phi^{-1}(F_1(x_1)), \dots, \phi^{-1}(F_d(x_d)) \right] \left[\Sigma^{-1} - I \right] \left[\phi^{-1}(F_1(x_1)), \dots, \phi^{-1}(F_d(x_d)) \right]^T}{2} \right\}}{2} \right\} \quad (5-36)$$

Le principal avantage de la copule Gaussienne est sa simplicité : les dépendances sont en fait simplement décrites à partir de la loi Gaussienne multivariée sur les données anamorphosées. Ceci permet son utilisation en dimension supérieure à deux ou trois, sans difficulté théorique particulière. Cependant, il est important de noter que ce modèle n'est pas

issu de la théorie des valeurs extrêmes multivariées, et donc que sa validité n'est pas garantie. Il conviendra donc de vérifier que l'adéquation du modèle de dépendance aux données est acceptable. De plus, un autre inconvénient est lié aux propriétés asymptotiques de cet outil : la copule Gaussienne modélise en effet des données asymptotiquement indépendantes. Les propriétés concernant la dépendance asymptotique de valeurs extrêmes sont généralement assez difficiles à mettre en avant sur les observations, à cause principalement de la taille limitée des échantillons. En conséquence, la validité du modèle en extrapolation sera clairement douteuse, même si l'ajustement aux observations semble satisfaisant. Nous discuterons plus en détail des limitations de cette approche dans une section ultérieure.

L'estimation des paramètres peut également s'avérer problématique : en effet, le nombre de paramètres à estimer augmente très rapidement avec le nombre de sites à étudier, puisque la matrice de variance de la copule comprend $d(d-1)/2$ termes à estimer. A titre d'exemple, si l'on souhaitait ajuster un modèle avec copule Gaussienne aux six stations hydrométriques étudiées dans le section III.2. , le nombre de paramètres à estimer serait égal à 29. Deux solutions sont envisageables pour contourner cette difficulté. La première consiste, dans un premier temps, à estimer la matrice de variance de la copule, en utilisant un estimateur basé sur les rangs, donc invariant par anamorphose, comme celui proposé par Phoon *et al.* (2004). Dans un second temps, les paramètres des distributions marginales peuvent être estimés conditionnellement à l'estimateur précédent. Notre but étant de quantifier le plus fidèlement possible la variance des estimateurs, et notamment de la tendance régionale, cette méthode ne peut pas être considérée comme satisfaisante, puisque la procédure d'estimation conditionnelle ne permet pas de prendre en compte toutes les incertitudes d'estimation. La seconde solution considérée ici consiste à simplifier le modèle de dépendance, en supposant que la dépendance entre deux sites est une fonction décroissante de la distance qui les sépare. Ce type d'hypothèse est courant en géostatistiques. Elle peut s'avérer acceptable pour l'étude des précipitations extrêmes. Par contre, l'application aux débits est plus problématique, puisque la notion de distance entre deux stations hydrométriques n'est pas très informative : il faudrait plutôt considérer une distance entre bassins versants. De telles distances existent (cf. Sauquet (2000) pour quelques exemples), mais ne suffisent pas forcément à expliquer les dépendances observées sur les phénomènes extrêmes. En conséquence, nous allons effectuer un cas d'étude relatif à des données de précipitations.

III.3.3. Exemple sur les pluies

Les données utilisées sont identiques à celles présentées dans l'article figurant en Annexe 3, dans lequel nous étudions l'impact de la dépendance spatiale dans un cadre stationnaire. Nous allons généraliser cette étude en incluant une tendance régionale au modèle.

Données et modèles

Six stations pluviométriques situées en région Parisienne sont utilisées (Figure 5-19), pour lesquelles 58 valeurs maximales annuelles de pluies journalières ont été extraites entre les années 1922 et 2003. Les années 1926, 1936, 1939-1944, 1948, 1949, 1952, 1955-1958, 1981-1983, 1991 et 1997-2001 n'ont pas été utilisées car toutes les stations n'étaient pas exemptes de lacunes. En première analyse, nous n'avons pas trouvé de fortes différences entre les ordres de grandeur des précipitations sur ces stations. Ainsi, la médiane des maxima annuels varie entre 28.5 mm et 31.75 mm. De plus, d'un point de vue météorologique, la zone d'étude est assez homogène, et les gradients d'altitude sont faibles. En conséquence, nous avons supposé que les données des six stations pouvaient être considérées comme issues d'une même distribution.

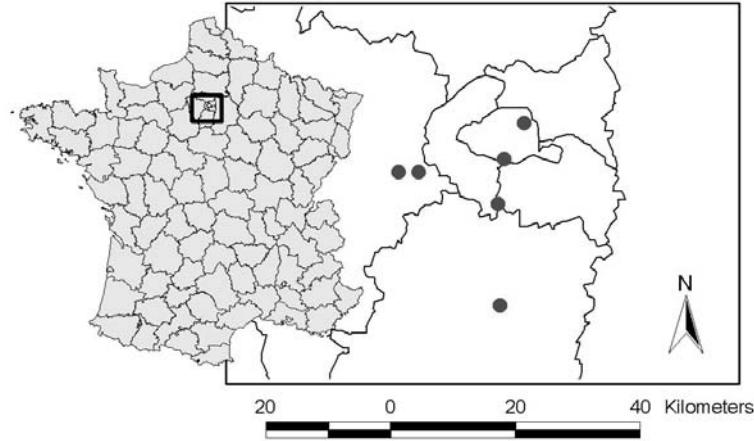


Figure 5-19. Localisations des stations pluviométriques étudiées.

Le modèle utilisé est similaire au modèle M2, à l'exception du fait que tous les paramètres sont supposés ici régionaux. Ainsi, en un site j et pour une année t , les maxima annuels $X_t^{(j)}$ sont modélisés de la façon suivante :

$$X_t^{(j)} \sim GEV(\alpha, \beta(1 + \delta(t-1921)), \xi) \quad (5-37)$$

En supposant qu'une copule Gaussienne est adéquate pour décrire les dépendances entre les observations, la densité d'un vecteur de maxima annuels en une année t peut s'écrire, d'après la formule (5-36) :

$$f(x_t^{(1)}, \dots, x_t^{(6)}; \alpha, \beta, \xi, \delta, \Sigma) = \left(\prod_{j=1}^6 GEV(x_t^{(j)}; \alpha, \beta(t), \xi) \right) |\Sigma|^{-1/2} \times \exp \left\{ - \frac{[V(x_t^{(1)}, \dots, x_t^{(6)}; \alpha, \beta(t), \xi)] [\Sigma^{-1} - I] [V(x_t^{(1)}, \dots, x_t^{(6)}; \alpha, \beta(t), \xi)]^T}{2} \right\}, \quad (5-38)$$

où $\beta(t) = \beta(1 + \delta(t-1921))$,

$$V(x_t^{(1)}, \dots, x_t^{(6)}; \alpha, \beta, \xi) = \left[\phi^{-1} \left(F(x_t^{(1)}; \alpha, \beta, \xi) \right), \dots, \phi^{-1} \left(F(x_t^{(6)}; \alpha, \beta, \xi) \right) \right]$$

et $F(x; \alpha, \beta, \xi)$ la fonction de répartition d'une loi GEV.

La vraisemblance des observations peut alors être calculée comme suit :

$$p(\mathbf{X} | \theta) = \prod_{t=1922}^{2003} f(x_t^{(1)}, \dots, x_t^{(6)}; \alpha, \beta, \xi, \delta, \Sigma) \quad (5-39)$$

Afin d'éviter de sur-paramétrer le modèle, la dépendance $\rho_{i,j}$ entre deux sites a été décrite comme une fonction décroissante de la distance $d_{i,j}$:

$$\begin{cases} \rho_{i,j} = \gamma_0 \exp(-\gamma_1 d_{i,j}) & \text{si } i \neq j \\ \rho_{i,i} = 1 \end{cases} \quad (5-40)$$

Un tel modèle est classique en géostatistiques. La discontinuité pour d proche de zéro est nommée effet pépite : elle traduit le fait que deux stations très proches ne sont pas parfaitement dépendantes, à cause notamment de l'incertitude liée à la mesure du phénomène.

Les distributions *a priori* ont été spécifiées de manière analogue à celle présentée dans la section III.2. Les mêmes distributions ont été utilisées pour les paramètres de forme et de tendance. Pour estimer les paramètres de position et d'échelle, les données de la période 1883-1921 sur la station Paris Montsouris, une des plus anciennes de France, ont été utilisées. La variance des estimateurs a été évaluée par Bootstrap, et une loi Gamma a été ajustée sur les estimateurs issus des échantillons répliqués. La distribution *a priori* des paramètres γ_0 et γ_1 , utilisés pour décrire la dépendance, a également été spécifiée par une approche Bootstrap utilisant les 23 années non communes à toutes les stations. A chaque réplcation, un jeu de données de 23 années est simulé par Bootstrap. Les corrélations des rangs sont alors calculées entre les stations deux à deux. La relation (5-40) est alors ajustée sur ces corrélations, par la méthode des moindres carrés. Un échantillon de N_{rep} valeurs de ces paramètres est ainsi créé. Une loi gamma est finalement ajustée pour le paramètre γ_1 . Pour le paramètre γ_0 , nous avons privilégié une loi Beta, car son support est égal à l'intervalle $[0 ; 1]$, ce qui est cohérent avec le domaine de définition du paramètre. En supposant l'indépendance *a priori* de tous les paramètres, la distribution *a priori* est finalement la suivante :

$$\begin{aligned} \pi(\theta) &= \pi(\alpha, \beta, \xi, \delta, \gamma_0, \gamma_1) \\ &= Ga(\alpha; a^{(\alpha)}, b^{(\alpha)}) Ga(\beta; a^{(\beta)}, b^{(\beta)}) N(\xi; 0, 0.3^2) \times \\ &N(\delta; 0, (1/80)^2) Beta(\gamma_0; a^{(\gamma_0)}, b^{(\gamma_0)}) Ga(\gamma_1; a^{(\gamma_1)}, b^{(\gamma_1)}) \end{aligned} \quad (5-41)$$

Les valeurs numériques des hyperparamètres sont précisées dans le Tableau 5-7.

Paramètre	α	β	γ_0	γ_1
Distribution	Gamma	Gamma	Beta	Gamma
a	42.9	308.8	7.69	0.58
b	5.5	12.3	6.49	75.05

Tableau 5-7. Spécification des distributions *a priori* marginales.

La distribution *a posteriori* est finalement calculée, à une constante près, comme le produit de la vraisemblance et de l'*a priori*. Afin de décrire cette distribution *a posteriori*, une stratégie MCMC en tout point identique à celle présentée à la section III.2.2. a été utilisée.

Validation du modèle

Par rapport au modèle régional présenté en section III.2. , une des principales différences est la prise en compte explicite des dépendances spatiales. Il convient donc de vérifier que la modélisation de ces dépendances que nous avons adoptée est acceptable. La Figure 5-20 représente les couples de valeurs transformées par normal score. La visualisation des nuages de points permet une première appréciation quant au modèle de dépendance utilisé. Par exemple, une copule Gaussienne ne sera pas adaptée à la description de données conduisant à des nuages de points de forme fortement non elliptique. Dans le cas présent, il semble difficile de rejeter, sur la base de cette visualisation, le modèle Gaussien de dépendance (ce qui ne constitue évidemment pas une preuve de sa validité). On remarque de plus la présence de fortes dépendances, entre les sites 4 et 5 notamment, qui justifient le rejet de l'hypothèse d'indépendance spatiale.

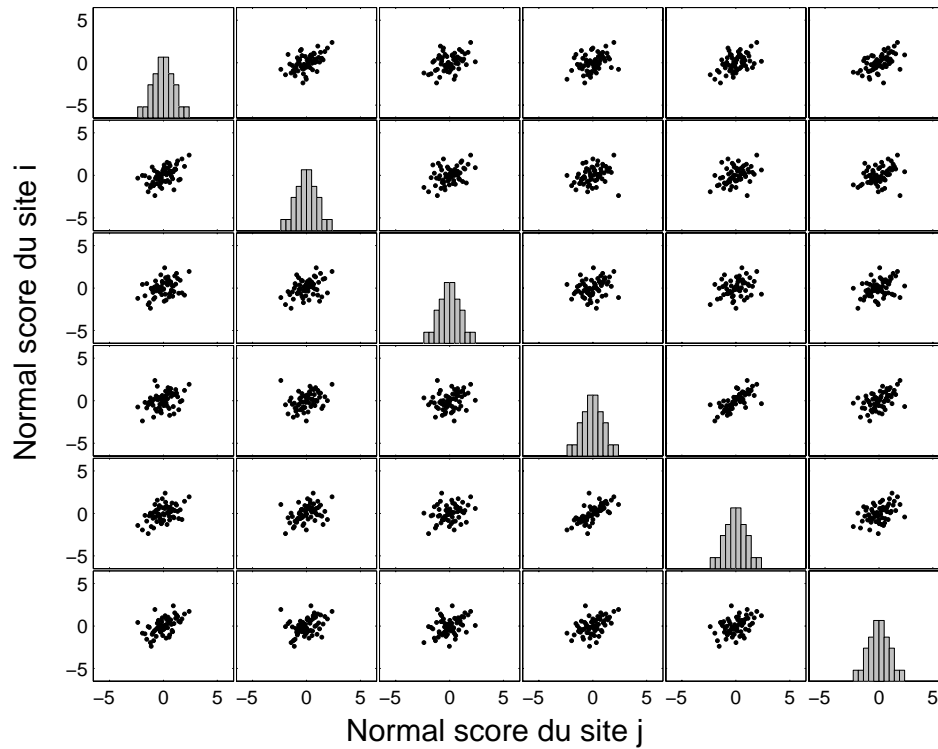


Figure 5-20. Matrice des nuages de points entre les maxima annuels de deux stations, transformés par normal score.

Nous avons également utilisé le graphique quantile-quantile basé sur une forme quadratique D devant suivre une loi du χ^2 (cf. Chapitre 4). Le résultat de cette analyse est présenté dans la Figure 5-21b. La matrice de variance a été estimée empiriquement à partir des données transformées par normal score. Le résultat semble globalement satisfaisant, malgré une légère déviation pour les faibles valeurs de D , ainsi que pour les deux quantiles les plus forts. Rappelons cependant que l'estimation de quantiles empiriques est très incertaine pour les plus fortes valeurs d'un échantillon. A titre de comparaison, le même graphique est réalisé (Figure 5-21a) en utilisant un modèle d'indépendance ($\Sigma = \mathbf{1}_d$). Les déviations par rapport à la première bissectrice sont alors plus systématiques, pour les faibles comme pour les fortes valeurs de D .

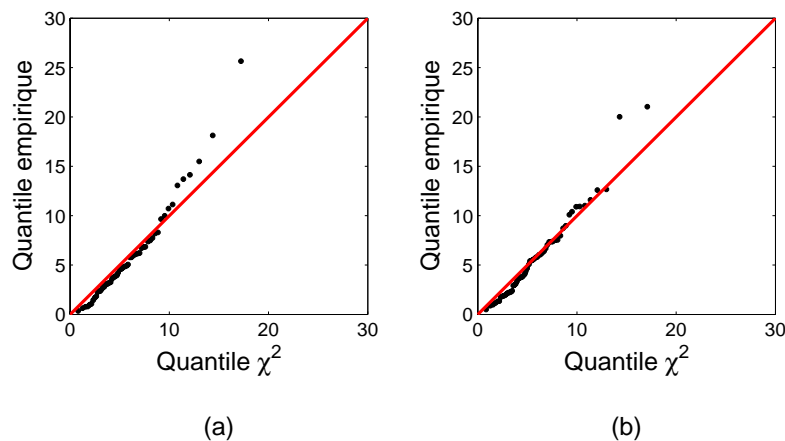


Figure 5-21. Graphiques quantile-quantile relatifs à la forme quadratique D .

(a) hypothèse d'indépendance ($\Sigma = \mathbf{1}_d$), (b) sans hypothèse d'indépendance ($\Sigma = \hat{\Sigma}$).

En admettant que le modèle Gaussien de dépendance soit acceptable pour la description des données, il est encore nécessaire de vérifier la pertinence de la relation dépendance-distance décrite par l'équation (5-40). La Figure 5-22 représente la courbe décrivant cette relation. Le trait plein représente la courbe médiane obtenue à partir de la loi *a posteriori* des paramètres γ_0 et γ_1 , et les traits pointillés représentent un intervalle de crédibilité à 90%. La relation semble cohérente avec des estimations des dépendances deux à deux. L'effet pépité conduit peut-être à sous-estimer la dépendance pour les faibles distances.

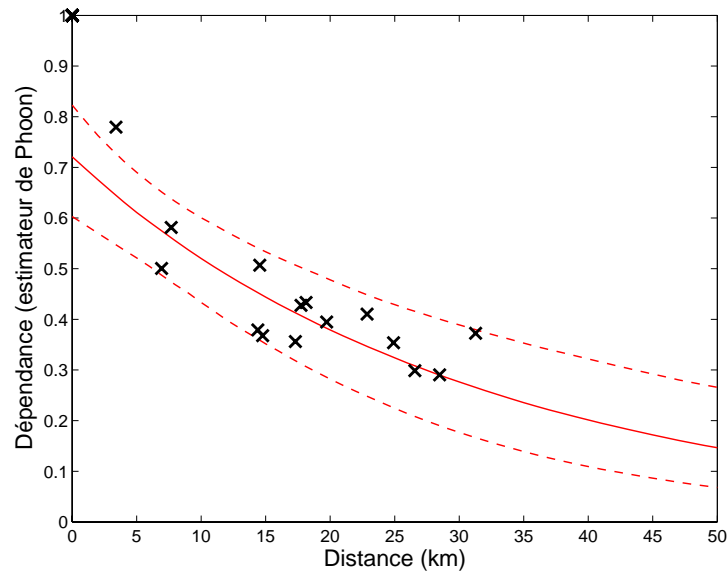


Figure 5-22. Relation dépendance-distance estimée à partir de la loi *a posteriori*. Les croix représentent des estimations des dépendances 2 à 2.

Pour compléter la validation du modèle, nous avons également calculé diverses *p*-valeurs Bayésiennes, comme expliqué en section III.2.4. En plus des statistiques utilisées dans cette précédente section, nous avons également évalué des statistiques liées aux caractéristiques de dépendance des observations. Plus précisément, nous avons dénombré le nombre d'années pour lesquelles toutes les stations ont connu un événement supérieur à une valeur x , et noté cet indicateur $NA(x)$. La valeur observée a ainsi été comparée à la distribution prédictive issue du modèle.

Les résultats de cette validation sont détaillés dans le Tableau 5-8. Il apparaît que le modèle offre une description satisfaisante des données. L'hypothèse d'une distribution identique sur tous les sites semble notamment acceptable, puisque aucun des quatre indicateurs décrivant les distributions locales (moyenne, écart-type, asymétrie, maximum) ne présente de *p*-valeur hors de l'intervalle [0.5 ; 0.95] (deux valeurs sont toutefois à la limite de cet intervalle, concernant la moyenne du site 2 et l'asymétrie du site 5). Les tendances observées sont également totalement compatibles avec celle issue du modèle régional (« tau de Kendall avec le temps »). Concernant les dépendances inter-sites (« tau de Kendall avec le site j »), le modèle parvient au moins à décrire l'intensité de ces dépendances deux à deux. Seule une *p*-valeur égale à 0.06 (dépendance entre les sites 4 et 5) est douteuse : elle correspond à la possible sous-estimation pour des sites proches évoquée précédemment. En dehors du cadre Gaussien, les dépendances deux à deux ne suffisent cependant pas à décrire les liaisons entre les variables marginales. Le modèle est donc également mis à l'épreuve *via* la statistique $NA(x)$, qui dépend de la structure globale de dépendance des observations. Pour des valeurs de x inférieures ou égales à 20 mm, le modèle semble en adéquation avec les données. On observe cependant une *p*-valeur égale à 0.01 pour $x=30$ mm. Plus précisément, on observe sur le jeu de données 14 années au cours desquelles toutes les stations ont connu

un événement supérieur à 30 mm. Si l'on observe la distribution prédictive de $NA(30\text{ mm})$ (Figure 5-23), on s'aperçoit que cette valeur de 14 années est considérée comme extrême par le modèle. On peut supposer que cette inadéquation soit liée à une mauvaise prise en compte par la copule Gaussienne des dépendances impliquant plus de deux variables, ou qu'il s'agisse d'un effet de l'hypothèse d'indépendance asymptotique : en effet, si les données sont asymptotiquement dépendantes, un modèle supposant l'indépendance asymptotique, comme le fait la copule Gaussienne, aura tendance à sous-estimer la probabilité d'occurrence d'événements extrêmes simultanément sur plusieurs composantes. Précisons que la p -valeur relative à $NA(40\text{ mm})$, qui semble acceptable, n'est que très peu informative, puisque seule une année durant laquelle tous les sites ont connu un événement supérieur à 40 mm a été observée.

	p -valeur					
	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
Moyenne	0.11	0.05	0.89	0.63	0.52	0.30
Ecart-type	0.13	0.17	0.71	0.67	0.67	0.50
Asymétrie	0.24	0.19	0.33	0.50	0.95	0.72
Maximum	0.15	0.13	0.48	0.57	0.86	0.73
tau de Kendall avec le temps	0.20	0.45	0.70	0.10	0.15	0.30
tau de Kendall avec le site 1	/	0.74	0.46	0.74	0.48	0.39
tau de Kendall avec le site 2	0.74	/	0.21	0.50	0.36	0.70
tau de Kendall avec le site 3	0.46	0.21	/	0.57	0.58	0.66
tau de Kendall avec le site 4	0.74	0.50	0.57	/	0.06	0.39
tau de Kendall avec le site 5	0.48	0.36	0.58	0.06	/	0.48
tau de Kendall avec le site 6	0.39	0.66	0.66	0.39	0.48	/
$NA(15\text{ mm})$			0.26			
$NA(20\text{ mm})$			0.16			
$NA(30\text{ mm})$			0.01*			
$NA(40\text{ mm})$			0.50			

Tableau 5-8. p -valeurs Bayésiennes pour différentes statistiques. Les astérisques dénotent les valeurs inférieures à 0.05 ou supérieures à 0.95.

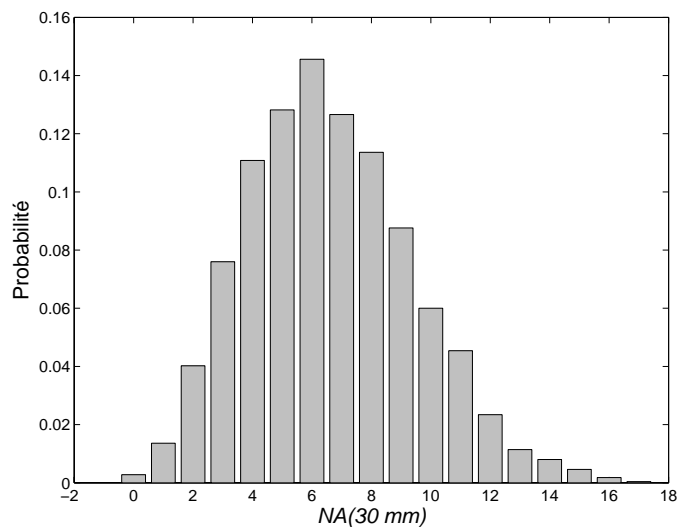


Figure 5-23. Densité prédictive de $NA(30\text{ mm})$.

Afin de préciser ces observations sur la faculté du modèle à reproduire des événements multivariés directement liés à la dépendance spatiale entre stations, nous avons comparé, pour diverses valeurs x , la probabilité empirique pour que le maximum annuel de toutes les stations soient supérieures à x avec la densité prédictive de cette probabilité. La probabilité pour qu'au moins une station ait un maximum annuel supérieur à x a également été étudiée. Les résultats sont présentés en Figure 5-24. Dans le premier cas (a), on observe que le modèle a tendance à légèrement sous-estimer cette probabilité. Pour la plupart des valeurs x , la densité prédictive n'est cependant pas incompatible avec la probabilité observée, la valeur $x=30$ mm faisant partie des exceptions. Dans le second cas (b), on observe une excellente cohérence entre modèle et observations. Globalement, on peut donc dire que la modélisation adoptée pour décrire les dépendances est pertinente dans le domaine des observations.

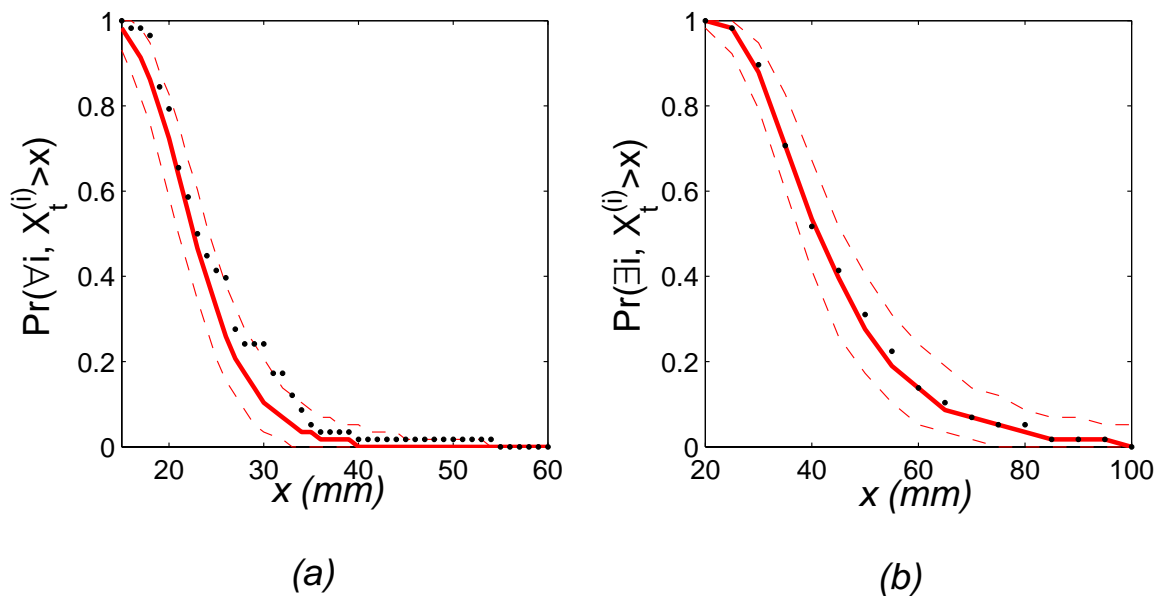


Figure 5-24. Probabilité pour que tous (a) ou au moins un (b) site(s) soi(en)t supérieur(s) à x . Les points représentent les probabilités empiriques observées, les traits la médiane (trait plein) et les quantiles à 5% et 95% (pointillés) de la densité prédictive.

Impact de la dépendance sur les distributions *a posteriori*

Nous allons à présent comparer les distributions *a posteriori* issues du modèle étudié, avec celles issues d'un modèle identique, mais où les stations ont été supposées indépendantes. Ces distributions, estimées par un noyau Gaussien à partir des échantillons simulés par les méthodes MCMC, sont illustrées dans la Figure 5-25. Les paramètres de position β et d'échelle α ont à peu près le même comportement : la prise en compte de la dépendance ne change pas sensiblement la valeur moyenne, mais augmente la variance *a posteriori*. Cette observation coïncide avec les conclusions formulées par Stedinger (1983), Hosking et Wallis (1988; 1997) et Madsen et Rosbjerg (1997) quant à l'effet de la dépendance inter-sites sur l'analyse régionale. Le paramètre de forme ζ présente un comportement différent : la prise en compte de la dépendance n'influe pas sur la variance, mais plutôt sur la localisation du mode de la distribution. Enfin, le paramètre de tendance δ semble être le plus sensible à la prise en compte de la dépendance : le mode est en effet déplacé, et la variance augmentée. Il est intéressant de noter que la valeur zéro possède une densité *a posteriori* relativement faible sous l'hypothèse d'indépendance, mais qu'elle devient beaucoup plus probable si la

dépendance est prise compte. On peut donc s'attendre à ce que la décision concernant le choix de l'un ou de l'autre des modèles soit également influencée, de même que l'analyse fréquentielle.

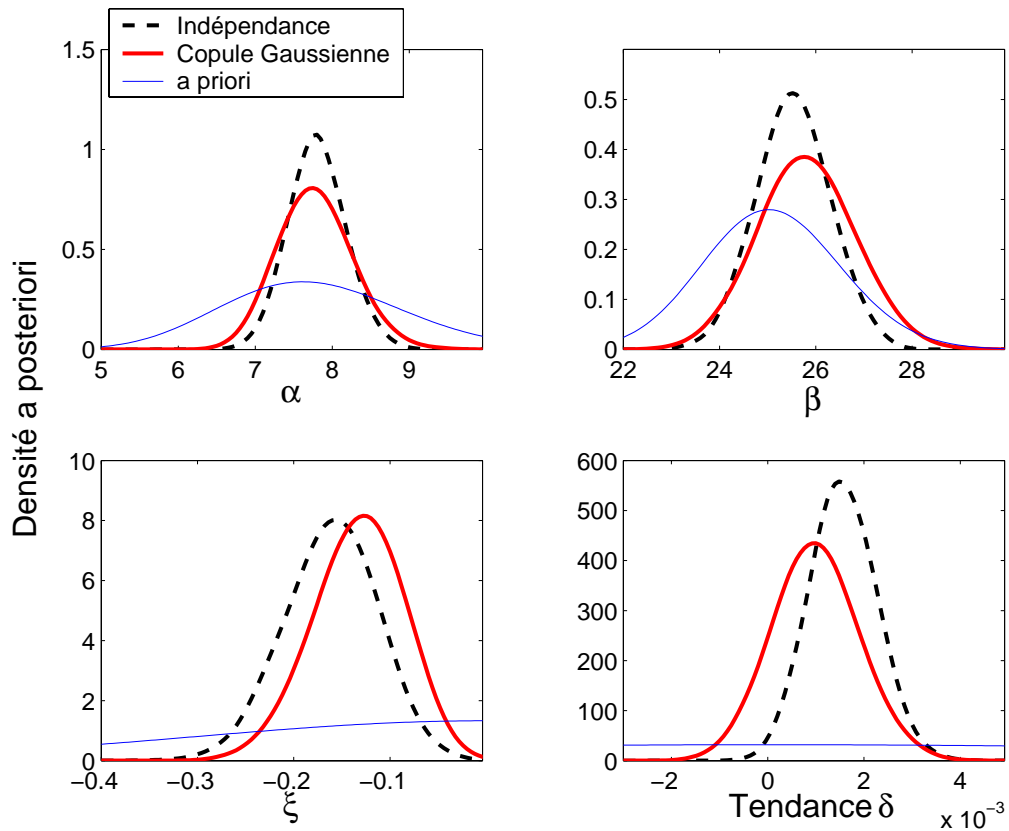


Figure 5-25. Comparaison des distributions a posteriori avec et sans l'hypothèse d'indépendance inter-sites.

Impact de la dépendance sur l'analyse fréquentielle

Afin de calculer les probabilités *a posteriori* des modèles à tendance régionale face à leurs homologues stationnaires, le modèle suivant a été utilisé, en faisant l'hypothèse d'indépendance ou en utilisant une copule Gaussienne :

$$X_t^{(i)} \sim GEV(\alpha, \beta, \xi) \quad (5-42)$$

En supposant que les modèles stationnaires et non stationnaires sont équiprobables *a priori*, les probabilités *a posteriori* du modèle à tendance régionale sont égales à 0.48 sous hypothèse d'indépendance et à 0.10 avec l'utilisation de la copule Gaussienne. Comme attendu, le fait de prendre en compte la dépendance spatiale tempère l'évidence quant à l'existence de cette non-stationnarité.

Les implications sur l'analyse fréquentielle des pluies sont décrites dans la Figure 5-26, qui décrit l'évolution du quantile de probabilité 0.9 au cours du temps, et dans la Figure 5-27, qui décrit la courbe des quantiles à une date fixée. Avec une copule Gaussienne, la non-stationnarité des pluies n'apparaît quasiment pas, alors qu'elle semble être une éventualité plus acceptable sous hypothèse d'indépendance. En conséquence, ce dernier modèle aura tendance à donner des quantiles plus importants pour le futur que le modèle prenant en compte la dépendance. Remarquons que la largeur des intervalles de crédibilité est sensiblement identique dans les deux cas : en effet, lorsque la dépendance est prise en compte, l'incertitude liée à l'estimation des paramètres est globalement plus importante, mais celle liée au choix du modèle est moindre, puisque le modèle a tendance à une faible probabilité (0.10). Ces deux phénomènes se compensent donc, conduisant à une incertitude globale comparable. Cependant, la confiance que l'on peut avoir vis à vis du modèle d'indépendance est limitée, puisque cette hypothèse n'est clairement pas vérifiée par les observations. A l'inverse, le modèle utilisant une copule Gaussienne est certes imparfait, mais semble plus pertinent.

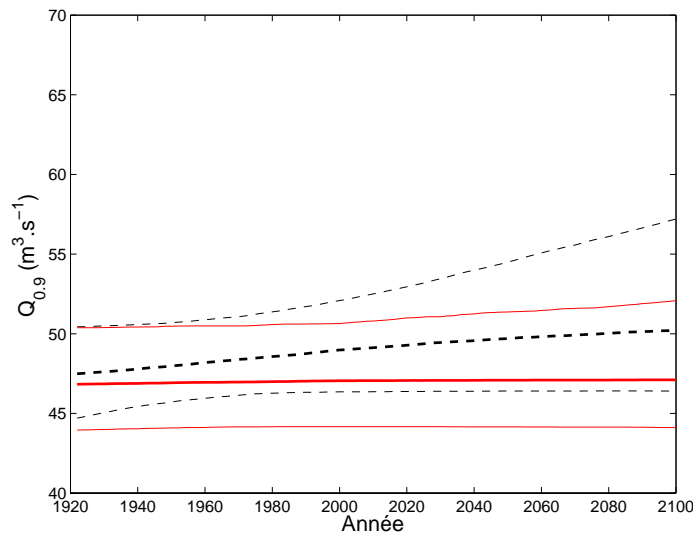


Figure 5-26. Evolution du quantile de probabilité 0.9 au cours du temps, sous hypothèse d'indépendance inter-sites (pointillés), et avec une copule Gaussienne (traits pleins).

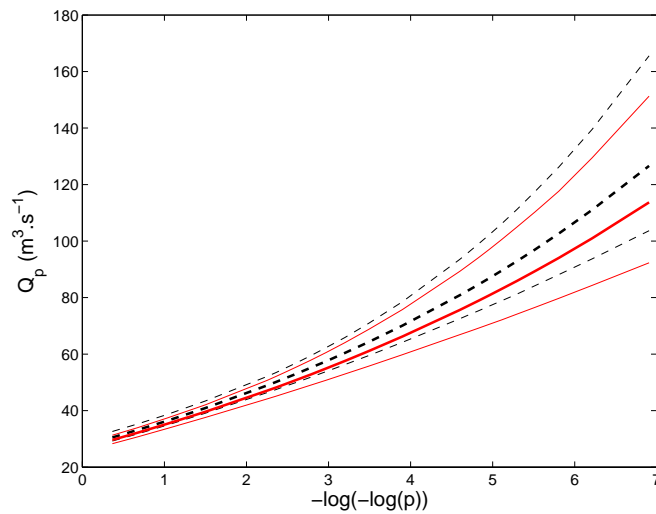


Figure 5-27. Courbe des quantiles en 2000, sous hypothèse d'indépendance inter-sites (pointillés), et avec une copule Gaussienne (traits pleins).

III.3.4. Discussion

Le cas d'étude précédent a permis d'illustrer le fait que la présence de dépendances entre les sites n'est pas anodine dans le cadre de l'analyse fréquentielle régionale, particulièrement si l'hypothèse de stationnarité doit être mise en doute. Cependant, les observations faites sur ce cas particulier ne peuvent pas être considérées comme des généralités : tout d'abord, d'autres jeux de données pourraient livrer des conclusions différentes ; de plus, nos résultats sont dépendants du modèle de copule Gaussienne adopté pour décrire la dépendance spatiale. Ce modèle n'est en effet qu'une description simplifiée de la dépendance, particulièrement dans le cadre des valeurs extrêmes. En effet, la dépendance entre deux sites n'est décrite qu'avec un unique paramètre, qui s'interprète comme une corrélation, alors que les relations liant deux variables extrêmes peuvent être bien plus complexes (la dépendance pouvant notamment dépendre des valeurs marginales). De plus, en dimension supérieure à 2, résumer toute la dépendance par des dépendances deux à deux est insuffisant hors du cadre Gaussien. Enfin, la copule Gaussienne possède une propriété d'indépendance asymptotique qui peut être problématique dans certains cas. Le choix de cette méthode est essentiellement lié à sa simplicité : elle fournit en effet une description paramétrique de phénomènes multivariés en dimension supérieure à 2 ou 3, alors que la généralisation vers des dimensions plus élevées est en général loin d'être triviale avec d'autres modèles (e.g. la généralisation multivariée du modèle logistique proposée par Tawn (1990)).

La question se pose donc de savoir si cet outil est adapté à la description de phénomènes spatiaux impliquant des valeurs extrêmes. Il convient de discerner deux objectifs bien différents : la description d'un jeu de données multivarié, et le calcul des probabilités de dépassement de phénomènes multivariés. Dans le premier cas, il convient de vérifier, comme dans toute modélisation, que le modèle est compatible avec les observations. Nous avons illustré quelques outils possibles dans le cas d'étude précédent. Le second problème est plus délicat, notamment en extrapolation, car la validité du modèle dépendra principalement des propriétés de dépendance asymptotique des observations. Or, il est en général difficile d'évaluer ces propriétés avec des données de type maxima marginaux, à cause principalement de la faible taille des échantillons.

En conclusion, nous pensons que la copule Gaussienne n'est généralement pas adaptée au calcul de faibles probabilités impliquant des phénomènes multivariés. A titre d'exemple, calculer la probabilité pour que tous les sites du cas d'étude précédent connaissent un événement supérieur à 100 mm avec le modèle de copule Gaussienne conduirait très certainement à des estimations peu convaincantes, voire fantaisistes. L'utilisation de maxima marginaux n'est de toute façon pas la meilleure approche dans ce type de problème, et des alternatives utilisant la totalité des données existent (De Haan et De Ronde, 1998). Par contre, nous pensons que si la description des dépendances fournie par une copule Gaussienne est compatible avec les données dans le domaine des observations, alors un tel modèle peut être bénéfique dans l'optique de l'estimation régionale. Un modèle prenant en compte même imparfaitement les dépendances conduira à une estimation plus réaliste, notamment en ce qui concerne la quantification des incertitudes, qu'un modèle d'indépendance, qui constitue dans tous les cas une description inappropriée à des données dépendantes.

SYNTHESE DU CHAPITRE 5

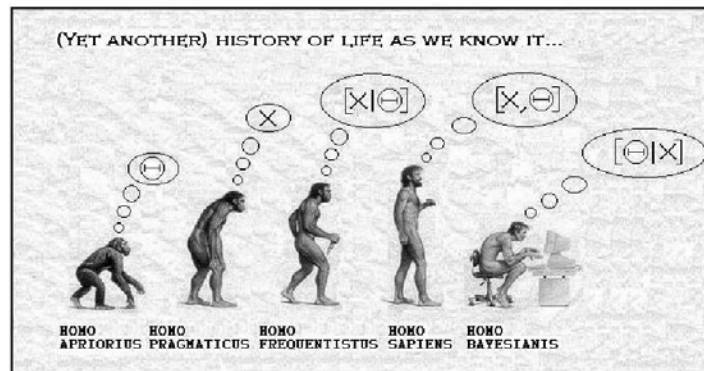
Ce chapitre a été consacré à la mise à jour des méthodes de prédétermination des risques pour prendre en compte d'éventuelles évolutions temporelles.

Dans un premier temps, quelques rappels concernant la définition du concept de période de retour ont été effectués. Nous avons mis en évidence des difficultés d'interprétation qui montrent qu'une échelle probabiliste semble préférable à une échelle temporelle pour quantifier le caractère exceptionnel d'événements hydro-climatiques extrêmes dans un contexte non stationnaire.

Au niveau méthodologique, les éventuelles évolutions temporelles ont été modélisées *via* des tendances sur les paramètres des distributions des variables d'intérêt. Nous avons choisi de nous placer dans le cadre de l'analyse Bayésienne pour l'estimation des paramètres. En effet, outre les avantages classiques de cette approche (introduction de connaissances *a priori*, accès direct à la distribution des paramètres, absence d'hypothèses asymptotiques, etc.), le formalisme Bayésien permet de prendre en compte au moins partiellement l'incertitude liée à la stationnarité des variables. Un cas d'étude détaillé illustre les différentes étapes de l'analyse au niveau local : introduction de connaissances *a priori* sur la distribution des quantiles, utilisation de la loi *a posteriori*, calcul des probabilités *a posteriori* des différents modèles et analyse fréquentielle.

La même approche a été appliquée à l'échelle régionale, en utilisant des modèles régionaux non stationnaires, pour lesquels certains paramètres (tendance et paramètre de forme) sont supposés identiques pour l'ensemble des stations appartenant à une région hydro-climatique homogène. La puissance des méthodes MCMC a permis d'estimer ce type de modèle complexe. Un premier cas d'étude a été traité en supposant l'indépendance spatiale des sites de la région. Nous avons ainsi pu illustrer les avantages potentiels de l'échelle régionale sur l'échelle locale, en terme de variabilité des estimateurs notamment. Dans un second temps, nous avons souhaité évaluer l'effet de l'hypothèse d'indépendance sur les résultats. A cet effet, les dépendances deux à deux entre sites ont été modélisées à l'aide d'une copule Gaussienne. Les résultats obtenus indiquent qu'ignorer la dépendance spatiale conduit à sous-estimer la variance des estimateurs, ce qui peut modifier assez nettement les probabilités *a posteriori* des différents modèles, et donc tempérer l'évidence quant à l'existence de non-stationnarités.

CHAPITRE 6. CONCLUSIONS ET PERSPECTIVES



Cette thèse avait pour objectif d'apporter des éléments de réponse à trois questions relatives à l'influence du changement climatique sur les valeurs extrêmes de débit : comment détecter des changements dans les séries hydro-météorologiques ? Peut-on discerner des impacts du réchauffement sur les observations hydrométriques ? Comment mettre à jour les méthodes de prédétermination pour intégrer l'éventuelle non-stationnarité des séries ?

Nous résumons dans cette section les principaux résultats obtenus au cours de ces travaux, tant d'un point de vue pratique que d'un point de vue méthodologique. Nous discutons également des perspectives qu'offrent ces résultats, et de manière plus générale, de l'intérêt d'utiliser des modèles statistiques évolués et adaptés pour améliorer la description et la compréhension des phénomènes hydro-climatiques extrêmes.

I. Acquis pratiques

I.1. Critique de longues séries de débits journaliers

Une attention particulière a été portée au cours de cette thèse à la critique des données. Il s'agit d'un travail long, parfois fastidieux, et coûteux, puisqu'il réclame de se déplacer chez les gestionnaires, qui sont géographiquement répartis sur tout le territoire français. Or, de nombreux travaux hydrologiques portent sur un grand nombre de stations hydrométriques. Il nous semble donc important de conserver une trace du travail réalisé, afin qu'il puisse profiter à d'autres utilisateurs. En effet, les stations appartenant aux échantillons réduits peuvent être considérées comme de bonne qualité, et exemptes de ruptures nettes expliquées par des biais météorologiques. Inversement, nous avons montré que de nombreuses stations présentaient des changements marqués liés à la météorologie, ce qui est susceptible d'introduire des biais dans les analyses relatives à ces bassins versants. A titre d'exemple, imaginons qu'un modèle pluie-débit soit calé sur une période donnée, puis appliqué à une autre période située après une rupture importante. Il est alors évident que les paramètres du modèle, qui permettaient de restituer correctement des débits avant la rupture, ne seront plus adaptés aux débits après rupture. De même, pour ces stations, les valeurs de débits de référence que l'on peut trouver dans la banque HYDRO sont erronées. Ces calculs sont en effet effectués automatiquement, en utilisant tous les débits publiés dans la banque.

I.2. Réflexions sur les techniques d'échantillonnage

L'extraction de variables décrivant les extrêmes hydrologiques à partir de chroniques journalières réclame d'effectuer un certain nombre de choix, particulièrement pour la sélection d'hydrogrammes ou de variables supérieures à un seuil. De plus, si les techniques d'échantillonnage sont en général efficaces sur des rivières au régime simple, des difficultés peuvent apparaître pour les régimes plus complexes, notamment si une composante nivale et une composante pluviale coexistent. Enfin, l'échantillonnage de nombreuses variables sur de nombreuses stations réclame des solutions informatiques robustes et entièrement automatisées.

En ce sens, les procédures utilisées dans cette thèse viennent compléter le travail déjà effectué au Cemagref par Lang (1995). Elles permettent notamment de détecter automatiquement les éventuels mélanges de populations relatifs aux cours d'eau dont le régime de crue est pluvio-nival. De plus, le champ de variables a été étendu, puisque des procédures spécifiques relatives aux ondes de fonte nivale et aux variables d'étiage ont été implémentées. D'un point de vue pratique, ces procédures se présentent comme un ensemble de routines programmées en FORTRAN. Elles pourraient être intégrées dans le logiciel PRETHY (Programme pour l'Echantillonnage et le Traitement de données Hydrométriques), développé au Cemagref par E. Sauquet, afin de permettre une utilisation conviviale des codes réalisés.

I.3. Impact du changement climatique sur les crues et les étiages

Le principal résultat pratique de cette thèse concerne le diagnostic effectué sur la stationnarité des débits extrêmes en France. La première conclusion est que la stationnarité des données hydrométriques est dans de nombreux cas une hypothèse irréaliste. Cependant, ces évolutions temporelles peuvent dans la plupart des cas être expliquées par des facteurs non climatiques, principalement des influences anthropiques directes et des problèmes météorologiques. Les variables décrivant le régime semblent être les moins sensibles à ces

changements, suivies par les variables de crue. Les variables d'étiage sont nettement les plus sensibles, particulièrement lorsqu'elles sont étudiées à pas de temps fin.

Sur les jeux réduits de stations pour lesquelles nous n'avons pas trouvé d'explication aux changements détectés, la présence d'évolutions temporelles apparaît bien moins évidente, bien que certains résultats significatifs localement demeurent. De plus, aucune cohérence n'apparaît à grande échelle. Ce résultat est en accord avec les observations effectuées sur les extrêmes de précipitations par les acteurs du projet IMFREX. Une analyse régionale a cependant permis de mettre en évidence des régions sur lesquelles plusieurs stations présentaient des tendances cohérentes, ce qui incite à penser que ces évolutions sont bien réelles. Nous avons alors tenté de relier ces évolutions à celles observées sur les précipitations et les températures. En utilisant des modèles hydrologiques, nous avons ainsi pu mettre en évidence un lien entre l'évolution des débits et l'évolution des variables climatiques sur deux régions (les Alpes et le Nord-Est de la France).

Ces résultats ne prouvent évidemment pas l'absence d'impacts du changement climatique, qui semble avéré en France pour les températures (Mestre, 2000; Spagnoli *et al.*, 2002), mais démontrent simplement la difficulté de discerner cette influence parmi toutes les autres sources de changement possibles. La variabilité naturelle des données extrêmes et la difficulté de les mesurer précisément compliquent encore ce type d'analyse. A une échelle plus réduite, les évolutions observées régionalement semblent bien réelles, mais les résultats contenus dans cette thèse ne suffisent pas à en déterminer la cause : variabilité climatique naturelle, impact du réchauffement climatique, ou modification du bassin versant.

II. Acquis méthodologiques

II.1. Détection de ruptures ou de tendances

La détection d'évolutions temporelles dans des données extrêmes nécessite à la fois l'obtention de données critiquées de bonne qualité et l'utilisation de procédures statistiques efficaces, qui sont à la fois non biaisées et puissantes. L'étude de ces méthodes a concerné à la fois l'échelle locale et l'échelle régionale.

Au niveau local, le problème n'est pas de mettre au point un test qui serait « meilleur » que les autres, mais de choisir, parmi les nombreuses méthodes disponibles dans la bibliographie, le test le plus adapté à la variable d'étude. Des études par simulation ont permis de clarifier les principaux facteurs influant sur le choix de la méthode, et de réaliser un schéma de sélection. *A posteriori*, les résultats de cette étude n'ont rien d'étonnant : les tests paramétriques sont plus puissants mais moins robustes, l'autocorrélation altère très sensiblement le risque de première espèce, et l'effectif de l'échantillon ne doit pas être trop faible lorsque la loi de la statistique de test est asymptotique. Certains points ont également été clarifiés : par exemple, pour les tests prenant en compte l'autocorrélation, la technique du *de-trending* nous paraît dangereuse, car elle a tendance à rendre les tests trop libéraux.

Au niveau régional, deux analyses complémentaires peuvent être réalisées : l'évaluation de la significativité régionale des résultats locaux, et la recherche de changements régionalement cohérents. Dans le premier cas, des simulations ont permis de montrer que la technique du Bootstrap par années proposée par Douglas *et al.* (2000) semblait efficace et relativement simple à appliquer. Dans le second cas, nous avons montré que les approches existantes permettaient d'évaluer une certaine significativité régionale, mais ne prenaient pas en compte la redondance d'information liée à la dépendance spatiale. Une procédure semi-paramétrique a donc été mise au point afin de différencier des changements cohérents et incohérents, compte tenu des dépendances observées entre les sites. La validité de cette approche est

cependant limitée par l'hypothèse de structure Gaussienne de la dépendance, qu'il est nécessaire de vérifier. Le développement de méthodes régionales plus robustes est certainement envisageable, en utilisant par exemple des permutations Bootstrap des années pour évaluer la loi de la statistique de test dans le cas où l'hypothèse de dépendance Gaussienne est problématique.

II.2. Solutions pour la gestion du risque en contexte non stationnaire

Là encore, les solutions proposées pour l'adaptation des méthodes de prédétermination au contexte non stationnaire concernent à la fois l'échelle locale et régionale. De manière générale, nous avons tout d'abord mis en avant les ambiguïtés liées à la définition d'une période de retour dans un contexte non stationnaire. L'interprétation temporelle classique est notamment problématique, puisque l'existence d'une période de retour n'est pas garantie. Bien que des définitions alternatives puissent être proposées, il semble préférable d'opter pour une échelle probabiliste afin de quantifier le caractère exceptionnel d'événements extrêmes dans un contexte non stationnaire.

L'approche que nous avons adoptée consiste à prendre en compte les évolutions temporelles en modélisant des tendances sur les paramètres des distributions des valeurs extrêmes. Cette approche est en partie empirique, car il n'est pas possible de prouver théoriquement que les extrêmes d'un processus non stationnaire suivront bien une loi des valeurs extrêmes. Cependant, les évolutions que nous cherchons à modéliser sont relativement lentes, si bien qu'à l'échelle de l'année, on peut faire l'approximation que les évolutions sont négligeables, ce qui justifie l'utilisation d'une loi des valeurs extrêmes. Cette solution pragmatique a d'ailleurs été proposée par plusieurs auteurs, par exemple Coles (2001) et Cunderlik et Burn (2003).

L'estimation des paramètres a été effectuée en utilisant le formalisme Bayésien. Au niveau local, nous avons adapté le travail de thèse réalisé par Luc Perrault (Perrault, 2000; Perrault *et al.*, 2000a, b) au contexte des événements extrêmes. En plus des avantages classiques de l'analyse Bayésienne (introduction d'informations *a priori*, absence d'hypothèses asymptotiques, etc.), cette approche permet de prendre en compte, au moins partiellement, l'incertitude liée à la stationnarité des séries étudiées.

Le passage à l'échelle régionale a été effectué en utilisant des modèles régionaux pour lesquels une partie des paramètres (et notamment le paramètre de tendance) est supposée identique pour tous les sites de la région étudiée. L'étude de tels modèles régionaux s'est alors déroulée en deux temps. Tout d'abord nous avons posé l'hypothèse d'indépendance spatiale des sites étudiés. Cette simplification permet de calculer la vraisemblance des observations, et d'effectuer une analyse Bayésienne comparable à celle effectuée à l'échelle locale. La principale différence concerne le nombre de paramètres à estimer, qui est bien plus important à l'échelle régionale. La puissance de méthodes MCMC permet de surmonter ce problème avec un nombre de sites raisonnable.

Dans un second temps, nous avons abordé le problème de la dépendance spatiale. Dans le cadre de travail que nous avons retenu, la principale difficulté est de pouvoir calculer la vraisemblance d'un vecteur d'observations : une loi multivariée, respectant les propriétés liées au caractère extrême des données locales, est en effet nécessaire. L'utilisation de copules est alors apparue naturelle dans ce contexte. La copule Gaussienne a été retenue car elle reste utilisable en dimension élevée (plus de deux ou trois dimensions) sans difficulté théorique particulière. La structure de dépendance imposée par ce modèle est par contre relativement rigide, et son adéquation avec les observations doit dans tous les cas être vérifiée. A partir d'un cas d'étude sur des précipitations, nous avons mis en avant l'impact de la dépendance spatiale sur les estimations. Il est notamment apparu, sur cet exemple particulier, qu'ignorer la

dépendance avait pour effet de diminuer artificiellement la variance des estimateurs, voire d'en décaler la valeur moyenne pour des paramètres de forme ou de tendance. En conséquence, l'évidence quant à l'existence d'évolutions temporelles est nettement différente, sur des données identiques, suivant que l'on prenne en compte ou pas la dépendance spatiale.

II.3. Mise en place d'outils statistiques dédiés à l'hydrologie

Afin de répondre aux objectifs fixés au début de la thèse, nous avons utilisé deux cadres d'analyse statistique, l'approche Bayésienne et les copules, qui sont couramment utilisés dans de nombreux domaines utilisant les statistiques appliquées, et commencent à intégrer la palette d'outil standard des hydrologues. Leur utilisation n'est cependant pas triviale, et a nécessité un important travail, tant théorique pour bien comprendre leur fonctionnement, que pratique, pour mettre en place les procédures informatiques afférentes. Nous espérons que les fruits de ce travail pourront profiter ultérieurement à d'autres utilisateurs.

L'analyse Bayésienne, associée aux méthodes MCMC, constitue une alternative intéressante à l'analyse fréquentielle classique. Elle permet en effet d'introduire des connaissances exogènes aux données *via* la distribution *a priori*, ce qui peut se révéler très intéressant en hydrologie, puisque le lien entre les écoulements et les processus physiques qui les génèrent ne sont pas totalement inconnus des hydrologues. Cette particularité pourrait permettre de répondre à une critique couramment formulée à l'encontre de l'approche statistique, par rapport à des approches plus physiques, à savoir la grande sensibilité des estimateurs à la présence de valeurs singulières. En intégrant plus de physique à la modélisation statistique *via* les *a priori*, on peut espérer diminuer cette sensibilité. Le fait de pouvoir accéder directement à une distribution des paramètres est également un avantage intéressant. En effet, les hypothèses asymptotiques (e.g. normalité asymptotique des estimateurs du maximum de vraisemblance) sont parfois difficilement admissibles. Enfin, le principal avantage de l'approche Bayésienne est sa grande flexibilité. Elle permet en effet de prendre en compte naturellement plusieurs sources d'incertitudes, notamment celle liée au choix d'un modèle. Signalons également la possibilité d'utiliser des modèles Bayésiens hiérarchiques, qui permettent de prendre en compte des variables latentes dans la modélisation. Le principal inconvénient, si l'on excepte les objections philosophiques qui ont longtemps déchiré Bayésiens et fréquentistes, est la difficulté calculatoire liée au caractère multidimensionnel de la distribution *a posteriori*. Grâce aux méthodes MCMC, cette difficulté peut aujourd'hui être surmontée, et devient même un avantage, puisque la puissance de ces méthodes permet de considérer des modèles de dimension relativement élevée, qui posent des problèmes numériques lorsqu'ils sont traités dans le cadre classique.

Les copules constituent le second outil statistique pris en main. Leur utilisation permet la prise en compte des dépendances existant dans les phénomènes multivariés. Bien que cet outil soit relativement simple à utiliser et puisse être utile si l'on s'intéresse simplement à la description d'un jeu de données multivarié, il ne peut être considéré que comme un premier pas vers une modélisation plus complète de la dépendance entre des valeurs extrêmes. La plupart des copules souffrent en effet d'un manque de connexion avec la théorie des valeurs extrêmes multivariées, ce qui rend leur utilisation en extrapolation problématique. Cependant, les aspects spatiaux, de par leur caractère fortement multidimensionnel, nécessiteront dans tous les cas d'effectuer un certain nombre d'hypothèses et d'approximations, dont la validité devra être évaluée sur les observations.

III. Perspectives

III.1. Organisation de la gestion des données hydrométriques en France

L'attention portée à la critique des données a mis en avant un certain nombre de difficultés, liées à l'accessibilité des informations relatives aux stations hydrométriques. La banque HYDRO permet une gestion des données hydrométriques à l'échelle nationale, ainsi qu'une mise à disposition fonctionnelle et efficace. Cependant, cette facilité d'accès aux valeurs de débits ne doit pas faire oublier que cette grandeur n'est pas mesurée au pas de temps publié, mais est en fait une hauteur transformée *via* la courbe de tarage. Les DIREN possèdent aujourd'hui une base de données, dénommée Barème, dans laquelle sont recensés les jaugeages effectués, ainsi que les différentes courbes de tarage utilisées depuis la dernière quinzaine d'années, parfois plus. De telles données sont très riches d'information, puisqu'elles permettent une première appréciation quant à la stabilité de la relation hauteur-débit, ainsi que sa validité en extrapolation. De plus, elles peuvent permettre de relier un changement de barème à une rupture observée sur les valeurs de débit publiées. Le développement et la diffusion de tels outils doivent donc être fortement encouragés, ainsi que la saisie des courbes plus anciennes. Les méta-données sont également une source d'information primordiale. Plus précisément, le recensement des faits marquants de la vie d'une station hydrométrique (déplacement, changement de gestionnaire, de matériel, mise en eau d'un aménagement, etc.) faciliterait grandement la validation des données. Evidemment, étant donné le nombre de stations gérées par les DIREN (ou les autres gestionnaires), ce type de recensement nécessiterait un travail colossal. Cependant, il est ponctuellement effectué lors d'une étude portant sur un cours d'eau particulier. Il serait bénéfique que les informations recueillies lors ces études ponctuelles ne soient pas perdues, et viennent alimenter une base de données.

L'approche adoptée par Météo France quant à l'homogénéité de ses longues séries de pluies ou de températures pourrait également être transposée aux débits. Les procédures d'homogénéisation développées par Mestre (2000) pourraient par exemple être adaptées pour un pas de temps relativement grossier (mensuel par exemple). Cependant, certaines caractéristiques propres aux débits rendent cette adaptation difficile. Tout d'abord, la variabilité spatiale est avant tout liée au réseau hydrographique, ce qui peut rendre problématique l'utilisation de stations de référence. De plus, les hétérogénéités s'expriment rarement de manière uniforme sur toute la gamme de débits. Par exemple, un changement d'extrapolation en hautes eaux peut créer des différences significatives sur les crues, mais ne pas affecter les débits plus modérés. Une alternative, récemment explorée par Naulet et Lang (2004), consiste à rechercher des changements dans les corrélations entre deux stations situées sur le même linéaire ou dans la même région. D'autres méthodes peuvent également voir le jour, puisque ce problème d'homogénéisation de longues chroniques est toujours du domaine de la recherche, notamment pour parvenir à homogénéiser des séries à pas de temps fin (journalier). L'autre solution consiste à créer un échantillon de séries de références, qui répondraient à un certain nombre de critères de qualité, et pour lesquelles l'hypothèse d'homogénéité pourrait être sinon garantie du moins acceptable.

Pour conclure, si nous n'avons pas été capables, aujourd'hui, de détecter un impact du réchauffement climatique dans les extrêmes hydrologiques, cela n'implique pas que cet impact n'existe pas. Immanquablement, les mêmes questions que celles auxquelles nous avons tenté de répondre dans cette thèse se reposeront à l'identique dans quelques années. Il n'est d'ailleurs pas impossible qu'une ré-analyse future des mêmes données délivre des résultats différents. Il pourrait donc être intéressant de définir un cadre de surveillance

climatique, afin d'affiner les études d'impact du réchauffement au cours des années. La mise en place d'un tel cadre nécessiterait d'une part de définir un réseau de stations de mesures utilisables, mais également de s'intéresser à des méthodes de détection séquentielles (permettant une mise à jour régulière des résultats), et non rétrospectives, comme celles que nous avons abordées ici.

III.2. Vers une meilleure prise en compte de la variabilité spatio-temporelle des événements extrêmes

Les méthodes statistiques prises en main au cours de la thèse présentent de nombreuses perspectives d'applications en hydrologie, dans le cadre d'analyses intégrant la variabilité spatio-temporelle des phénomènes, ainsi que les incertitudes liées à leur modélisation. Leur application a toutefois montré qu'elles nécessitaient encore des développements pour représenter correctement les propriétés statistiques des variables étudiées.

Une première piste consisterait à travailler sur les trois hypothèses des modèles régionaux utilisés au Chapitre 5. Tout d'abord, la prise en compte de la dépendance *via* une expression qui est fonction de la distance entre les sites ne peut être utilisée pour les débits. Une distance entre bassins versants doit donc être définie. En première approximation, on peut penser que la dépendance entre des débits de crue relevés en deux sites est avant tout le reflet de la dépendance des fortes pluies qui affectent les bassins versants et qui génèrent les événements de crue. Cette distance pourrait ainsi être dérivée des dépendances observées sur les pluies de bassin. La seconde limitation actuelle du modèle est liée à l'hypothèse d'égalité stricte de certains paramètres sur toute la région. Une hypothèse moins contraignante consisterait à autoriser ces paramètres à varier lentement et de manière continue dans l'espace. Un tel modèle peut être implémenté à l'aide d'un modèle Bayésien hiérarchique (Cooley, 2005; Cooley *et al.*, 2005). Le principe est de supposer que les paramètres sont eux-mêmes la réalisation d'un champ spatial Gaussien, dont la corrélation est décrite comme une fonction de la distance. Ainsi, pour deux sites proches, les paramètres des distributions ponctuelles seront également voisins puisque fortement corrélés. Le modèle proposé par Cooley permet également d'introduire des covariables pour expliquer des variations plus brusques de ces paramètres dans l'espace. Typiquement, la présence d'un relief aura tendance à modifier la distribution des pluies. Ce type de modèle permet donc d'assouplir considérablement la condition de région homogène.

Enfin, la dernière limitation concerne la modélisation de la dépendance entre les sites. L'utilisation d'un modèle plus adéquat que celui issu d'une copule Gaussienne serait souhaitable. Plus précisément, une caractérisation de la dépendance spatiale cohérente avec la théorie des valeurs extrêmes multivariées peut être effectuée dans le cas de processus Max-stables (Schalther et Tawn, 2003). Pour un ensemble de sites donné, la dépendance est caractérisée par un ensemble d'indices, nommés les coefficients extrémaux, qui décrivent la liaison entre tous les sous-ensembles possibles de sites. Cette caractérisation est cependant assez complexe, puisque pour un ensemble de p sites, il faut calculer $2^p - 1$ coefficients extrémaux. Une première approche consiste donc à caractériser la dépendance entre paires de sites. Ceci conduit à un outil comparable au variogramme couramment utilisé en géostatistiques, nommé un Madogramme (Cooley, 2005; Naveau *et al.*, 2005). Il serait donc intéressant d'utiliser des distributions multivariées qui respectent ces caractérisations issues de la théorie. Rappelons toutefois que cette théorie n'est valable qu'asymptotiquement, et que les données réelles, et notamment les données hydro-climatologiques, peuvent présenter des comportements particuliers. L'hypothèse de Max-stabilité, ou encore le caractère asymptotiquement dépendant ou indépendant d'un jeu de données spatialisé sont notamment problématiques, et peuvent être difficiles à vérifier sur des observations.

Au delà de l'aspect spatial, l'analyse de valeurs extrêmes multivariées constitue une perspective prometteuse en hydrologie. Ce domaine est actuellement abondamment étudié par les statisticiens, et les résultats obtenus laissent entrevoir de nombreuses applications. Nous ne l'avons abordé que partiellement dans l'optique de l'analyse régionale, puisque nous ne nous intéressons pas au domaine de l'extrapolation. Pourtant, la plupart des événements hydrologiques extrêmes sont intrinsèquement multivariés (Adamson *et al.*, 1999), et leur gravité est souvent le résultat d'une combinaison de valeurs extrêmes simultanément sur plusieurs variables. Estimer la probabilité relative à ce type de dépassement, y compris au-delà du domaine observé, est donc particulièrement intéressant. Par exemple, la construction d'hydrogrammes de référence pourrait être envisagée en étudiant la liaison existant entre plusieurs descripteurs (débit de pointe, volume, durée, asymétrie, etc.). Des approches telles que celle proposée par De Haan et De Ronde (1998) ou par Heffernan et Tawn (2004) devraient pouvoir être adaptées à la description multivariée d'hydrogrammes. Parmi les nombreuses autres applications possibles, citons l'analyse fréquentielle à une confluence, la modélisation multi-durées des pluies ou des débits, ou encore l'étude de la persistance des valeurs extrêmes (clusters).

De nombreuses perspectives d'application émanent également de l'utilisation du formalisme Bayésien. L'incorporation de variables latentes dans un modèle Bayésien permet ainsi de prendre en compte de façon plus large les incertitudes présentes dans l'analyse, et notamment celles liées à la mesure des phénomènes étudiés. Par exemple, dans un modèle pluie-débit conceptuel, les paramètres sont calés en utilisant un critère d'adéquation entre les débits observés et les débits issus du modèle. Or, les pluies utilisées comme données d'entrée du modèle sont généralement issues de postes pluviométriques, qui ne fournissent qu'une description ponctuelle du champ pluvieux ayant réellement touché le bassin versant. En conséquence, il existe une incertitude importante sur le calcul de la pluie de bassin, pouvant biaiser le calage des paramètres. Kavetski *et al.* (2006) proposent de reconnaître explicitement cette source d'incertitude, en considérant que pour un épisode pluvieux donné, la pluie qui est utilisée comme donnée d'entrée est le produit de la pluie réellement reçue par le bassin, qui est inconnue, et d'une erreur multiplicative commune à l'épisode pluvieux, et liée notamment au caractère ponctuel du réseau pluviométrique.

Kuczera (1999) utilise également le formalisme Bayésien pour prendre en compte l'incertitude liée à la transformation hauteur/débit dans l'analyse fréquentielle des crues. Cette dernière approche ouvre une perspective particulièrement intéressante, qui concerne l'ajout d'information historique pour améliorer la prédétermination des événements extrêmes. En effet, la reconstitution des débits de crue est généralement réalisée par extrapolation de la courbe de tarage, avec une incertitude qui varie selon la méthode utilisée (ajustement graphique ou modélisation hydraulique) et l'écart entre le niveau maximum jaugeé et celui de la crue considérée (Lang *et al.*, 2006). La prise en compte explicite de cette erreur potentielle de reconstitution permettrait de juger de manière plus réaliste de l'intérêt d'ajouter des informations historiques, qui sont moins précises que les données actuelles, mais permettent d'élargir considérablement le cadre temporel d'analyse. Le formalisme Bayésien et les méthodes MCMC constituent encore une fois des outils adaptés, sinon indispensables, à la réalisation d'une telle analyse (Reis *et al.*, 2005).

BIBLIOGRAPHIE

- Adamson, P. T., A. V. Metcalfe, et B. Parmentier. 1999. Bivariate extreme value distributions: an application of the Gibbs sampler to the analysis of floods. *Water Resources Research* **35**:2825-2832.
- Ambroise, B. 1999. *La Dynamique du Cycle de l'Eau dans un Bassin Versant – Processus, Facteurs, Modèles*. 200 p. HGA, Bucarest.
- Andréassian, V., C. Michel, C. Perrin, M. Tangara, et E. Parent. 2001. Détection de changements dans le comportement hydrologique des bassins versants : cadre méthodologique et exemples issus de l'hydrologie forestière. Rapport présenté pour l'évaluation du Thème de Recherche Transfeau Cemagref. 40 p.
- Andréassian, V., E. Parent, et C. Michel. 2003. A distribution-free test to detect gradual changes in watershed behavior. *Water Resources Research* **39**.
- Arnell, N. W. 1999. Climate change and global water resources. *Global Environmental Change-Human and Policy Dimensions* **9**:S31-S49.
- Arnell, N. W. 2003. Effects of IPCC SRES emissions scenarios on river runoff: a global perspective. *Hydrology and Earth System Sciences* **7**:619-641.
- Arora, V. K., et G. J. Boer. 2001. Effects of simulated climate change on the hydrology of major river basins. *Journal of Geophysical Research-Atmospheres* **106**:3335-3348.
- Bainville, V., et P. Ladoy. 1995. Préoccupations environnementales au début du XIXe siècle; la circulaire n° 18 du 25 avril 1821. *La Météorologie Volume spécial Histoire*:88-94.
- Barnett, T. P., K. Hasselmann, M. Chelliah, T. Delworth, G. Hegerl, P. Jones, E. Rasmusson, E. Roeckner, C. Ropelewski, B. Santer, et S. Tett. 1999. Detection and attribution of recent climate change: A status report. *Bulletin of the American Meteorological Society* **80**:2631-2659.
- Bayazit, M., et B. Onoz. 2004. Comment on "Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test" by Sheng Yue and Chun Yuan Wang. *Water Resources Research* **40**.
- Bayley, G. V., et J. M. Hammersley. 1946. The effective number of independent observations in an autocorrelated time series. *Journal of the Royal Statistical Society* **8**:184-197.
- Berger, J. O. 1985. *Statistical decision theory and bayesian analysis*. 617 p. New York.
- Bernier, J. 1994. Statistical detection of changes in geophysical series. *Engineering Risk in Natural Resources Management* **275**:159 - 176.
- Bernier, J., E. Parent, et J.-J. Boreux. 2000. *Statistique pour l'environnement: traitement bayésien de l'incertitude*. 364 p. Technique & Documentation.
- Birsan, M. V., P. Molnar, P. Burlando, et M. Pfaundler. 2005. Streamflow trends in Switzerland. *Journal of Hydrology* **314**:312-329.
- Black, A. R. 1996. Major Flooding and Increased Flood Frequency in Scotland since 1988. *Physical and Chemistry of the Earth* **20**:463-468.
- Black, A. R., et J. C. Burns. 2002. Re-assessing the flood risk in Scotland. *Science of the Total Environment* **294**:169-184.
- Bloomfield, P., et D. Nychka. 1992. Climate Spectra and Detecting Climate Change. *Climatic Change* **21**:275-287.
- Brillinger, D. R. 1989. Consistent detection of a monotonic trend superposed on a stationary time series. *Biometrika* **76**:23-30.

- Bronstert, A. 1996. River Flooding in Germany: Influenced by Climate Change? *Physical and Chemistry of the Earth* **20**:445-450.
- Buishand, T. A. 1982. Some methods for testing the homogeneity of rainfall records. *Journal of Hydrology* **58**:11 - 27.
- Burn, D. H., et M. A. H. Elnur. 2002. Detection of hydrologic trends and variability. *Journal of Hydrology* **255**:107-122.
- Caballero, Y., F. Habets, J. Noilhan, et J. F. Moor. 2003. Le changement climatique: étude de l'impact sur les ressources en eau du bassin Adour-Garonne. *Revue de l'agence de l'eau* **88**:5-11.
- Cassou, C. 2004. Du changement climatique aux régimes de temps: l'oscillation nord-atlantique. *La Météorologie* **45**:21-32.
- Champeaux, J. L., et A. Tamburini. 1995. Zonage climatique de la France à partir des séries de précipitations [1971-1990] du réseau climatologique d'État. *La Météorologie* **14**.
- Champion, M. 1858-1864. *Les inondations en France du VI^e siècle à nos jours*. 3000 p. Cemagref Editions.
- Chernoff, H., et S. Zacks. 1964. Estimating the current mean of a normal distribution which is subjected to change in time. *Annals of Mathematical Statistics* **35**:99 - 1018.
- Chetail, F. 2004. Caractérisation des régimes de crue des cours d'eau de la région Rhône-Alpes selon l'approche QdF. Rapport de stage. Cemagref. 85 p, Lyon.
- Chib, S. 1995. Marginal Likelihood From the Gibbs Output. *Journal of American Statistical Association* **90**:1313-1321.
- Chiew, F. H., et T. A. McMahon. 1993. Detection of trend or change in annual flow of Australian rivers. *International Journal of Climatology* **13**:643-653.
- Cigizoglu, H. K., M. Bayazit, et B. Onoz. 2005. Trends in the Maximum, Mean and Low Flows of Turkish Rivers. *Journal of Hydrometeorology* **6**:280-290.
- Coles, S. 2001. *An Introduction to Statistical Modeling of Extreme Values*. 210 p. Springer-Verlag, London.
- Coles, S., et L. Pericchi. 2003. Anticipating catastrophes through extreme value modelling. *Journal of the Royal Statistical Society Series C-Applied Statistics* **52**:405-416.
- Coles, S., L. R. Pericchi, et S. Sisson. 2003. A fully probabilistic approach to extreme rainfall modelling. *Journal of Hydrology* **273**:35-50.
- Coles, S. G. 1999. Statistical inference: Bayesian inference. <http://www.stats.bris.ac.uk/~masgc/teaching/bayes.ps>.
- Coles, S. G., et E. A. Powell. 1996. Bayesian methods in extreme value modelling: A review and new developments. *International Statistical Review* **64**:119-136.
- Coles, S. G., et J. A. Tawn. 1996. A Bayesian analysis of extreme rainfall data. *Applied Statistics-Journal of the Royal Statistical Society Series C* **45**:463-478.
- Cooley, D. 2005. Statistical Analysis of Extremes Motivated by Weather and Climate Studies: Applied and Theoretical Advances. University of Colorado. 122 p.
- Cooley, D., D. Nychka, et P. Naveau. 2005. A Spatial Bayesian Hierarchical Model for a Precipitation Return Levels Map. *in*. Extreme Value Analysis, Gothenburg, Sweden.
- Cox, D. R. 1966. *Théorie du renouvellement*. 186 p. Dunod, Paris.
- Cox, D. R., et P. A. W. Lewis. 1966. *The Statistical Analysis of Series of Events*. 285 p. London.
- CTGREF, Srae, et S. H. Diame. 1980-1982. Synthèse nationale sur les crues des petits bassins versants. Fascicule 2 : la méthode Socose ; Information Technique n° 38-2 (Juin 1980) ; Fascicule 3 : la méthode Crupedix.
- Cunderlik, J. M., et D. H. Burn. 2003. Non-stationary pooled frequency analysis. *Journal of Hydrology* **276**:210-223.

- Cunderlik, J. M., et T. Ouarda. 2006. Regional flood-duration-frequency modeling in the changing environment. *Journal of Hydrology* **318**:276-291.
- Dalrymple, T. 1960. Flood frequency analyses. in Water-supply paper 1553-A. US Geological Survey.
- Dansgaard, W., S. J. Johnsen, H. B. Clausen, D. Dahljensen, N. S. Gundestrup, C. U. Hammer, C. S. Hvidberg, J. P. Steffensen, A. E. Sveinbjornsdottir, J. Jouzel, et G. Bond. 1993. Evidence for General Instability of Past Climate from a 250-Kyr Ice-Core Record. *Nature* **364**:218-220.
- De Haan, L., et J. De Ronde. 1998. Sea and Wind: Multivariate Extremes at Work. *Extremes* **1**:7-45.
- Dindar, L. 2003. Analyse comparative de la puissance de tests de détection de ruptures dans des séries chronologiques de débit. Rapport de stage DESS. CEMAGREF. 84 p, Lyon.
- Douglas, E. M., R. M. Vogel, et C. N. Kroll. 2000. Trends in floods and low flows in the United States: impact of spatial correlation. *Journal of Hydrology* **240**:90-105.
- Ducharne, A., S. Théry, P. Viennot, E. Ledoux, E. Gomez, et M. Déqué. 2003. Influence du changement climatique sur l'hydrologie du bassin de la Seine. *VertigO - La revue en sciences de l'environnement* **4**:1-13.
- Easterling, D. R., J. L. Evans, P. Y. Groisman, T. R. Karl, K. E. Kunkel, et P. Ambenje. 2000. Observed Variability and Trends in Extreme Climate Events: A Brief Review. *Bulletin of the American Meteorological Society* **81**:417-426.
- El Adlouni, S., A. C. Favre, et B. Bobee. 2006. Comparison of methodologies to assess the convergence of Markov chain Monte Carlo methods. *Computational Statistics and Data Analysis*. In press.
- Embrechts, P., C. Klüppelberg, et T. Mikosch. 1997. *Modelling extremal events for insurance and finance*. 648 p. Springer.
- Etchevers, P., C. Golaz, F. Habets, et J. Noilhan. 2002. Impact of a climate change on the Rhone river catchment hydrology. *Journal of Geophysical Research-Atmospheres* **107**.
- Fisher, R. A., et L. H. Tippett. 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Cambridge Phil. Soc.* **24**.
- Frich, P., L. V. Alexander, P. Della-Marta, B. Gleason, M. Haylock, A. Tank, et T. Peterson. 2002. Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research* **19**:193-212.
- Galéa, G., et C. Prudhomme. 1997. Notions de base et concepts utiles pour la compréhension de la modélisation synthétique des régimes de crue des bassins versants au sens des modèles QdF. *Revue des Sciences de l'Eau* **1**:83-101.
- Garçon, R. 1999. Overall rain-flow model for flood forecasting and pre-determination. *Houille Blanche-Revue Internationale De l'Eau* **54**:88-95.
- Garçon, R., E. Bossard, M. Lang, et S. Le Clerc. 2002. Revisiter la notion de scénario hydrologique de référence pour la caractérisation du risque d'inondation. EDF-DTG, CEMAGREF. 76 p.
- Gardner, L. A. 1969. On detecting changes in the mean of normal variates. *The annals of mathematical statistics* **40**:116-126.
- Garreta, V. 2005. Analyse Bayésienne des débits de crue: méthodes MCMC et cas d'études. Rapport de stage de fin d'étude. Cemagref. 97 p.
- Gelfand, A. E., et A. F. M. Smith. 1990. Sampling-based approaches to calculating marginal densities. *Journal of American Statistical Association* **85**:398-409.
- Gelman, A., J. B. Carlin, H. S. Stern, et D. B. Rubin. 1995. *Bayesian data analysis*. 526 p. Chapman & Hall.

- Gelman, A., G. O. Roberts, et W. R. Gilks. 1996. Efficient Metropolis Jumping Rules. Pages 599-607 in *Bayesian Statistics 5*. Oxford University Press.
- Geman, S., et D. Geman. 1984. Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**:721-741.
- Gilard, O., et N. Gendreau. 1998. Inondabilité : une méthode de prévention raisonnable du risque d'inondation pour une gestion mieux intégrée des bassins versants. *Revue des Sciences de l'Eau* **11**:429-444.
- Gombay, E., et L. Horvath. 1994. An Application of the Maximum-Likelihood Test to the Change-Point Problem. *Stochastic Processes and Their Applications* **50**:161-171.
- Gombay, E., et L. Horvath. 1996a. Approximations for the time of change and the power function in change-point models. *Journal of Statistical Planning and Inference* **52**:43-66.
- Gombay, E., et L. Horvath. 1996b. On the rate of approximations for maximum likelihood tests in change-point models. *Journal of Multivariate Analysis* **56**:120-152.
- Gombay, E., et L. Horvath. 1997. An application of the likelihood method to change-point detection. *Environmetrics* **8**:459-467.
- GREHYS. 1996a. Inter-comparaison of regional flood frequency procedures for Canadian rivers. *Journal of Hydrology* **186**:85-103.
- GREHYS. 1996b. Presentation and review of some methods for regional flood frequency analysis. *Journal of Hydrology* **186**:63-84.
- Grelot, F. 2004. Gestion collective des inondations : peut-on tenir compte de l'avis de la population dans la phase d'évaluation économique a priori ? Mémoire de thèse. Ecole nationale supérieure d'arts et métiers Paris, Cemagref Lyon. 383 p.
- Groisman, P. Y., R. W. Knight, D. R. Easterling, T. R. Karl, G. C. Hegerl, et V. A. N. Razuvaev. 2005. Trends in intense precipitation in the climate record. *Journal of Climate* **18**:1326-1350.
- Groisman, P. Y., R. W. Knight, et T. R. Karl. 2001. Heavy precipitation and high streamflow in the contiguous United States: Trends in the twentieth century. *Bulletin of the American Meteorological Society* **82**:219-246.
- Gruza, G., E. Rankova, V. Razuvaev, et O. Bulygina. 1999. Indicators of climate change for the Russian Federation. *Climatic Change* **42**:219-242.
- Hamed, K. H., et A. R. Rao. 1998. A modified Mann-Kendall trend test for autocorrelated data. *Journal of Hydrology* **204**:182-196.
- Hansen, J., R. Ruedy, J. Glascoe, et M. Sato. 1999. GISS analysis of surface temperature change. *Journal of Geophysical Research-Atmospheres* **104**:30997-31022.
- Hasselmann, K. 1993. Optimal Fingerprints for the Detection of Time-Dependent Climate-Change. *Journal of Climate* **6**:1957-1971.
- Hasselmann, K. 1998. Conventional and Bayesian approach to climate-change detection and attribution. *Quarterly Journal of the Royal Meteorological Society* **124**:2541-2565.
- Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**:97-109.
- Hawkins, D. M. 1977. Testing a sequence of observations for a shift in location. *Journal of the American Statistical Association* **72**:180 - 186.
- Haylock, M., et N. Nicholls. 2000. Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910-1998. *International Journal of Climatology* **20**:1533-1541.
- Heffernan, J. E., et J. A. Tawn. 2004. A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society* **66**:497-546.

- Hennessey, K. J., R. Suppiah, et C. M. Page. 1999. Australian rainfall changes, 1910-1995. *Australian Meteorological Magazine* **48**:1-13.
- Hirsch, R. M., R. B. Alexander, et R. A. Smith. 1991. Selection of methods for the detection and estimation of trends in water quality. *Water Resources Research* **27**:803 - 813.
- Hisdal, H., K. Stahl, L. M. Tallaksen, et S. Demuth. 2001. Have streamflow droughts in Europe become more severe or frequent? *International Journal of Climatology* **21**:317-333.
- Hoeting, J. A., D. Madigan, A. E. Raftery, et C. T. Volinsky. 1999. Bayesian model averaging: A tutorial. *Statistical Science* **14**:382-401.
- Horvath, L. 1989. The limit distribution of the likelihood ratio and cumulative sum tests for a change in binomial probability. *Journal of Multivariate Analysis* **31**:148-159.
- Horvath, L. 1993. The maximum likelihood method for testing changes in the parameters of normal observations. *Annals of Statistics* **21**:671-680.
- Hosking, J. R. M., R. Wallis James, et F. Wood Eric. 1985. An appraisal of the regional flood frequency procedure in the UK flood studies report. *Hydrological Sciences Journal* **30**:85-109.
- Hosking, J. R. M., et J. R. Wallis. 1988. The effect of intersite dependence on regional flood frequency analysis. *Water Resources Research* **24**:588-600.
- Hosking, J. R. M., et J. R. Wallis. 1997. *Regional Frequency Analysis: an approach based on L-Moments*. 226 p. Cambridge University Press, Cambridge, UK.
- Hourcade, J. C. 2005. Scénario d'émissions pour le XXI^e siècle. Quels déterminants? Quelles incertitudes? in. *Changement climatique: le probable, le possible, l'improbable*, Toulouse, France.
- Hubbard, E. F., J. M. Landwehr, et A. R. Barker. 1997. Temporal variability in the hydrologic regimes of the United States. in. *FRIEND'97 - Regional Hydrology: Concepts and Models for Sustainable Water Resource Management*, Postojna, Slovenia. A. Gustard. 97-103.
- Hubert, P., J.-P. Carbonnel, et A. Chaouche. 1989. Segmentation des séries hydrométéorologiques - Application à des séries de précipitations et de débits de l'Afrique de l'Ouest. *Journal of Hydrology* **110**:349-367.
- IPCC. 2001. *Climate change 2001: synthesis report*. Cambridge University Press. 408 p, Cambridge.
- James, B., K. L. James, et D. Siegmund. 1992. Asymptotic Approximations for Likelihood Ratio Tests and Confidence-Regions for a Change-Point in the Mean of a Multivariate Normal-Distribution. *Statistica Sinica* **2**:69-90.
- Jaruskova, D. 1997. Some problems with application of change-point detection methods to environmental data. *Environmetrics* **8**:469-483.
- Javelle, P. 2001. *Caractérisation du régime des crues: le modèle débit-durée-fréquence convergent. Approche locale et régionale*. Mémoire de thèse. INP Grenoble, Cemagref Lyon. 268 p, Grenoble.
- Javelle, P., J. M. Grésillon, et G. Galéa. 1999. Discharge-duration-Frequency curves modeling for floods and scale invariance. *Comptes Rendus de l'Académie des Sciences, Sciences de la terre et des planètes*. **329**:39-44.
- Javelle, P., T. Ouarda, M. Lang, B. Bobee, G. Galea, et J. M. Gresillon. 2002. Development of regional flood-duration-frequency curves based on the index-flood method. *Journal of Hydrology* **258**:249-259.
- Jenkinson, A. F. 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society* **81**:158-171.

- Jones, P. D., T. J. Osborn, K. R. Briffa, C. K. Folland, E. B. Horton, L. V. Alexander, D. E. Parker, et N. A. Rayner. 2001. Adjusting for sampling density in grid box land and ocean surface temperature time series. *Journal of Geophysical Research-Atmospheres* **106**:3371-3380.
- Karl, T. R., et R. W. Knight. 1998. Secular trends of precipitation amount, frequency, and intensity in the United States. *Bulletin of the American Meteorological Society* **79**:231-241.
- Kass, R. E., et A. E. Raftery. 1995. Bayes Factors. *Journal of the American Statistical Association* **90**:773-795.
- Katz, R. W., M. B. Parlange, et P. Naveau. 2002. Statistics of extremes in hydrology. *Advances in Water Resources* **25**:1287-1304.
- Kavetski, D., G. Kuczera, et S. W. Franks. 2006. Bayesian analysis of input uncertainty in hydrological modeling: 2. Application. *Water Resources Research* **42**.
- Kehagias, A. 2004. A hidden Markov model segmentation procedure for hydrological and environmental time series. *Stochastic Environmental Research and Risk Assessment* **18**:117-130.
- Keim, B. D., et J. F. Cruise. 1998. A technique to measure trends in the frequency of discrete random events. *Journal of Climate* **11**:848-855.
- Kendall, M. G. 1975. *Rank correlation methods*. 202 p. Griffin, London.
- Kendall, M. G., et A. Stuart. 1943. *The advanced theory of statistics*. 451 p.
- Kharin, V. V., et F. W. Zwiers. 2000. Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM. *Journal of Climate* **13**:3760-3788.
- Kharin, V. V., et F. W. Zwiers. 2005a. Change in Temperature and precipitation extreme in the IPCC AR4 multi-model ensemble. *in*. Extreme Value Analysis, Gothenburg, Sweden.
- Kharin, V. V., et F. W. Zwiers. 2005b. Estimating extremes in transient climate change simulations. *Journal of Climate* **18**:1156-1173.
- Kharin, V. V., F. W. Zwiers, et X. B. Zhang. 2005. Intercomparison of near-surface temperature and precipitation extremes in AMIP-2 simulations, reanalyses, and observations. *Journal of Climate* **18**:5201-5223.
- Khodja, H., H. Lubès-Niel, J. M. Sabatier, E. Servat, et J. E. Paturel. 1998. Analyse spatio-temporelle de données pluviométriques en Afrique de l'Ouest. Recherche d'une rupture en moyenne. Une alternative intéressante: les tests de permutations. *Revue de Statistique Appliquée* **46**:95 -110.
- Kiely, G. 1999. Climate change in Ireland from precipitation and streamflow observations. *Advances in Water Resources* **23**:141-151.
- Kim, H. J., et L. J. Cai. 1993. Robustness of the Likelihood Ratio Test for a Change in Simple Linear-Regression. *Journal of the American Statistical Association* **88**:864-871.
- Krasovskaia, I., L. Gottschalk, et E. Leblois. 2002. Signature of changing climate in river flow regimes of Rhone-Mediterranean-Corsica region. *Houille Blanche-Revue Internationale De L Eau*:25-30.
- Kuczera, G. 1999. Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian inference. *Water Resources Research* **35**:1551-1557.
- Kumar, K. K., B. Rajagopalan, et M. A. Cane. 1999. On the weakening relationship between the Indian monsoon and ENSO. *Science* **284**:2156-2159.
- Kundzewicz, Z. W. 2004. Searching for change in hydrological data. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* **49**:3-6.

- Kundzewicz, Z. W., D. Graczyk, T. Maurer, I. Pinskiwar, M. Radziejewski, C. Svensson, et M. Szwed. 2005. Trend detection in river flow series: 1. Annual maximum flow. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* **50**:797-810.
- Kundzewicz, Z. W., et A. J. Robson. 2000. *Detecting trend and other changes in hydrological data*. 160 p. Wallingford, UK.
- Kundzewicz, Z. W., et A. J. Robson. 2004. Change detection in hydrological records - a review of the methodology. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* **49**:7-19.
- Landsea, C. W., R. A. Pielke, A. Mestas-Nunez, et J. A. Knaff. 1999. Atlantic basin hurricanes: Indices of climatic changes. *Climatic Change* **42**:89-129.
- Lang, M. 1995. Les chroniques en hydrologie: Modélisation comparée par un système de gestion de bases de données relationnel et orienté-objet. Traitements de base et intervalles de confiance des quantiles de crues. Techniques d'échantillonnage par la méthode du renouvellement. Mémoire de thèse. Université Joseph Fourier. 296 p, Grenoble.
- Lang, M. 2000. Analyse probabiliste des crues pour la prévention du risque d'inondation. Habilitation à diriger des recherches. Institut polytechnique de Grenoble. 153 p, Grenoble.
- Lang, M., T. B. M. J. Ouarda, et B. Bobée. 1999. Towards operational guidelines for over-threshold modeling. *Journal of Hydrology* **225**:103-117.
- Lang, M., C. Perret, E. Renouf, E. Sauquet, et A. Paquier. 2006. Incertitudes sur les débits de crue. *in*. Valeurs rares et extrêmes de précipitations et de débits, Lyon, France. SHF. 103-112.
- Lang, M., P. F. Rasmussen, G. Oberlin, et B. Bobee. 1997. Over-threshold sampling: modeling of occurrences by renewal processes. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* **3**:279-320.
- Lang, M., B. Renard, L. Dindar, F. Lemaitre, et P. Bois. 2004. Use of a statistical test based on Poisson process for the detection of changes in peak-over-threshold series. *in*. Hydrology : Science & Practice for the 21st Century, London. British Hydrological Society. 158-164.
- Le Clerc, S. 2004. Revisiter la notion de scénario hydrologique de référence pour la caractérisation des inondations. Université Joseph Fourier Grenoble, Cemagref Lyon. 374 p, Grenoble.
- Leadbetter, M. R. 1983. Extremes and local dependence in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **65**:291-306.
- Leblois, E. 2002. Evaluation of the possible impacts of climatic change by distributed models (Gewex-Rhone et Gicc-Rhone projects). *Houille Blanche-Revue Internationale De l'Eau*:78-83.
- Lee, A. F. S., et S. M. Heghinian. 1977. A shift of the mean level in a sequence of independant normal random variables - a Bayesian approach. *Technometrics* **19**:503-506.
- Lemaitre, F. 2002. Recensement des tests de détection de tendances ou de ruptures adaptés à l'analyse de stationnarité des régimes de crues en France. Rapport de fin d'étude ENTPE Cemagref. 94 p, Lyon.
- Lettenmaier, D. P., A. W. Wood, R. N. Palmer, E. F. Wood, et E. Z. Stakhiv. 1999. Water resources implications of global warming: A US regional perspective. *Climatic Change* **43**:537-579.
- Lettenmaier, D. P., E. F. Wood, et J. R. Wallis. 1994. Hydro-Climatological Trends in the Continental United-States, 1948-88. *Journal of Climate* **7**:586-607.

- Levitus, S., J. I. Antonov, T. P. Boyer, et C. Stephens. 2000. Warming of the world ocean. *Science* **287**:2225-2229.
- Lindstrom, G., et S. Bergstrom. 2004. Runoff trends in Sweden 1807-2002. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* **49**:69-83.
- Lins, H. F., et J. R. Slack. 1999. Streamflow trends in the United States. *Geophysical Research Letters* **26**:227-230.
- Livezey, R. E., et W. Y. Chen. 1983. Statistical field significance and its determination by Monte Carlo techniques. *Monthly Weather Review* **111**:46-59.
- Llasat, M. D., M. Barriendos, et T. Rigo. 2002. Frequency and temporal analysis of mediterranean heavy rainfalls producing severe floods in Spain and southern France. *Houille Blanche-Revue Internationale De l'Eau*:139-144.
- Lu, Z. Q., et L. M. Berliner. 1999. Markov switching time series models with application to a daily runoff series. *Water Resources Research* **35**:523-534.
- Lubès-Niel, H., J. M. Masson, J.-E. Paturel, et E. Servat. 1998. Variabilité climatique et statistiques. Etude par simulation de la puissance et de la robustesse de quelques tests utilisés pour vérifier l'homogénéité de chroniques. *Revue des Sciences de l'Eau* **3**:383 - 408.
- Ludwig, W., P. Serrat, L. Cesmat, et J. Garcia-Esteves. 2004. Evaluating the impact of the recent temperature increase on the hydrology of the Tet River (Southern France). *Journal of Hydrology* **289**:204-221.
- Madsen, H., et D. Rosbjerg. 1997. The partial duration series method in regional index-flood modeling. *Water Resources Research* **33**:737-746.
- Mann, H. B. 1945. Nonparametric tests against trend. *Econometrica* **13**:245-259.
- Mann, M. E., R. S. Bradley, et M. K. Hughes. 1998. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* **392**:779-787.
- Manton, M. J., P. M. Della-Marta, M. R. Haylock, K. J. Hennessy, N. Nicholls, L. E. Chambers, D. A. Collins, G. Daw, A. Finet, D. Gunawan, K. Inape, H. Isobe, T. S. Kestin, P. Lefale, C. H. Leyu, T. Lwin, L. Maitrepierre, N. Ouprasitwong, C. M. Page, J. Pahalad, N. Plummer, M. J. Salinger, R. Suppiah, V. L. Tran, B. Trewin, I. Tibig, et D. Yee. 2001. Trends in extreme daily rainfall and temperature in Southeast Asia and the South Pacific: 1961-1998. *International Journal of Climatology* **21**:269-284.
- Mardia, K. V. 1980. Tests of univariate and multivariate normality. Pages 279-320 in P.R. Krishnaiah, editor. *Handbook of Statistics 1: analysis of variance*, Amsterdam, Holland.
- Maronna, R., et V. J. Yohai. 1978. A bivariate test for the detection of a systematic change in mean. *Journal of American Statistical Association* **73**:640 - 645.
- Martins, E. S., et J. R. Stedinger. 2000. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research* **36**:737-744.
- Matalas, N. C., et W. B. Langbein. 1962. Information content on the mean. *Journal of Geophysical Research* **67**:3441-3448.
- McGilchrist, C. A., et K. D. Woodyer. 1975. Note on a distribution-free CUSUM technique. *Technometrics* **17**:321-325.
- Mears, C. A., et F. J. Wentz. 2005. The effect of diurnal correction on satellite-derived lower tropospheric temperature. *Science* **309**:1548-1551.
- Mestre, O. 2000. Méthodes statistiques pour l'homogénéisation de longues séries climatiques. Mémoire de thèse. Université Paul Sabatier. 229 p, Toulouse.
- Mestre, O. 2003. Modélisation des paramètres de la GEV: application à l'étude des crues. in. XVIe entretiens du centre Jacques Cartier., Lyon.

- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, et E. Teller. 1953. Equation of state calculations by fast computing machines. *Journal of chemical physics* **21**:1087-1092.
- Metropolis, N., et S. Ulam. 1949. The Monte Carlo method. *Journal of the American Statistical Association* **44**:335-341.
- Meunier, M. 2001. Regional flow-duration-frequency model for tropical island of Martinique. *Journal of Hydrology* **247**:31-53.
- Moberg, A., et P. D. Jones. 2005. Trends in indices for extremes in daily temperature and precipitation in central and western Europe, 1901-99. *International Journal of Climatology* **25**:1149-1171.
- Moisselin, J. M., M. Schneider, C. Canellas, et O. Mestre. 2002. Les changements climatiques en France au vingtième siècle. *La Météorologie* **38**:45-56.
- Molnar, P., et J. A. Ramirez. 2001. Recent trends in precipitation and streamflow in the Rio Puerco Basin. *Journal of Climate* **14**:2317-2328.
- Mosselman, E. 2006. Les valeurs rares et extrêmes dans la gestion des risques d'inondation aux Pays-Bas. *in*. Valeurs rares et extrêmes de précipitations et de débits, Lyon, France. SHF.
- Moyeed, R. A., et R. T. Clarke. 2005. The use of Bayesian methods for fitting rating curves, with case studies. *Advances in Water Resources* **28**:807-818.
- Mudelsee, M., M. Borngen, G. Tetzlaff, et U. Grunewald. 2003. No upward trends in the occurrence of extreme floods in central Europe. *Nature* **425**:166-169.
- Nash, J. E., et J. V. Sutcliffe. 1970. River flow forecasting through conceptual models: 1. A discussion of principles. *Journal of Hydrology* **10**:257-274.
- Naulet, R., et M. Lang. 2004. Critique d'un jeu de données hydrométriques pour la Diren Ile-de-France. Cemagref. 37 p, Lyon, France.
- Naveau, P., D. Cooley, et P. Poncet. 2005. Spatial extremes analysis in climate studies. *in*. Extreme Value Analysis, Gothenburg, Sweden.
- Nogaj, M., S. Yiou, S. Parey, F. Malek, et P. Naveau. 2006. Intensity and Frequency of Temperature Extremes over the North Atlantic Region. *Geophysical Research Letters*. *In press*.
- Olsen, J. R., J. R. Stedinger, N. C. Matalas, et E. Z. Stakhiv. 1999. Climate variability and flood frequency estimation for the Upper Mississippi and Lower Missouri Rivers. *Journal of the American Water Resources Association* **35**:1509-1523.
- Ouarda, T. B. M. J., C. Girard, G. Cavadias, et B. Bobee. 2001. Regional flood frequency estimation with canonical correlation analysis. *Journal of Hydrology* **254**:157-173.
- Ouarda, T. B. M. J., M. Lang, B. Bobee, J. Bernier, et P. Bois. 1999. Analysis of regional flood models utilized in France and Quebec. *Revue des Sciences de l'Eau* **12**:155-182.
- Paquet, E. 2004. A new version of the hydrological model MORDOR: snowpack model at different elevations. *Houille Blanche-Revue Internationale De l'Eau*:75-82.
- Pardé, M. 1955. *Fleuves et rivières*. 224 p. Armand Colin, Paris.
- Parent, E., et J. Bernier. 2003. Encoding prior experts judgments to improve risk analysis of extreme hydrological events via POT modeling. *Journal of Hydrology* **283**:1-18.
- Parey, S., F. Malek, C. Laurent, et D. Dacunha-Castelle. 2006. Trends and climate evolution: statistical approach for very high temperatures in France. *submitted to Climatic Change*.
- Parkinson, C. L., D. J. Cavalieri, P. Gloersen, H. J. Zwally, et J. C. Comiso. 1999. Arctic sea ice extents, areas, and trends, 1978-1996. *Journal of Geophysical Research-Oceans* **104**:20837-20856.
- Paturel, J. E., M. Ouedraogo, E. Servat, G. Mahe, A. Dezetter, et J. F. Boyer. 2003. The concept of rainfall and streamflow normals in West and Central Africa in a context of

- climatic variability. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* **48**:125-137.
- Paturel, J. E., E. Servat, B. Kouame, H. Lubes, M. Ouedraogo, et J. M. Masson. 1997. Climatic variability in humid Africa along the Gulf of Guinea .2. An integrated regional approach. *Journal of Hydrology* **191**:16-36.
- Perreault, L. 2000. Analyse bayésienne rétrospective d'une rupture dans les séquences de variables aléatoires hydrologiques. Mémoire de thèse. ENGREF, INRS-Eau. 200 p.
- Perreault, L., J. Bernier, B. Bobee, et E. Parent. 2000a. Bayesian change-point analysis in hydrometeorological time series. Part 1. The normal model revisited. *Journal of Hydrology* **235**:221-241.
- Perreault, L., J. Bernier, B. Bobee, et E. Parent. 2000b. Bayesian change-point analysis in hydrometeorological time series. Part 2. Comparison of change-point models and forecasting. *Journal of Hydrology* **235**:242-263.
- Perreault, L., et V. Fortin. 2003. Mixture and Hidden Markov models for peak flow analysis. *in*. Seizièmes entretiens du centre Jacques Cartier, Lyon, France.
- Perreault, L., M. Hache, M. Slivitzky, et B. Bobee. 1999. Detection of changes in precipitation and runoff over eastern Canada and US using a Bayesian approach. *Stochastic Environmental Research and Risk Assessment* **13**:201-216.
- Perreault, L., E. Parent, J. Bernier, B. Bobee, et M. Slivitzky. 2000c. Retrospective multivariate Bayesian change-point analysis: A simultaneous single change in the mean of several hydrological sequences. *Stochastic Environmental Research and Risk Assessment* **14**:243-261.
- Perrin, C. 2000. Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative. Mémoire de thèse. Institut National Polytechnique de Grenoble, Cemagref Antony. 530 p.
- Peterson, T. C., D. R. Easterling, T. R. Karl, P. Groisman, N. Nicholls, N. Plummer, S. Torok, I. Auer, R. Boehm, D. Gullett, L. Vincent, R. Heino, H. Tuomenvirta, O. Mestre, T. Szentimrey, J. Salinger, E. J. Forland, I. Hanssen-Bauer, H. Alexandersson, P. Jones, et D. Parker. 1998. Homogeneity adjustments of in situ atmospheric climate data: A review. *International Journal of Climatology* **18**:1493-1517.
- Peterson, T. C., et R. S. Vose. 1997. An overview of the global historical climatology network temperature database. *Bulletin of the American Meteorological Society* **78**:2837-2849.
- Pettitt, A. N. 1979. A Non-parametric Approach to the Change-point Problem. *Applied Statistics* **28**:126-135.
- Phoon, K. K., S. T. Quek, et H. W. Huang. 2004. Simulation of non-Gaussian processes using fractile correlation. *Probabilistic Engineering Mechanics* **19**:287-292.
- Pickands, J. 1975. Statistical inference using extreme order statistics. *Annals of Statistics* **3**:119-131.
- Piervitali, E., M. Colacino, et M. Conte. 1997. Signals of climatic change in the central-western Mediterranean basin. *Theoretical and Applied Climatology* **58**:211-219.
- Plummer, N., M. J. Salinger, N. Nicholls, R. Suppiah, K. J. Hennessy, R. M. Leighton, B. Trewin, C. M. Page, et J. M. Lough. 1999. Changes in climate extremes over the Australian region and New Zealand during the twentieth century. *Climatic Change* **42**:183-202.
- Polansky, A. M., et C. E. Check. 2002. Testing trends in environmental compliance. *Journal of Agricultural Biological and Environmental Statistics* **7**:452-468.
- Prudhomme, C. 1995. Modèles synthétiques des connaissances en hydrologie. Mémoire de thèse. Université Montpellier II, Cemagref Lyon. 400 p, Montpellier.
- Prudhomme, C., D. Jakob, et C. Svensson. 2003. Uncertainty and climate change impact on the flood regime of small UK catchments. *Journal of Hydrology* **277**:1-23.

- Radziejewski, M., A. Bardossy, et Z. W. Kundzewicz. 2000. Detection of change in river flow using phase randomization. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* **45**:547-558.
- Radziejewski, M., et Z. W. Kundzewicz. 2004. Detectability of changes in hydrological records. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* **49**:39-51.
- Ramos, M. H. 2002. Analyse de la pluviométrie sous des systèmes nuageux convectifs. Etude de cas sur des données de la ville de Marseille et de la méthode ISIS de Météo-France. Mémoire de thèse. Université Joseph Fourier, LTHE, Grenoble. 165 p.
- Reis, D. S., et J. R. Stedinger. 2005. Bayesian MCMC flood frequency analysis with historical information. *Journal of Hydrology* **313**:97-116.
- Reis, D. S., J. R. Stedinger, et E. S. Martins. 2005. GEV Flood Quantile Estimators with Bayesian Shape-Parameter GLS Regression. *in*. Extreme Value Analysis, Gothenburg.
- Renard, B., V. Garreta, et M. Lang. 2006a. An empirical comparison of MCMC methods used in Bayesian inference. Application for regional trend estimation. *submitted to Water Resources Research*.
- Renard, B., et M. Lang. 2006. Use of copulas for multivariate extreme value analysis: some case studies in hydrology. *submitted to Advances in Water Resources*.
- Renard, B., M. Lang, et P. Bois. 2006b. Statistical analysis of extreme events in a non-stationary context via a Bayesian framework. *Stochastic Environmental Research and Risk Assessment*. *Under press*.
- Ribatet, M., E. Sauquet, J. M. Gresillon, et T. B. M. J. Ouarda. 2006. A regional Bayesian POT model for flood frequency analysis. *Stochastic Environmental Research and Risk Assessment*. *submitted*.
- Robinson, D. A., K. F. Dewey, et R. R. Heim. 1993. Global Snow Cover Monitoring - an Update. *Bulletin of the American Meteorological Society* **74**:1689-1696.
- Robson, A. J. 2002. Evidence for trends in UK flooding. *Philosophical Transactions of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences* **360**:1327-1343.
- Robson, A. J., T. K. Jones, D. W. Reed, et A. C. Bayliss. 1998. A study of national trend and variation in UK floods. *International Journal of Climatology* **18**:165-182.
- Robson, A. J., et D. W. Reed. 1996. Non-stationarity in UK flood records. Report to MAFF Project Centre for Ecology and Hydrology. 63 p.
- Rosbjerg, D., et H. Madsen. 2004. Advanced approaches in PDS/POT modelling of extreme hydrological events. *in*. Hydrology : Science & Practice for the 21st Century, London. British Hydrological Society. 217-221.
- Saporta, G. 1990. *Probabilités, analyse des données et statistiques*. 528 p. Technip.
- Sauquet, E. 2000. Une cartographie des écoulements annuels et mensuels d'un grand bassin versant structurée par la topologie du réseau hydrographique. Mémoire de thèse. Université Joseph Fourier Grenoble, Cemagref Lyon. 356 p, Grenoble.
- Sauquet, E. 2003. Examen de la stationnarité des écoulements du Rhône en lien avec la variabilité climatique et les actions humaines. *in*. Colloque technique barrages et développement durable, MEDD.
- Schalther, M., et J. A. Tawn. 2003. A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika* **90**:139-156.
- Schertzer, D., P. Bernadara, A. Biaou, I. Tchiguirinskaia, M. Lang, E. Sauquet, H. Bendjoudi, P. Hubert, S. Lovejoy, et J. M. Veysseire. 2006. Extrêmes et multifractals en hydrologie: résultats, validations et perspectives. *in*. Valeurs rares et extrêmes de précipitations et de débits, Lyon, France. SHF. 81-90.

- Schönwiese, C. D., et J. Rapp. 1997. *Climate trend atlas of Europe based on observations 1891-1990*. 228 p. Kluwer Academic Publishers.
- Schtroumpf, G. 2006. I win. *Stupid bets with Cemagref friends. :-)*.
- Sen, A., et M. S. Srivastava. 1975. On tests for detecting change in mean. *The Annals of Statistics* **3**:98-108.
- Sen, P. K. 1968. Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association* **63**:1379-1389.
- Servat, E., J. E. Paturel, H. Lubes, B. Kouame, M. Ouedraogo, et J. M. Masson. 1997a. Climatic variability in humid Africa along the Gulf of Guinea .1. Detailed analysis of the phenomenon in Cote d'Ivoire. *Journal of Hydrology* **191**:1-15.
- Servat, E., J. E. Paturel, H. LubesNiel, B. Kouame, M. Travaglio, et B. Marieu. 1997b. Regarding the reduction of flow in Western and Central Africa. *Comptes Rendus de l'Academie des Sciences Serie II Fascicule a-Sciences de la Terre et des Planetes* **325**:679-682.
- Severinghaus, J. P., T. Sowers, E. J. Brook, R. B. Alley, et M. L. Bender. 1998. Timing of abrupt climate change at the end of the Younger Dryas interval from thermally fractionated gases in polar ice. *Nature* **391**:141-146.
- Sharma, A., et U. Lall. 2004. Model averaging and its use in probabilistic forecasting of hydrological variables. *in*. Hydrology: Science & Practice for the 21st Century, London, England. British Hydrological Society. 372-378.
- Sherwood, S., J. Lanzante, et C. Meyer. 2005. Radiosonde daytime biases and late 20th century warming. *Science* **309**.
- Shorthouse, C., et N. W. Arnell. 1999. The effects of climatic variability on spatial characteristics of European river flows. *Physics and Chemistry of the Earth. Part B-Hydrology Oceans and Atmosphere* **24**:7-13.
- Smith, R. L. 1989. Extreme value analysis of environmental time series:an application to trend detection in ground-level ozone. *Statistical Science* **4**:367-393.
- Smith, T. M., et R. W. Reynolds. 2005. A global merged land and sea surface temperature reconstruction based on historical observations (1880-1997). *Journal of Climate* **18**:2021-2036.
- Spagnoli, B., S. Planton, M. Deque, O. Mestre, et J. M. Moisselin. 2002. Detecting climate change at a regional scale: The case of France. *Geophysical Research Letters* **29**.
- Stainforth, D. A., T. Aina, C. Christensen, M. Collins, N. Faull, D. J. Frame, J. A. Kettleborough, S. Knight, A. Martin, J. M. Murphy, C. Piani, D. Sexton, L. A. Smith, R. A. Spicer, A. J. Thorpe, et M. R. Allen. 2005. Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature* **433**:403-406.
- Stedinger, J. R. 1983. Estimating a regional flood frequency distribution. *Water Resources Research* **19**:503-510.
- Stewart, I. T., D. R. Cayan, et M. D. Dettinger. 2005. Changes toward earlier streamflow timing across western North America. *Journal of Climate* **18**:1136-1155.
- Stott, P. A., S. F. B. Tett, G. S. Jones, M. Allen, J. F. B. Mitchell, et G. J. Jenkins. 2000. External control of twentieth century temperature variations by natural and anthropogenic forcings. *Science* **290**:2133-2137.
- Strupczewski, W. G., et Z. Kaczmarek. 2001. Non-stationary approach to at-site flood frequency modelling II. Weighted least squares estimation. *Journal of Hydrology* **248**:143-151.
- Strupczewski, W. G., V. P. Singh, et W. Feluch. 2001. Non-stationary approach to at-site flood frequency modelling I. Maximum likelihood estimation. *Journal of Hydrology* **248**:123-142.

- Tallaksen, L. M., et H. A. J. Van Lanen. 2004. *Hydrological Drought: processes and estimation methods for streamflow and groundwater*. 580 p. Elsevier.
- Tank, A., et G. P. Konnen. 2003. Trends in indices of daily temperature and precipitation extremes in Europe, 1946-99. *Journal of Climate* **16**:3665-3680.
- Tank, A., G. P. Konnen, et F. M. Selten. 2005. Signals of anthropogenic influence on European warming as seen in the trend patterns of daily temperature variance. *International Journal of Climatology* **25**:1-16.
- Tapsoba, D., M. Hache, L. Perreault, et B. Bobee. 2004. Bayesian Rainfall Variability Analysis in West Africa along Cross Sections in Space-Time grid Boxes. *Journal of Climate* **17**:1069-1082.
- Tawn, J. A. 1990. Modelling multivariate extreme value distributions. *Biometrika* **77**:245-253.
- Thyer, M., et G. Kuczera. 2000. Modeling long-term persistence in hydroclimatic time series using a hidden state Markov model. *Water Resources Research* **36**:3301-3310.
- Thyer, M., et G. Kuczera. 2003a. A hidden Markov model for modelling long-term persistence in multi-site rainfall time series 1. Model calibration using a Bayesian approach. *Journal of Hydrology* **275**:12-26.
- Thyer, M., et G. Kuczera. 2003b. A hidden Markov model for modelling long-term persistence in multi-site rainfall time series. 2. Real data analysis. *Journal of Hydrology* **275**:27-48.
- Thyer, M., G. Kuczera, et Q. J. Wang. 2002. Quantifying parameter uncertainty in stochastic models using the Box-Cox transformation. *Journal of Hydrology* **265**:246-257.
- Vincent, C. 2002. Influence of climate change on French glaciers mass balance over the 20th century. *Houille Blanche-Revue Internationale De l'Eau*:20-24.
- Vincent, C., G. Kappenberger, F. Valla, A. Bauder, M. Funk, et E. Le Meur. 2004. Ice ablation as evidence of climate change in the Alps over the 20th century. *Journal of Geophysical Research-Atmospheres* **109**.
- Von Storch, H., et A. Navarra. 1999. *Analysis of Climate Variability*. 342 p. Springer.
- Voss, R., W. May, et E. Roeckner. 2002. Enhanced resolution modelling study on anthropogenic climate change: Changes in extremes of the hydrological cycle. *International Journal of Climatology* **22**:755-777.
- Wang, Q. J. 2001. A Bayesian joint probability approach for flood record augmentation. *Water Resources Research* **37**:1707-1712.
- Wang, S. L., H. J. Jin, S. X. Li, et L. Zhao. 2000. Permafrost degradation on the Qinghai-Tibet Plateau and its environmental impacts. *Permafrost and Periglacial Processes* **11**:43-53.
- Werritty, A. 2002. Living with uncertainty: climate change, river flows and water resource management in Scotland. *Science of the Total Environment* **294**:29-40.
- Wetherald, R. T., et S. Manabe. 1999. Detectability of summer dryness caused by greenhouse warming. *Climatic Change* **43**:495-511.
- Woodward, W. A., et H. L. Gray. 1993. Global Warming and the Problem of Testing for Trend in Time Series Data. *Journal of Climate* **6**:953-962.
- Worsley, K. J. 1979. On the likelihood ratio test for a shift in location of normal populations. *Journal of American Statistical Association* **74**:365 - 367.
- Worsley, K. J. 1983. The power of likelihood ratio and cumulative sum tests for a change in a binomial probability. *Biometrika* **70**:455-464.
- Yagouti, A., I. Abi-Zeid, T. B. M. J. Ouarda, et B. Bobee. 2001. Revue de processus ponctuels et synthèse de tests statistiques pour le choix d'un type de processus. *Revue des Sciences de l'Eau* **14**:323-361.

- Yoshino, F. 1999. Studies on the characteristics of variation and spatial correlation of the long-term annual runoff in the world rivers. *Journal of the Japanese Society for Hydrology and Water Resources* **12**:109-120.
- Yue, S., et P. Pilon. 2003. Interaction between deterministic trend and autoregressive process. *Water Resources Research* **39**.
- Yue, S., et P. Pilon. 2004. A comparison of the power of the t test, Mann-Kendall and bootstrap tests for trend detection. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* **49**:21-37.
- Yue, S., P. Pilon, et G. Cavadias. 2002a. Corrigendum to "Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series (vol 259, pg 254, 2002)". *Journal of Hydrology* **264**:262-263.
- Yue, S., P. Pilon, et G. Cavadias. 2002b. Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. *Journal of Hydrology* **259**:254-271.
- Yue, S., P. Pilon, et B. Phinney. 2003. Canadian streamflow trend detection: impacts of serial and cross-correlation. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* **48**:51-63.
- Yue, S., P. Pilon, B. Phinney, et G. Cavadias. 2002c. The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrological Processes* **16**:1807-1829.
- Yue, S., et C. Y. Wang. 2002a. Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test. *Water Resources Research* **38**.
- Yue, S., et C. Y. Wang. 2002b. Regional streamflow trend detection with consideration of both temporal and spatial correlation. *International Journal of Climatology* **22**:933-946.
- Yue, S., et C. Y. Wang. 2004. The Mann-Kendall test modified by effective sample size to detect trend in serially correlated hydrological series. *Water Resources Management* **18**:201-218.
- Zeng, L., et P. J. Kelly. 1997. A preliminary investigation of the trend of flood events in the United States. *in*. National Association of Real Estate Investment Managers Senior Officer Property Insurance Forum, Boston, US. 6.
- Zhai, P. M., A. J. Sun, F. M. Ren, X. N. Liu, B. Gao, et Q. Zhang. 1999. Chances of climate extremes in China. *Climatic Change* **42**:203-218.
- Zhang, X. B., L. A. Vincent, W. D. Hogg, et A. Niitsoo. 2000. Temperature and precipitation trends in Canada during the 20th century. *Atmosphere-Ocean* **38**:395-429.
- Zhang, X. B., et F. W. Zwiers. 2004. Comment on "Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test" by Sheng Yue and Chun Yuan Wang. *Water Resources Research* **40**.
- Zhang, X. B., F. W. Zwiers, et G. L. Li. 2004. Monte Carlo experiments on the detection of trends in extreme values. *Journal of Climate* **17**:1945-1952.
- Zwiers, F. W., et V. V. Kharin. 1998. Changes in the extremes of the climate simulated by CCC GCM2 under CO2 doubling. *Journal of Climate* **11**:2200-2222.
- Zwiers, F. W., et X. B. Zhang. 2003. Toward regional-scale climate change detection. *Journal of Climate* **16**:793-797.

ANNEXES

ARTICLE 1 : ANALYSE BAYESIENNE NON STATIONNAIRE

Article soumis à *Stochastic Environmental Research and Risk assessment*.

Accepté, sous presse.

Statistical analysis of extreme events in a non-stationary context via a Bayesian framework. Case study with peak-over-threshold data.

B. RENARD¹, M. LANG¹, P. BOIS²

(1) *Cemagref Centre de Lyon, U.R. Hydrologie-Hydraulique, 3 bis Quai Chauveau, CP 220, 69336 Lyon cedex 09, France.*

telephone:33 4 72 20 87 72

fax:33 4 78 47 78 75

e-mail : renard@lyon.cemagref.fr

(2) *LTHE, Laboratoire d'études des Transferts en Hydrologie et Environnement, BP53, 38041 Grenoble cedex 09, France.*

Abstract

Statistical analysis of extremes currently assumes that data arise from a stationary process, although such an hypothesis is not easily assessable and should therefore be considered as an uncertainty. The aim of this paper is to describe a Bayesian framework for this purpose, considering several probabilistic models (stationary, step-change and linear trend models) and four extreme values distributions (exponential, generalized Pareto, Gumbel and GEV). Prior distributions are specified by using regional prior knowledge about quantiles. Posterior distributions are used to estimate parameters, quantify the probability of models and derive a realistic frequency analysis, which takes into account estimation, distribution and stationarity uncertainties. MCMC methods are needed for this purpose, and are described in the article. Finally, an application to a POT discharge series is presented, with an analysis of both occurrence process and peak distribution.

Keywords

Bayesian analysis; extreme events; stationarity; uncertainties; climate change; frequency analysis.

1. Introduction

Since Gumbel's precursory work (Gumbel, 1958), a great number of methods have been developed in the statistics of extreme events studies (see Katz et al. (2002), for a recent review). Among them, Bayesian analysis recently received attention, thanks to advances in computational and numerical tools (Coles and Powell, 1996; Coles and Tawn, 1996; Coles and Pericchi, 2003; Coles et al., 2003; Diebolt et al., 2003; Parent and Bernier, 2003).

In most cases, these methods assume that data are independent and stationary, which means that the statistical properties of the studied process are invariant with respect to time. Nevertheless, this assumption sometimes appears very doubtful. For example, with discharges series, changes in land use can modify the floods and droughts regimes. More generally, in a global climate change context, hydro-meteorological data could exhibit trends or step-changes, as has been demonstrated for global temperature, as an example (IPCC, 2001). Strupczewski et al. (2001) therefore derived a likelihood-based approach to model trends in the distribution parameters, including a criterion for the "best" model choice. Cunderlick and Burn (2003) used the same kind of method in order to include data from several sites and to estimate quantiles from pooled frequency analysis. An alternative of interest was proposed by Perreault et al. (2000a; 2000b), by quantifying the probability of each model in a Bayesian framework, and combining their results with appropriate weightings. In this manner, model choice uncertainty can be taken into account, which seems to be a very pragmatic approach as it is impossible to verify what the "true" model is – and even if such a model exists. However, Perreault et al. made an assumption of normality, which prevents this methodology from being applied to extreme events data.

The aim of this article is to derive a Bayesian framework for extreme events analysis, based on a comparison of models. In section 2, we describe the successive steps of the Bayesian approach. The probabilistic models and prior distributions used in this paper are described in sections 3 and 4. Particular attention will be paid to parameter estimation, detection of changes and frequency analysis (section 5). The particular case of the exponential distribution for peak over threshold (POT) or inter-arrival time data is discussed in section 6. Finally, we present an application to hydrological data, with exponential and generalized Pareto distributions (section 7), before drawing some conclusions and discussing potential improvements (section 8).

2. An overview of the Bayesian approach

A complete and detailed presentation of Bayesian analysis and its applications can be found in Berger (1985) or in Gelman et al. (1995). Figure 1 just shows the general principle of the method. As a first step, probabilistic models for the data must be chosen. In the context of the potential changes described here, these models consist of including trends or step changes in the distribution parameters. Given an independent observations vector \mathbf{X} , it is then possible to compute the likelihood $p(\mathbf{X}|\theta)$ relating to each model, as is usually done in more standard approaches. The fundamental difference of Bayesian analysis concerns the parameters vector θ : it is not viewed as an unknown but constant quantity, which has to be estimated, but as a random vector, whose distribution can be computed from observations. For this purpose, a prior distribution $\pi(\theta)$ is assumed for θ . This distribution reflects the prior knowledge about the parameters, and must be assessed without using the observations. In some cases, the lack of information about a phenomenon leads to the use of so-called non-informative priors. We won't consider this particular case in this paper, thus assuming that prior information is available, however vague it may be. This is not a drastic restriction, as environmental variables can often be connected with not totally unknown physical processes.

Given the likelihood and the prior distribution, Bayes theorem provides the parameters posterior distribution:

$$p(\boldsymbol{\theta}|\mathbf{X}) = \frac{\pi(\boldsymbol{\theta})p(\mathbf{X}|\boldsymbol{\theta})}{\int \pi(\boldsymbol{\theta})p(\mathbf{X}|\boldsymbol{\theta})d\boldsymbol{\theta}} \quad (1)$$

The posterior distribution can be viewed as an updating of the prior information about $\boldsymbol{\theta}$, thanks to the observations. It provides a more complete estimation than the classical point estimates. However, as soon as the dimension of $\boldsymbol{\theta}$ becomes large, the use of this distribution is no longer straightforward, and specific computational methods are needed. The next sections describe the successive steps to achieve a complete analysis framework.

3. Models for extreme values analysis

Table 1 summarizes the probability density functions (pdf) considered in this paper. These distributions arise from theoretical results, thus applying to a wide range of situations involving extreme values. Two main sampling methods are concerned. Firstly, the block maximum consists of selecting the maximum value within a block. Normally, a one year block is chosen, leading to the annual maxima (AM) series. Secondly, peak over threshold (POT) data consists of selecting all the events exceeding a high threshold (see Rosbjerg and Madsen (2004) for an up-to-date review of literature). The extreme value theorem (Fisher and Tippett, 1928) shows that when the block size tends to infinity, the limit distribution, if it exists, of the block maximum series computed from *iid* variables is a GEV distribution. Similarly, when the threshold value increases, the limit distribution of a POT series can be approximated with a Generalized Pareto distribution (Pickands, 1975). In real-world problems, the data from which the extreme variables are extracted may be too complex to meet the requirements of the above theorems. As an example, daily discharges or rainfalls are neither independent nor identically distributed in case of non-stationarity. Moreover, the block size is usually equal to 365 days, which might be insufficient to apply the asymptotic arguments of theorems' proofs. However, these distributions have been extensively applied in studies involving extreme values (e.g. floods, heavy rainfalls, wave heights), and have demonstrated their usefulness in most hydro-meteorological applications.

Moreover, if a Poisson process is assumed for the occurrence of POTs, it can be shown that a Poisson + Exponential model leads to a Gumbel distribution for the annual maxima, whereas Poisson + Generalized Pareto similarly leads to the GEV distribution (Lang et al., 1999). Finally, the inter-arrivals time between two events can be modelled with a one-parameter exponential distribution when a Poisson process is assumed for the occurrences (Lang, 1999).

With our notations, σ is the threshold value, which is supposed to be known, and μ , λ , and ζ are the location, scale and shape parameters. The two-parameter distributions (Exponential and Gumbel) can be considered as limiting cases of the corresponding three-parameter ones (Generalized Pareto and GEV), when $\zeta \rightarrow 0$.

In order to take into account non-stationary data, a step change or a linear trend can be assumed for the parameters, as formalized in table 2. t denotes time, in a discrete (AM) or continuous (POT) way. Notice that the shape parameter ζ is always assumed to be constant (although the same kind of changes could theoretically be modelled), because of the difficulty of precisely estimating this parameter. Moreover, other kinds of changes could be modelled, such as a linear trend on the log-transformed scale parameter (Katz et al., 2002), non-simultaneous step changes in the location and the scale parameters, multiple step changes or polynomial trends. Although the methodology presented here can take into account such kind of models, it always has to be kept in mind that complex models can improve the description

of a phenomenon, but can dramatically decrease the accuracy of the estimations or increase the numerical difficulties.

Finally, for each of the 12 models presented in table 2, the likelihood of an observed vector $\mathbf{X} = (x_1, \dots, x_n)$ can be computed as follows:

$$p(\mathbf{X} | \boldsymbol{\theta}) = \prod_{i=1}^n p(x_i | \boldsymbol{\theta}) \quad (2)$$

4. Specifying priors

Scientists usually have to estimate some environmental characteristics at one site of interest, without any data being available at this particular location. Standard statistical analysis is then impossible, and alternative methods have been developed for this purpose. Regionalization can thus be used to transfer knowledge from gauged to ungauged sites (GREHYS, 1996; Javelle et al., 2002; Madsen et al., 2002). A more physical approach consists of deducing some estimates from underlying physical processes. As an example, in the hydrology field, a large number of methods have been used to relate stream flow to explanatory variables, such as rainfall, topography of the watershed, soil properties, etc (see Reed (1999) for an example in UK).

When specifying priors, the situation is the same, as data must not be used at this step. The non-exhaustive above list of examples shows that regional knowledge is often available. Such information just has to be transferred to the probabilistic model parameters. In extreme values analysis, a variable of current concern is the p -quantile q_p , that is the value which is exceeded with probability $(1-p)$. Coles and Tawn (1996) thus proposed a method for producing prior distributions from these quantiles. Let $p(x|\boldsymbol{\theta})$ be one of the stationary pdf (models M_θ), with dimension k ($\boldsymbol{\theta} \in \mathbb{R}^k$). Let $p_1 < \dots < p_k$, and $q_{p_1} < \dots < q_{p_k}$ be the corresponding quantiles, which can be expressed as $\boldsymbol{\theta}$ and p functions. As an example, with the M_0^{exp} model, $q_p = \sigma - \lambda \log(1-p)$. As a natural order exists between q_{p_1}, \dots, q_{p_k} , thus leading to strongly non-independent values, prior production is done using the following modified quantiles:

$$\tilde{q}_1 = q_{p_1}, \tilde{q}_2 = q_{p_2} - q_{p_1}, \dots, \tilde{q}_k = q_{p_k} - q_{p_{k-1}} \quad (3)$$

Let $G(y|a, b)$ be the pdf of a gamma distribution with parameters a and b :

$$G(y|a, b) = \frac{b^a}{\Gamma(a)} y^{a-1} e^{-by} 1_{\{t>0\}}(y) \quad (4)$$

where $\Gamma(\cdot)$ denotes the gamma function and $1_\Omega(y)$ is the indicator function of the Ω set ($1_\Omega(y) = 1$ if $y \in \Omega$, 0 otherwise). The $(\tilde{q}_i)_{i=1, \dots, k}$ are supposed to follow independent gamma distributions:

$$\pi(\tilde{q}_1, \dots, \tilde{q}_k) = G(\tilde{q}_1 | a_1, b_1) \dots G(\tilde{q}_k | a_k, b_k) \quad (5)$$

The gamma distribution is used because of its flexibility, but alternative choices are possible. The hypothesis of independence is used here for convenience. Nevertheless, this is not a drastic restriction, as the joint prior distribution must reflect diffuse knowledge, which will be updated in the following step. Moreover, structural dependence between parameters will be preserved when transferring prior knowledge from the quantiles space to the parameters space.

The prior distribution is finally computed by replacing the \tilde{q}_i by their expression in terms of θ , and multiplying by the Jacobian of the transformation $\theta \mapsto (\tilde{q}_1, \dots, \tilde{q}_k)$:

$$\pi(\theta) = G(\tilde{q}_1(\theta) | a_1, b_1) \dots G(\tilde{q}_k(\theta) | a_k, b_k) |\mathbf{J}| \quad (6)$$

where \mathbf{J} is the Jacobian matrix of the above transformation, and the couples (a_i, b_i) are called the hyperparameters of the model, which have to be fixed at a particular value, reflecting the prior knowledge about the modified quantiles $(\tilde{q}_i)_{i=1, \dots, k}$. Tables 3a and 3b show the $q_p(\theta)$ mathematical expressions in each model M_0 , M_1 and M_2 , and the corresponding Jacobian matrixes. The determinants of these matrixes can be computed numerically with any mathematical software.

In the case of a stationary model M_0 , prior specification is thus related to at most three modified quantiles $(\tilde{q}_1, \tilde{q}_2, \tilde{q}_3)$. Prior specification can be achieved in the same manner with non-stationary models M_1 and M_2 , despite an increase in the θ dimension. In the step change model (M_1), a discrete distribution $\pi(\tau)$ has to be assumed for the change point τ (uniform on $1, \dots, n-1$ for instance). Let r be an integer such as $1 \leq r < k$. Prior specifications are the following:

$$\left\{ \begin{array}{l} \text{Before change, } (\tilde{q}_i^{(1)})_{i=1, \dots, r} \text{ follow independent gamma distributions.} \\ \text{After change, } (\tilde{q}_i^{(2)})_{i=1, \dots, k-r} \text{ follow independent gamma distributions.} \end{array} \right.$$

Assuming independence between the change point τ and the \tilde{q}_i , prior distribution is then:

$$\begin{aligned} \pi(\theta) = & G(\tilde{q}_1^{(1)}(\theta) | a_1^{(1)}, b_1^{(1)}) \dots G(\tilde{q}_r^{(1)}(\theta) | a_r^{(1)}, b_r^{(1)}) \\ & \times G(\tilde{q}_1^{(2)}(\theta) | a_1^{(2)}, b_1^{(2)}) \dots G(\tilde{q}_{k-r}^{(2)}(\theta) | a_{k-r}^{(2)}, b_{k-r}^{(2)}) \pi(\tau) |\mathbf{J}| \end{aligned} \quad (7)$$

If no prior knowledge about the existence of a change is available, the same specifications for quantiles with same probability before and after change can be chosen:

$$(a_i^{(1)}, b_i^{(1)}) = (a_i^{(2)}, b_i^{(2)}), \forall i = 1, \dots, \min(r, k-r) \quad (8)$$

If a trend is assumed in the parameters (M_2), let $t_1 < t_2$ be two time points, and $1 \leq r < k$. Prior specifications are then:

$$\left\{ \begin{array}{l} \text{At } t = t_1, (\tilde{q}_i(t_1))_{i=1, \dots, r} \text{ follow independent gamma distributions.} \\ \text{At } t = t_2, (\tilde{q}_i(t_2))_{i=1, \dots, k-r} \text{ follow independent gamma distributions.} \end{array} \right.$$

The prior distribution can thus be written:

$$\begin{aligned} \pi(\theta) = & G(\tilde{q}_1(\theta(t_1)) | a_1^{(1)}, b_1^{(1)}) \dots G(\tilde{q}_r(\theta(t_1)) | a_r^{(1)}, b_r^{(1)}) \\ & \times G(\tilde{q}_1(\theta(t_2)) | a_1^{(2)}, b_1^{(2)}) \dots G(\tilde{q}_{k-r}(\theta(t_2)) | a_{k-r}^{(2)}, b_{k-r}^{(2)}) |\mathbf{J}| \end{aligned} \quad (9)$$

As previously explained in the M_1 case, equation (8) can be used to complete the prior specifications.

At last, the approach proposed by Coles and Tawn (1996) was exposed here because regional or empirical knowledge is often expressed in terms of quantiles values, rather than parameters values. However, alternative prior distributions can be used without additional difficulties within the presented framework.

5. Computations on posterior distributions

The methods described in sections 3 and 4 allow Bayes theorem to be used to compute the posterior distribution $p(\boldsymbol{\theta}|\mathbf{X})$. However, the integral of the denominator of equation (1) is generally intractable, $p(\boldsymbol{\theta}|\mathbf{X})$ is thus only known up to a constant of proportionality:

$$p(\boldsymbol{\theta} | \mathbf{X}) \propto p(\mathbf{X} | \boldsymbol{\theta})\pi(\boldsymbol{\theta}) \quad (10)$$

Moreover, the high dimensionality of $p(\boldsymbol{\theta}|\mathbf{X})$ makes this function difficult to handle in practice. Next sections present some methods to overcome this drawback. Readers interested in further exploration of posterior distributions than those described below can refer to Tanner (1996).

5.1. Parameters estimation via marginal densities

The $p(\boldsymbol{\theta}|\mathbf{X})$ j^{th} marginal is defined by:

$$p(\theta_j | \mathbf{X}) = \int p(\theta_1, \dots, \theta_k | \mathbf{X}) d\theta_1 \dots d\theta_{j-1} d\theta_{j+1} \dots d\theta_k \quad (11)$$

Once again, this integration is often intractable. No analytical expression of the posterior marginal is then available. An alternative is to simulate a sample from the joint posterior distribution $p(\boldsymbol{\theta}|\mathbf{X})$. Two main methods can be used for this purpose: the Metropolis-Hastings algorithm (Metropolis and Ulam, 1949; Metropolis et al., 1953; Hastings, 1970) and the gridy Gibbs sampling (Ritter and Tanner, 1992). The latter is conceptually easier to understand, and is thus presented in detail in Annex 1, although the former method may be more efficient in some situations.

Given a sample of N vectors $(\boldsymbol{\theta}^{(i)})_{i=1, \dots, N} = (\theta_1^{(i)}, \dots, \theta_k^{(i)})_{i=1, \dots, N}$, a sample from the j^{th} marginal is simply obtained by considering $(\theta_j^{(i)})_{i=1, \dots, N}$. Estimation can then be made by means of this sample: the mean or the median gives a point estimate, and a confidence interval can be calculated by computing lower and upper empirical quantiles. Notice that the values $(\theta_j^{(i)})_{i=1, \dots, N}$ are not independent, a large sample size is thus required to avoid any bias due to autocorrelation. A sensitivity analysis could be necessary to determine a critical sample size, above which estimates stay stable.

5.2. Detection of changes

We applied the methodology described by Perreault et al. (2000b), based on Bayes factors (Kass and Raftery, 1995), to detect changes in environmental series. Let us define the marginal distribution of \mathbf{X} as follow:

$$p(\mathbf{X} | M) = \int p(\mathbf{X}, \boldsymbol{\theta} | M) d\boldsymbol{\theta} \quad (12)$$

A simple calculation shows that $p(\mathbf{X}|M)$ is the denominator of the Bayes theorem applied to a particular model M :

$$p(\mathbf{X} | M) = \int p(\mathbf{X} | \boldsymbol{\theta}, M) \pi(\boldsymbol{\theta} | M) d\boldsymbol{\theta} \quad (13)$$

As stated previously, this quantity can generally not be explicitly computed. Chib (1995) proposed a simple method for estimating the marginal distribution of \mathbf{X} , based on gridy Gibbs sampling. A detailed description of this approach is made in Annex 2. Let $\Omega = \{M_1, \dots, M_q\} = \Omega_1 \cup \Omega_1^c$, where the exponent c denotes the complement of a set, be a

partition of the collection of models considered. As an example, $\Omega = \{M_0^{\text{exp}}, M_1^{\text{exp}}, M_2^{\text{exp}}, M_0^{\text{gp}}, M_1^{\text{gp}}, M_2^{\text{gp}}\} = \{M_0^{\text{exp}}, M_0^{\text{gp}}\} \cup \{M_1^{\text{exp}}, M_1^{\text{gp}}, M_2^{\text{exp}}, M_2^{\text{gp}}\}$ could be a partition for POT data models. More generally, the partition $\Omega_1 \cup \Omega_1^c$ can take the form “stationary vs non stationary models”, or “2 parameters vs 3 parameters models”. Assume that $\pi(\cdot)$ is a prior probability distribution on Ω , the posterior probability of a model M_j can then be computed with the Bayes theorem:

$$p(M_j | \mathbf{X}) = \frac{p(\mathbf{X} | M_j)\pi(M_j)}{\sum_{i=1}^q p(\mathbf{X} | M_i)\pi(M_i)} \quad (14)$$

The Bayes factor between models M_i and M_j is then defined by:

$$B_{i,j} = \left(\frac{p(M_i | \mathbf{X})}{p(M_j | \mathbf{X})} \right) / \left(\frac{\pi(M_i)}{\pi(M_j)} \right) = \frac{p(\mathbf{X} | M_i)}{p(\mathbf{X} | M_j)} \quad (15)$$

A composite Bayes factor between Ω_1 et Ω_1^c can also be computed:

$$B_{\Omega_1, \Omega_1^c} = \left(\frac{\sum_{M \in \Omega_1} p(M | \mathbf{X})}{\sum_{M \in \Omega_1^c} p(M | \mathbf{X})} \right) / \left(\frac{\sum_{M \in \Omega_1} \pi(M)}{\sum_{M \in \Omega_1^c} \pi(M)} \right) \quad (16)$$

The Bayes factor provides a measurement of the pertinence of a hypothesis compared to another. It can thus be compared to classical likelihood ratios. However, model likelihood is not computed by replacing the parameters by their ML estimates, but by integrating on the prior distribution. More generally, this approach is not a classical test procedure. The Bayes factor value depends on prior specifications in addition to sampling variance. Thus, it cannot be interpreted as a risk in terms of limit error percentage when resampling size tends to infinity, but rather in terms of a false decision cost (see Perreault (2000) for further explanations). Classically, the two error types are assumed to have the same cost. The Bayes factor is then compared to 1, with greater certainty for high or near zero values (Kass and Raftery, 1995).

5.3. Frequency analysis.

Section 5.1 provides a way of simulating a sample for the $p(\theta | \mathbf{X})$ distribution, which relates to estimation uncertainty. Section 5.2 describes the computing of posterior probabilities of the models: as a collection of models cannot be considered totally exhaustive, a part of the modelling uncertainty is thus controlled. These two kinds of uncertainty can be taken into account in frequency analysis, by the following procedure:

Repeat N times:

1. Random choice of a model M^* , with probabilities $p(M_1 | \mathbf{X}), \dots, p(M_k | \mathbf{X})$.
2. Random choice of a parameter vector θ^* from the sample of $p(\theta | \mathbf{X}, M^*)$.
3. Computation of $q_p(t)$, the p -quantile in model M^* with parameter θ^* at time t .

The quantiles collection obtained can be considered as a sample of the quantile distribution at time t , including estimation and modelling uncertainties. It can thus be used to provide a point estimate, a confidence interval or even a distribution estimate for $q_p(t)$. If data

stationarity is questionable, the function $t \mapsto q_p(t)$ provides a description of the possible change in terms of quantile drift with time.

6. A particular case: the exponential distribution

Up to this point, the methods used in this paper have been described as generally as possible. The framework is thus still applicable with other models. Nevertheless, distributions belonging to the exponential family can be treated in a simpler way, with analytical results being available without using MCMC methods. In the next sections, posterior marginal distributions and the marginal distribution of observations are computed for exponential models, leading to substantial simplifications. Notice that Parent and Bernier (2003) also succeeded in deriving a MCMC-free method in the case of a stationary generalized Pareto distribution.

6.1. Stationary model

The likelihood for M_0^{exp} model is:

$$p(\mathbf{X} | \boldsymbol{\theta}) = \prod_{i=1}^n p(x_i | \boldsymbol{\theta}) = \frac{1}{\lambda^n} \exp\left(-\frac{n(\bar{X} - \sigma)}{\lambda}\right) \quad (17)$$

Viewed as a function of λ , this likelihood is proportional to an inverse gamma pdf $IG(y|a,b)$:

$$IG(y|a,b) = \frac{b^a}{\Gamma(a)} \frac{1}{y^{a+1}} \exp\left(-\frac{b}{y}\right) 1_{\{t>0\}}(y) \quad (18)$$

This particularity is known as the conjugacy property. Thus, if $\pi(\lambda)$ is supposed to follow an inverse gamma distribution with parameters a and b , the posterior distribution of λ is:

$$\begin{aligned} p(\lambda | \mathbf{X}) &\propto \pi(\lambda) p(\mathbf{X} | \lambda) \\ &= \frac{b^a}{\Gamma(a)} \frac{1}{\lambda^{n+a+1}} \exp\left(-\frac{b+n(\bar{X}-\sigma)}{\lambda}\right) \\ &= f(\lambda) \end{aligned} \quad (19)$$

Up to proportionality, λ posterior distribution is still an inverse gamma distribution, with updated parameters:

$$\begin{aligned} a' &= a + n \\ b' &= b + n(\bar{X} - \sigma) \end{aligned} \quad (20)$$

The marginal distribution of observations is finally simply obtained by integrating $f(\lambda)$:

$$\begin{aligned} p(\mathbf{X} | M_0^{\text{exp}}) &= \int f(\lambda) d\lambda \\ &= \frac{b^a}{\Gamma(a)} \int \frac{1}{\lambda^{a+1}} \exp\left(-\frac{b'}{\lambda}\right) d\lambda \\ &= \frac{b^a}{\Gamma(a)} \frac{\Gamma(a')}{b'^{a'}} \end{aligned} \quad (21)$$

6.2. Step change model

Following the table 2 notations, the M_1^{exp} likelihood is:

$$p(\mathbf{X} | \boldsymbol{\theta}) = \left(\frac{1}{\lambda_1}\right)^\tau \left(\frac{1}{\lambda_2}\right)^{n-\tau} \exp\left(-\sum_{i=1}^{\tau} \frac{x_i - \sigma}{\lambda_1}\right) \exp\left(-\sum_{i=\tau+1}^n \frac{x_i - \sigma}{\lambda_2}\right) \quad (22)$$

For fixed τ , the likelihood is proportional to the product of two inverse gamma distributions. Assuming independence between the change point and λ_i values, the prior distribution is:

$$\pi(\boldsymbol{\theta}) = IG(\lambda_1 | a_1, b_1) IG(\lambda_2 | a_2, b_2) \pi(\tau) \quad (23),$$

where $\pi(\tau)$ is any discrete distribution on $1, \dots, n-1$. Once again, a simple calculation shows that, for a fixed τ the posterior distribution is still proportional to a product of inverse gamma densities:

$$\begin{aligned} p(\boldsymbol{\theta} | \mathbf{X}) &\propto IG(\lambda_1 | a_1, b_1) IG(\lambda_2 | a_2, b_2) \pi(\tau) p(\mathbf{X} | \boldsymbol{\theta}) \\ &= \frac{b_1^{a_1}}{\Gamma(a_1)} \frac{b_2^{a_2}}{\Gamma(a_2)} IG(\lambda_1 | a'_1, b'_1) IG(\lambda_2 | a'_2, b'_2) \pi(\tau) \\ &= f(\lambda_1, \lambda_2, \tau) \end{aligned} \quad (24)$$

with:

$$\begin{aligned} a'_1 &= a_1 + \tau \\ a'_2 &= a_2 + n - \tau \\ b'_1 &= b_1 + \sum_{i=1}^{\tau} (x_i - \sigma) \\ b'_2 &= b_2 + \sum_{i=\tau+1}^n (x_i - \sigma) \end{aligned} \quad (25)$$

The marginal posterior distribution can directly be computed by integration from the joint posterior distribution:

$$\begin{aligned}
 p(\tau | \mathbf{X}) &\propto \int f(\lambda_1, \lambda_2, \tau) d\lambda_1 d\lambda_2 \\
 &\propto \frac{\Gamma(a'_1)\Gamma(a'_2)}{(b'_1)^{a'_1} (b'_2)^{a'_2}} \pi(\tau) \\
 p(\lambda_1 | \mathbf{X}) &\propto \int f(\lambda_1, \lambda_2, \tau) d\lambda_2 d\tau \\
 &\propto \sum_{\tau=1}^{n-1} \frac{\Gamma(a'_2)}{(b'_2)^{a'_2}} \pi(\tau) \left(\frac{1}{\lambda_1}\right)^{a'_1+1} \exp\left(-\frac{b'_1}{\lambda_1}\right) \\
 p(\lambda_2 | \mathbf{X}) &\propto \int f(\lambda_1, \lambda_2, \tau) d\lambda_1 d\tau \\
 &\propto \sum_{\tau=1}^{n-1} \frac{\Gamma(a'_1)}{(b'_1)^{a'_1}} \pi(\tau) \left(\frac{1}{\lambda_2}\right)^{a'_2+1} \exp\left(-\frac{b'_2}{\lambda_2}\right)
 \end{aligned} \tag{26}$$

Finally, the marginal distribution of observations is:

$$\begin{aligned}
 p(\mathbf{X} | M_1^{\text{exp}}) &= \int f(\lambda_1, \lambda_2, \tau) d\lambda_1 d\lambda_2 d\tau \\
 &= \sum_{\tau=1}^{n-1} \frac{\Gamma(a'_1)\Gamma(a'_2)}{\Gamma(a_1)\Gamma(a_2)} \frac{(b_1)^{a_1} (b_2)^{a_2}}{(b'_1)^{a'_1} (b'_2)^{a'_2}} \pi(\tau)
 \end{aligned} \tag{27}$$

6.3. Trend model

The likelihood can be written as:

$$p(\mathbf{X} | \boldsymbol{\theta}) = \left(\frac{1}{\lambda_0}\right)^n \prod_{i=1}^n \left(\frac{1}{1 + \lambda_1 t_i}\right) \exp\left(-\frac{1}{\lambda_0} \sum_{i=1}^n \frac{x_i - \sigma}{1 + \lambda_1 t_i}\right) \tag{28}$$

For fixed λ_i , this likelihood viewed as a λ_0 function is proportional to an inverse gamma pdf. Unfortunately, no conjugacy property appears for the λ_i parameter. Any prior distribution can thus be chosen for λ_i , for instance a normal distribution with parameters m and s . The joint prior distribution is then:

$$\pi(\boldsymbol{\theta}) = IG(\lambda_0 | a, b) N(\lambda_1 | m, s) \tag{29},$$

leading to the following posterior distribution:

$$\begin{aligned}
 p(\boldsymbol{\theta} | \mathbf{X}) &\propto p(\mathbf{X} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) \\
 &= \left(\frac{1}{\lambda_0}\right)^{n+a+1} \frac{b^a}{\Gamma(a)} \frac{1}{s\sqrt{2\pi}} \prod_{i=1}^n \left(\frac{1}{1 + \lambda_1 t_i}\right) \exp\left(-\frac{1}{\lambda_0} \left(b + \sum_{i=1}^n \frac{x_i - \sigma}{1 + \lambda_1 t_i}\right) - \frac{(\lambda_1 - m)^2}{2s^2}\right) \\
 &= f(\lambda_0, \lambda_1)
 \end{aligned} \tag{30}$$

The f function can be explicitly integrated with respect to λ_0 , thus giving the marginal posterior distribution of λ_i :

$$\begin{aligned}
p(\lambda_1 | \mathbf{X}) &\propto \int f(\lambda_0, \lambda_1) d\lambda_0 \\
&= \frac{b^a}{\Gamma(a)} \frac{\Gamma(n+a)}{\left(b + \sum_{i=1}^n \frac{x_i - \sigma}{1 + \lambda_1 t_i}\right)^{n+a}} \frac{1}{s\sqrt{2\pi}} \prod_{i=1}^n \left(\frac{1}{1 + \lambda_1 t_i}\right) \exp\left(-\frac{(\lambda_1 - m)^2}{2s^2}\right) \\
&= f_1(\lambda_1)
\end{aligned} \tag{31}$$

The marginal posterior distribution of λ_0 and the marginal distribution of observations can be computed by 1-D numerical integration:

$$\begin{aligned}
p(\lambda_0 | \mathbf{X}) &\propto \int f(\lambda_0, \lambda_1) d\lambda_1 \\
p(\mathbf{X} | M_2^{\text{exp}}) &= \int f(\lambda_0, \lambda_1) d\lambda_0 d\lambda_1 \\
&= \int f_1(\lambda_1) d\lambda_1
\end{aligned} \tag{32}$$

7. Application

The Bayesian method was applied to a stream flow series from the Drôme river at Luc en Diois (194 km², South East of France). Daily stream flows were available between the years 1907 and 2003. We extracted a series of independent peaks above a 23 m³.s⁻¹ threshold, leading to a sample of 93 values. Over-threshold values and inter-arrivals series were studied (figure 2). The former was studied with the six models $M_0^{\text{exp}}, M_1^{\text{exp}}, M_2^{\text{exp}}, M_0^{\text{gp}}, M_1^{\text{gp}}$ and M_2^{gp} , the latter with $M_0^{\text{exp}}, M_1^{\text{exp}}$ and M_2^{exp} .

7.1. Prior specification

The first step consisted of specifying a discrete prior distribution on the models collection we considered. In order to avoid any bias in the hypothesis we wished to evaluate (stationary vs. non-stationary and Exponential vs. Generalized Pareto), the following probabilities were chosen:

$$\begin{cases} \pi(M_0^{\text{exp}}) = \pi(M_0^{\text{gp}}) = 0.25 \\ \pi(M_1^{\text{exp}}) = \pi(M_2^{\text{exp}}) = \pi(M_1^{\text{gp}}) = \pi(M_2^{\text{gp}}) = 0.125 \end{cases}$$

for the over-threshold series, and

$$\begin{cases} \pi(M_0^{\text{exp}}) = 0.5 \\ \pi(M_1^{\text{exp}}) = \pi(M_2^{\text{exp}}) = 0.25 \end{cases}$$

for the inter-arrivals series. In this manner, each hypothesis had an equal prior probability.

The second step consisted of evaluating our prior knowledge on quantiles, as explained in section 4. We applied the methodology proposed by Galéa and Prudhomme (1997) and Javelle et al. (1999), known as the QdF (flow-duration-frequency) models. As a first step, we estimated the instantaneous discharge quantile with probability 0.9 by using the Crupedix formula (CTGREF et al., 1980-1982):

$$q_{0,9} = S^{0.8} \left(\frac{r_{0,9}}{80} \right)^2 R \quad (33)$$

where S is the watershed area, $r_{0,9}$ is the daily rainfall quantile with probability 0.9, and R is a regional index. The Socose duration D (CTGREF et al., 1980-1982), which characterizes the flood duration, was then computed:

$$\log(D) = -0.69 + 0.32 \log(S) + 2.2 \left(\frac{\bar{r}}{r_{0,9} \cdot \bar{t}} \right)^{1/2} \quad (34)$$

where \bar{r} is the mean annual rainfall and \bar{t} the mean annual temperature. All the parameters needed for the calculation of $q_{0,9}$ and D are available on maps. The QdF methodology can then be used to deduce flood quantiles with other durations or frequencies, by comparison with three reference data sets, representative of the three main groups of hydrological flood regimes in France. The reference data set suitable for the Drôme river had already been chosen by Prudhomme (1995).

We thus obtained 0.9- and 0.99-quantiles of daily stream flows, leading to \hat{q}_1 and \hat{q}_2 values as defined in section 4. In order to fit a two-parameters prior distribution to each modified quantile, an indication about their variability was also needed. In the Crupedix method, a 90% confidence interval is $[q/2; 2q]$. Because of the lack of information about the final uncertainty in our prior specification process, we decided to use the same shape of confidence intervals for the modified quantiles. We thus evaluated the hyperparameters a and b of the Gamma distribution, so that the mean prior value of \tilde{q}_i was equal to \hat{q}_i and the 0.95-quantile of the gamma distribution was equal to $2 \times \hat{q}_i$. In the exponential case, the simplified procedure described in section 6 was applied. The estimates $\hat{\lambda} = \frac{\sigma - \hat{q}_1}{\log(1 - p_1)}$ and

$\hat{\lambda}' = \frac{\sigma - 2 \times \hat{q}_1}{\log(1 - p_1)}$ were computed, and the hyperparameters a and b of the Inverse Gamma

distribution were evaluated so that the mean was equal to $\hat{\lambda}$ and the 0.95-quantile was equal to $\hat{\lambda}'$. Finally, the same value of the a hyperparameter was chosen for the prior distribution relative to inter-arrivals (leading to the same coefficient of variation), and the b hyperparameter was computed to obtain a mean value of 1 for the λ parameter (leading to one event per year on average). Equation (8) was used in models M_1 and M_2 . Table 4 gives the results of the prior specification process.

7.2. Results

Table 5 summarizes the models' posterior probabilities and associated Bayes factors. According to the scale proposed by Kass and Raftery (1995), evidence for non-stationarity of the inter-arrivals series appears to be positive, a linear trend on the exponential distribution being the most probable model. Figure 3 shows the process mean for each of the three models. For this purpose, parameter estimates are the means of marginal posterior distributions (figure 4), except for the change point in model M_1^{exp} , which is the mode of the discrete posterior distribution.

No clear evidence appears concerning the stationarity of the over-threshold series, with a Bayes factor near to 1. Conversely, the generalized Pareto distribution seems to be most efficient, as shown by figure 5 for models M_0^{exp} and M_0^{gp} . Surprisingly, the model M_1^{gp} appears to be the more probable, but does not provide an accurate estimate of the change-point (Figure 6). One possible explanation could be the sensitivity of Chib approach to the starting point: the accuracy of the method may be lower when no clear isolated mode appears in the posterior distribution of τ . Nevertheless, the posterior distribution of the scale parameter after change is very close to that of the scale parameter in model M_0^{gp} (Figure 7). Consequently, results of the frequency analysis should also be very close for these two models.

Figure 8 shows the result of such an analysis. In order to take into account changes in the inter-arrivals process Y (which is supposed to be independent from the over-thresholds X), the procedure presented in section 5.3 was modified as follows:

Repeat N times:

1. Random choice of a model M^* for the over-thresholds, with probabilities $p(M_1|X), \dots, p(M_k|X)$.
2. Random choice of a parameter vector θ^* from the sample of $p(\theta|X, M^*)$.
3. Random choice of a model M^+ for the inter-arrivals, with probabilities $p(M_1|Y), \dots, p(M_k|Y)$.
4. Random choice of a parameter vector θ^+ from the sample of $p(\theta|Y, M^+)$.
5. Computation of $\mu(t)$ ($=1/\lambda(t)$), the mean number of events per year in model M^+ at time t , with parameter θ^+ .
6. Computation of the probability relative to a return period T at time t : $p = 1 - 1/T\mu(t)$.
7. Computation of $q_p(t)$, the p -quantile in model M^* with parameter θ^* at time t .

Flood quantiles have decreased by about 19% between years 1907 and 2003. Figure 9 focuses on the trajectory of the 10-year return flood. The curve shape in the early years is due to the step-change model: once the posterior probability of a step change has become too weak, this model assumes that the process will stay in the second state, thus behaving like a stationary model. The curve is thus smoothed after such a date. Prediction of quantile future values must be regarded with caution, because we neglect the uncertainty due to change persistency in the future. Nevertheless, such a figure emphasizes the ambiguity relative to the use of “return periods” for characterizing frequency in a non-stationary context. In such a case, it would be better to use the 0.9-quantile, which has a 0.1 probability of being exceeded in the next year, this value being different each year.

8. Conclusion and discussion

The aim of this article was to derive a Bayesian framework to deal with data for which stationarity is questionable. The method presented here can take into account modelling uncertainties, in order to obtain a more realistic quantification of the results accuracy. This should result in wider confidence intervals, compared to methods ignoring the modelling uncertainties. Conversely, prior specifications can be used to inject physical or regional knowledge about a phenomenon, which may increase the accuracy of estimates. Coles and Powell (1996) thus noted that prior spatial information can be very useful for estimating scale and shape parameters of a GEV distribution, even with few data, whereas the location parameter is mostly site-specific and data-dependent. More information about the advantages and drawbacks of Bayesian methods, compared to likelihood-based ones, in extreme values analysis are needed. More precisely, future studies could compare the accuracy of estimates, the robustness to outliers, and investigate the influence of prior specifications (or misspecifications).

Although we focused on 12 particular models, including two sampling methods, two kinds of distributions and three change possibilities, the framework can easily be used in other contexts, like polynomial changes, multiple step changes, seasonal distributions mixture, use of covariates, etc. Some hypotheses we made could also be relaxed by means of appropriate models. Firstly, we assumed that data arise from an independent sample, which can sometimes be very doubtful. ARMA models are then good candidates to overcome this difficulty (see Ray and Tsay (2002) for instance). Secondly, we assumed any change to be final. Thus, when forecasting a future quantile value, the uncertainty due to change persistency is not taken into account. In the step-change case, Hidden Markov Chains can be used for this purpose (Robert et al., 2000; Perreault and Fortin, 2003). Nevertheless, most of the published approaches assumed data normality. Some improvements are thus still foreseeable in these topics.

Finally, one of the more promising tasks would be to derive a Bayesian multivariate method for extreme values analysis. More precisely, in a global climate change context, at-site analysis is not satisfactory. Unfortunately, generalization of existing methods (Perreault et al., 2000c) is not straightforward for extreme values distributions. The use of copulas (Favre et al., 2004) might be considered for a restricted number of locations. Adaptation of the flood-duration-frequency approach (Javelle et al., 2002) in a non-stationary context is also one of the possibilities. At last, hierarchical Bayesian models, including trends or step-changes parameters, may be useful for regional analysis of extremes (Cooley, 2005; Cooley et al., 2005), as well as for measurement uncertainties handling (Reis and Stedinger, 2005).

Acknowledgements

This work was conducted as part of a national program of hydrological research (PNRH), which associates Cemagref (Lyon), LTHE (Grenoble), Hydrosiences (Montpellier), Meteo France (Toulouse), and Electricité de France (EDF Chatou and Grenoble). The authors would like to thank all members of this project. The financial support provided by Cemagref and EDF for the PhD research of B. Renard is gratefully acknowledged. We also thank two anonymous reviewers for their helpful comments.

Bibliography

- Berger, J. O. 1985. Statistical decision theory and bayesian analysis. Springer-Verlag. 617 p. New York.
- Chib, S. 1995. Marginal Likelihood From the Gibbs Output. *Journal of American Statistical Association* **90**:1313-1321.
- Coles, S., and L. Pericchi. 2003. Anticipating catastrophes through extreme value modelling. *Journal of the Royal Statistical Society Series C-Applied Statistics* **52**:405-416.
- Coles, S., L. R. Pericchi, and S. Sisson. 2003. A fully probabilistic approach to extreme rainfall modelling. *Journal of Hydrology* **273**:35-50.
- Coles, S. G., and E. A. Powell. 1996. Bayesian methods in extreme value modelling: A review and new developments. *International Statistical Review* **64**:119-136.
- Coles, S. G., and J. A. Tawn. 1996. A Bayesian analysis of extreme rainfall data. *Applied Statistics-Journal of the Royal Statistical Society Series C* **45**:463-478.
- Cooley, D. 2005. *Statistical Analysis of Extremes Motivated by Weather and Climate Studies: Applied and Theoretical Advances*. University of Colorado. 122 p.
- Cooley, D., D. Nychka, and P. Naveau. 2005. A Spatial Bayesian Hierarchical Model for a Precipitation Return Levels Map. *in*. *Extreme Value Analysis*, Gothenburg, Sweden.

- CTGREF, Srae, and S. H. Diame. 1980-1982. Synthèse nationale sur les crues des petits bassins versants. Fascicule 2 : la méthode Socose ; Information Technique n° 38-2 (Juin 1980) ; Fascicule 3 : la méthode Crupedix.
- Cunderlik, J. M., and D. H. Burn. 2003. Non-stationary pooled frequency analysis. *Journal of Hydrology* **276**:210-223.
- Diebolt, J., M. A. El-Aroui, M. Garrido, and S. Girard. 2003. Quasi-conjugate Bayes estimates for GPD parameters and application to heavy tails modelling. Rapport de recherche INRIA. 29 p.
- Favre, A. C., S. El Adlouni, L. Perreault, N. Thiemonge, and B. Bobee. 2004. Multivariate hydrological frequency analysis using copulas. *Water Resources Research* **40**.
- Fisher, R. A., and L. H. Tippett. 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Cambridge Phil. Soc.* **24**.
- Galéa, G., and C. Prudhomme. 1997. Notions de base et concepts utiles pour la compréhension de la modélisation synthétique des régimes de crue des bassins versants au sens des modèles QdF. *Revue des Sciences de l'Eau* **1**:83-101.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. Bayesian data analysis. C. Hall. 526 p. Chapman & Hall.
- GREHYS. 1996. Presentation and review of some methods for regional flood frequency analysis. *Journal of Hydrology* **186**:63-84.
- Gumbel, E. J. 1958. Statistics of extremes. 375 p. Columbia University Press, New York.
- Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**:97-109.
- IPCC. 2001. Climate change 2001: synthesis report. Cambridge University Press. 408 p, Cambridge.
- Javelle, P., J. M. Grésillon, and G. Galéa. 1999. Discharge-duration-Frequency curves modeling for floods and scale invariance. *Comptes Rendus de l'Académie des Sciences, Sciences de la terre et des planètes.* **329**:39-44.
- Javelle, P., T. Ouarda, M. Lang, B. Bobee, G. Galéa, and J. M. Grésillon. 2002. Development of regional flood-duration-frequency curves based on the index-flood method. *Journal of Hydrology* **258**:249-259.
- Kass, R. E., and A. E. Raftery. 1995. Bayes Factors. *Journal of the American Statistical Association* **90**:773-795.
- Katz, R. W., M. B. Parlange, and P. Naveau. 2002. Statistics of extremes in hydrology. *Advances in Water Resources* **25**:1287-1304.
- Lang, M. 1999. Theoretical discussion and Monte-Carlo simulations for a Negative Binomial process paradox. *Stochastic Environmental Research and Risk Assessment* **13**:183-200.
- Lang, M., T. B. M. J. Ouarda, and B. Bobée. 1999. Towards operational guidelines for over-threshold modeling. *Journal of Hydrology* **225**:103-117.
- Madsen, H., P. S. Mikkelsen, D. Rosbjerg, and H. P. 2002. Regional estimation of rainfall intensity-duration-frequency curves using generalized least squares regression of partial duration series statistics. *Water Resources Research* **38**:1239.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *Journal of chemical physics* **21**:1087-1092.
- Metropolis, N., and S. Ulam. 1949. The Monte Carlo method. *Journal of the American Statistical Association* **44**:335-341.
- Parent, E., and J. Bernier. 2003. Encoding prior experts judgments to improve risk analysis of extreme hydrological events via POT modeling. *Journal of Hydrology* **283**:1-18.
- Perreault, L. 2000. Analyse bayésienne rétrospective d'une rupture dans les séquences de variables aléatoires hydrologiques. ENGREF, INRS-Eau. 200 p.

- Perreault, L., J. Bernier, B. Bobee, and E. Parent. 2000a. Bayesian change-point analysis in hydrometeorological time series. Part 1. The normal model revisited. *Journal of Hydrology* **235**:221-241.
- Perreault, L., J. Bernier, B. Bobee, and E. Parent. 2000b. Bayesian change-point analysis in hydrometeorological time series. Part 2. Comparison of change-point models and forecasting. *Journal of Hydrology* **235**:242-263.
- Perreault, L., and V. Fortin. 2003. Mixture and Hidden Markov models for peak flow analysis. *in*. Seizièmes entretiens du centre Jacques Cartier, Lyon, France.
- Perreault, L., E. Parent, J. Bernier, B. Bobee, and M. Slivitzky. 2000c. Retrospective multivariate Bayesian change-point analysis: A simultaneous single change in the mean of several hydrological sequences. *Stochastic Environmental Research and Risk Assessment* **14**:243-261.
- Pickands, J. 1975. Statistical inference using extreme order statistics. *Annals of Statistics* **3**:119-131.
- Prudhomme, C. 1995. Modèles synthétiques des connaissances en hydrologie. Université Montpellier II, CEMAGREF Lyon. 400 p, Montpellier.
- Ray, B. K., and R. S. Tsay. 2002. Bayesian methods for change-point detection in long-range dependent processes. *Journal of Time Series Analysis* **23**:687-705.
- Reed, D. W. 1999. Flood Estimation Handbook. Volume 1: Overview. I. o. Hydrology. 108 p. Wallingford.
- Reis, D. S., and J. R. Stedinger. 2005. Bayesian MCMC flood frequency analysis with historical information. *Journal of Hydrology* **313**:97-116.
- Ritter, C., and M. A. Tanner. 1992. Facilitating the Gibbs Sampler - the Gibbs Stopper and the Griddy-Gibbs Sampler. *Journal of the American Statistical Association* **87**:861-868.
- Robert, C. P., T. Ryden, and D. M. Titterton. 2000. Bayesian inference in hidden Markov models through the reversible jump Markov chain Monte Carlo method. *Journal of the Royal Statistical Society Series B-Statistical Methodology* **62**:57-75.
- Rosbjerg, D., and H. Madsen. 2004. Advanced approaches in PDS/POT modelling of extreme hydrological events. *in*. *Hydrology : Science & Practice for the 21st Century*, London. B. H. Society. 217-221.
- Strupczewski, W. G., and Z. Kaczmarek. 2001. Non-stationary approach to at-site flood frequency modelling II. Weighted least squares estimation. *Journal of Hydrology* **248**:143-151.
- Strupczewski, W. G., V. P. Singh, and W. Feluch. 2001. Non-stationary approach to at-site flood frequency modelling I. Maximum likelihood estimation. *Journal of Hydrology* **248**:123-142.
- Tanner, M. A. 1996. *Tools for Statistical Inference*. Springer-Verlag. 208 p. New York.

Annex 1. Griddy Gibbs sampling

The aim of this method is to simulate a sample from the posterior distribution $p(\boldsymbol{\theta}|\mathbf{X})$. The Gibbs sampling algorithm can be written as follow:

- Choose a starting value $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \dots, \theta_k^{(0)})$, and set $j = 0$.
- Repeat N times :
 - $j = j+1$;
 - Sample $\theta_1^{(j)} \sim p(\theta_1 | \theta_2^{(j-1)}, \dots, \theta_k^{(j-1)}, \mathbf{X})$
 - Sample $\theta_2^{(j)} \sim p(\theta_2 | \theta_1^{(j)}, \theta_3^{(j-1)}, \dots, \theta_k^{(j-1)}, \mathbf{X})$
 - ...
 - Sample $\theta_q^{(j)} \sim p(\theta_q | \theta_1^{(j)}, \dots, \theta_{q-1}^{(j)}, \theta_{q+1}^{(j-1)}, \dots, \theta_k^{(j-1)}, \mathbf{X})$
 -
 - Sample $\theta_k^{(j)} \sim p(\theta_k | \theta_1^{(j)}, \dots, \theta_{k-1}^{(j)}, \mathbf{X})$
- end

The vectors series $(\boldsymbol{\theta}^{(j)})$ converges to the target posterior distribution as j tends to infinity. In order to decrease the influence of the starting point, the first m iterations are usually deleted, and inference is made with the last $N-m$ iterations. Sensitivity analysis may be necessary to determine acceptable values for m and N (usually, at least a few thousands of iterations are used for m and N). More generally, this method presents the same drawbacks as other iterative simulation techniques. Convergence thus has to be monitored, by choosing several starting points or computing convergence indices. Gelman et al. (1995) or Tanner (1996) provide some guidelines for improving the numerical simulations efficiency.

Unfortunately, the Gibbs sampling algorithm can usually not be used in this raw version, because it involves being able to sample from the full conditional densities $p(\theta_q | \theta_1^{(j)}, \dots, \theta_{q-1}^{(j)}, \theta_{q+1}^{(j-1)}, \dots, \theta_k^{(j-1)}, \mathbf{X})$. Ritter and Tanner (1992) proposed the use of a discrete approximation of the cumulative density functions (cdf) of these distributions:

- Choose a grid of points y_1, \dots, y_p
- Evaluate $p(\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_q, \mathbf{X})$ on this grid, to obtain w_1, \dots, w_p
- Compute the cumulative sums of w_1, \dots, w_p to obtain an approximation of the CDF.
- Sample u from a uniform distribution on $[0, 1]$.
- Transform u by the inverse of the approximate CDF.

This algorithm must be added at each step of the Gibbs iteration. If the full conditional densities are known only up to proportionality, $p(\theta_q | \theta_1^{(j)}, \dots, \theta_{q-1}^{(j)}, \theta_{q+1}^{(j-1)}, \dots, \theta_k^{(j-1)}, \mathbf{X})$ can be replaced by $f(\theta_1^{(j)}, \dots, \theta_{q-1}^{(j)}, y, \theta_{q+1}^{(j-1)}, \dots, \theta_k^{(j-1)})$, the product of prior and likelihood, and the cumulative sums of w_1, \dots, w_p must then be divided by the total sum to provide the CDF approximation. Inversion of the CDF can be made with a linear interpolation between two grid points. The grid choice is the most important issue of this technique: it has to be broad enough to cover the range of the distributions, and fine enough to ensure a sufficient accuracy, keeping in mind that this supplementary step is computing-time expensive. Some improvements of the method are described in Ritter and Tanner (1992).

Annex 2. Chib method

The aim of this method is to compute the marginal distribution of the observations, which is the normalizing constant of the Bayes theorem:

$$p(\mathbf{X} | M) = \frac{\pi(\boldsymbol{\theta})p(\mathbf{X} | \boldsymbol{\theta})}{p(\boldsymbol{\theta} | \mathbf{X})} = \frac{f(\boldsymbol{\theta})}{p(\boldsymbol{\theta} | \mathbf{X})} \quad (\text{A1})$$

This relationship being true for any vector $\boldsymbol{\theta}$, let us consider a particular $\boldsymbol{\theta}^* = (\theta_1^*, \dots, \theta_k^*)$. $f(\boldsymbol{\theta}^*)$ is directly computable, which is not the case of the denominator. Consider the following relationship:

$$\begin{aligned} p(\theta_1^*, \dots, \theta_k^* | \mathbf{X}) &= p(\theta_1^* | \mathbf{X})p(\theta_2^* | \theta_1^*, \mathbf{X}) \dots p(\theta_q^* | \theta_1^*, \theta_2^*, \dots, \theta_{q-1}^*, \mathbf{X}) \\ &\dots \times p(\theta_k^* | \theta_1^*, \dots, \theta_{k-1}^*, \mathbf{X}) \end{aligned} \quad (\text{A2})$$

The first term can be evaluated thanks to the sample of the first marginal distribution, by using a Gaussian kernel as an example. The last term can be computed by 1-D numerical integration:

$$p(\theta_k^* | \theta_1^*, \dots, \theta_{k-1}^*, \mathbf{X}) = \frac{f(\boldsymbol{\theta}^*)}{\int f(\theta_1^*, \dots, \theta_{k-1}^*, \theta_k) d\theta_k} \quad (\text{A3})$$

The gridy Gibbs sampling can be used to compute intermediary terms. $p(\theta_q^* | \theta_1^*, \theta_2^*, \dots, \theta_{q-1}^*, \mathbf{X})$ is indeed the first marginal of the distribution $p(\theta_q, \theta_{q+1}, \dots, \theta_k | \theta_1^*, \theta_2^*, \dots, \theta_{q-1}^*, \mathbf{X})$, evaluated at θ_q^* . The gridy Gibbs algorithm can thus be applied to the non-normalized posterior density with the first $q-1$ components being fixed, that is $f(\theta_1^*, \theta_2^*, \dots, \theta_{q-1}^*, \theta_q, \theta_{q+1}, \dots, \theta_k)$.

Although this approach is theoretically valid for any value $\boldsymbol{\theta}^*$ with non-zero posterior probability, Chib recommends the use of a high-density point to increase the method accuracy.

List of tables

Table 1 : Probability density functions used in extreme values analysis

Table 2: Models of changes on the distributions parameters

Table 3a: Quantiles calculation and Jacobian matrixes for prior specifications : Exponential (a), Generalized Pareto (b) and Gumbel (c) distributions

Table 3b: Quantiles calculation and Jacobian matrixes for prior specifications : GEV distribution

Table 4 : Prior specification for the Drôme river at Luc en Diois

Table 5: Posterior probabilities of models and relatives Bayes factors

List of figures

Figure 1: Bayesian analysis overview

Figure 2: POT events for the Drôme river at Luc en Diois (a), and associated peaks (b) and inter-arrivals (c) series.

Figure 3: Mean of the inter-arrivals process as a function of time, for models M_0^{exp} , M_1^{exp} and M_2^{exp} (resp. dotted, dashed and solid line)

Figure 4: Marginal densities of the inter-arrivals process parameters, for models M_0^{exp} (a), M_1^{exp} (b) and M_2^{exp} (c)

Figure 5: Quantiles curves with models M_0^{exp} and M_0^{gp} (resp. solid and dashed line). Empirical frequencies are represented by crosses

Figure 6: Prior (dotted line) and posterior (bars) marginal densities of the change point in model M_1^{gp}

Figure 7: Posterior marginal densities of the scale parameters in models M_0^{gp} and M_1^{gp} (after change).

Figure 8: Quantiles mean curves (solid lines) and 90% confidence intervals (dotted lines) at dates 1907 (grey) and 2003 (black)

Figure 9: Evolution of the 10-year return flood (solid line) with time and 90% confidence interval (dotted lines)

	POT data	Block Maximum data
2-parameters distributions	<p>Exponential</p> $p(x \theta) = \frac{1}{\lambda} \exp\left(-\frac{x-\sigma}{\lambda}\right)$ <p>$\theta = (\sigma, \lambda)$</p> <p>$\sigma$ known (threshold value)</p> <p>$\lambda > 0$</p>	<p>Gumbel</p> $p(x \theta) = \frac{1}{\lambda} \exp\left(-\frac{x-\mu}{\lambda} - \exp\left(-\frac{x-\mu}{\lambda}\right)\right)$ <p>$\theta = (\mu, \lambda)$</p> <p>$\mu > 0$</p> <p>$\lambda > 0$</p>
	<p>Generalized Pareto</p> $p(x \theta) = \frac{1}{\lambda} \left(1 - \frac{\xi(x-\sigma)}{\lambda}\right)^{\frac{1}{\xi}-1}$ <p>$\theta = (\sigma, \lambda, \xi)$</p> <p>$\sigma$ known (threshold value)</p> <p>$\lambda > 0$</p> <p>$\xi \neq 0$</p> <p>$1 - \frac{\xi(x-\sigma)}{\lambda} > 0$</p>	<p>Generalized Extreme Value (GEV)</p> $p(x \theta) = \frac{1}{\lambda} \left(1 - \frac{\xi(x-\mu)}{\lambda}\right)^{\frac{1}{\xi}-1} \exp\left(-\left(1 - \frac{\xi(x-\mu)}{\lambda}\right)^{\frac{1}{\xi}}\right)$ <p>$\theta = (\mu, \lambda, \xi)$</p> <p>$\mu > 0$</p> <p>$\lambda > 0$</p> <p>$\xi \neq 0$</p> <p>$1 - \frac{\xi(x-\mu)}{\lambda} > 0$</p>

Table 1

Sampling method	Distribution	Model		
		stationary	step-change	linear trend
POT	Exponential	$\mathbf{M}_0^{\text{exp}}$ $\theta(t)=\theta$	$\mathbf{M}_1^{\text{exp}}$ $\lambda(t) = \begin{cases} \lambda_1 & \text{if } t \leq \tau \\ \lambda_2 & \text{if } t > \tau \end{cases}$	$\mathbf{M}_2^{\text{exp}}$ $\lambda(t) = \lambda_0 (1 + \lambda_1 t)$
		\mathbf{M}_0^{gp} $\theta(t)=\theta$	\mathbf{M}_1^{gp} $(\lambda(t), \xi(t)) = \begin{cases} (\lambda_1, \xi) & \text{if } t \leq \tau \\ (\lambda_2, \xi) & \text{if } t > \tau \end{cases}$	\mathbf{M}_2^{gp} $\lambda(t) = \lambda_0 (1 + \lambda_1 t)$ $\xi(t) = \xi$
		\mathbf{M}_0^{gu} $\theta(t)=\theta$	\mathbf{M}_1^{gu} $(\mu(t), \lambda(t)) = \begin{cases} (\mu_1, \lambda_1) & \text{if } t \leq \tau \\ (\mu_2, \lambda_2) & \text{if } t > \tau \end{cases}$	\mathbf{M}_2^{gu} $\mu(t) = \mu_0 (1 + \mu_1 t)$ $\lambda(t) = \lambda_0 (1 + \lambda_1 t)$
Block Maximum	GEV	$\mathbf{M}_0^{\text{gev}}$ $\theta(t)=\theta$	$\mathbf{M}_1^{\text{gev}}$ $(\mu(t), \lambda(t), \xi(t)) = \begin{cases} (\mu_1, \lambda_1, \xi) & \text{if } t \leq \tau \\ (\mu_2, \lambda_2, \xi) & \text{if } t > \tau \end{cases}$	$\mathbf{M}_2^{\text{gev}}$ $\mu(t) = \mu_0 (1 + \mu_1 t)$ $\lambda(t) = \lambda_0 (1 + \lambda_1 t)$ $\xi(t) = \xi$

Table 2

		Model		
		stationary	step-change	linear trend
(a)	$\mathbf{M}_0^{\text{exp}}$ $q_p(t) = \sigma - \lambda \log(1-p)$ Specification on \tilde{q}_1 $J = -\log(1-p_1)$	$\mathbf{M}_1^{\text{exp}}$ $q_p^{(i)}(t) = \sigma - \lambda_i \log(1-p)$ Specification on $(\tilde{q}_1^{(1)}, \tilde{q}_1^{(2)})$ $J = \begin{pmatrix} -\log(1-p_1) & 0 \\ 0 & -\log(1-p_1) \end{pmatrix}$	$\mathbf{M}_2^{\text{exp}}$ $q_p(t) = \sigma - \lambda_0(1 + \lambda_1 t) \log(1-p)$ Specification on $(\tilde{q}_1(t_1), \tilde{q}_1(t_2))$ $J = \begin{pmatrix} -(1 + \lambda_1 t_1) \log(1-p_1) & -\lambda_0 t_1 \log(1-p_1) \\ -(1 + \lambda_1 t_2) \log(1-p_1) & -\lambda_0 t_2 \log(1-p_1) \end{pmatrix}$	
	(b)	\mathbf{M}_0^{gp} $q_p(t) = \sigma + \frac{\lambda}{\xi}(1 - (1-p)^\xi)$ $y_i = (1-p_i)^\xi$ Specification on $(\tilde{q}_1, \tilde{q}_2)$ $J = \begin{pmatrix} \frac{1-y_1}{\xi} & \frac{\lambda}{\xi^2}(y_1(1-\log(y_1))-1) \\ \frac{y_1-y_2}{\xi} & \frac{\lambda}{\xi^2}(y_1(\log(y_1)-1)-y_2(\log(y_2)-1)) \\ 0 & \frac{1-y_1}{\xi} \end{pmatrix}$	\mathbf{M}_1^{gp} $q_p^{(i)}(t) = \sigma + \frac{\lambda_i}{\xi}(1 - (1-p)^\xi)$ Specification on $(\tilde{q}_1^{(1)}, \tilde{q}_2^{(1)}, \tilde{q}_1^{(2)})$ $J = \begin{pmatrix} \frac{1-y_1}{\xi} & 0 & \frac{\lambda_1}{\xi^2}(y_1(1-\log(y_1))-1) \\ \frac{y_1-y_2}{\xi} & 0 & \frac{\lambda_1}{\xi^2}(y_1(\log(y_1)-1)-y_2(\log(y_2)-1)) \\ 0 & \frac{1-y_1}{\xi} & \frac{\lambda_2}{\xi^2}(y_1(1-\log(y_1))-1) \end{pmatrix}$	\mathbf{M}_2^{gp} $q_p(t) = \sigma + \frac{\lambda_0(1 + \lambda_1 t)}{\xi}(1 - (1-p)^\xi)$ Specification on $(\tilde{q}_1(t_1), \tilde{q}_2(t_1), \tilde{q}_1(t_2))$ $J = \begin{pmatrix} \frac{(1 + \lambda_1 t_1)(1-y_1)}{\xi} & \frac{\lambda_0 t_1(1-y_1)}{\xi} & \frac{\lambda_0(1 + \lambda_1 t_1)}{\xi^2}(y_1(1-\log(y_1))-1) \\ \frac{(1 + \lambda_1 t_1)(y_1-y_2)}{\xi} & \frac{\lambda_0 t_1(y_1-y_2)}{\xi} & \frac{\lambda_0(1 + \lambda_1 t_1)}{\xi^2}(y_1(\log(y_1)-1)-y_2(\log(y_2)-1)) \\ \frac{(1 + \lambda_1 t_2)(1-y_1)}{\xi} & \frac{\lambda_0 t_2(1-y_1)}{\xi} & \frac{\lambda_0(1 + \lambda_1 t_2)}{\xi^2}(y_1(1-\log(y_1))-1) \end{pmatrix}$
		(c)	\mathbf{M}_0^{gu} $q_p(t) = \mu - \lambda \log(-\log(p))$ Specification on $(\tilde{q}_1, \tilde{q}_2)$ $J = \begin{pmatrix} 1 & -\log(-\log(p_1)) \\ 0 & \log\left(\frac{\log(p_1)}{\log(p_2)}\right) \end{pmatrix}$	\mathbf{M}_1^{gu} $q_p^{(i)}(t) = \mu_i - \lambda_i \log(-\log(p))$ Specification on $(\tilde{q}_1^{(1)}, \tilde{q}_2^{(1)}, \tilde{q}_1^{(2)}, \tilde{q}_2^{(2)})$ $J = \begin{pmatrix} 1 & -\log(-\log(p_1)) & 0 & 0 \\ 0 & \log\left(\frac{\log(p_1)}{\log(p_2)}\right) & 0 & 0 \\ 0 & 0 & 1 & -\log(-\log(p_1)) \\ 0 & 0 & 0 & \log\left(\frac{\log(p_1)}{\log(p_2)}\right) \end{pmatrix}$

Table 3a

	$\mathbf{M}_0^{\text{gev}}$
	$q_p(t) = \mu + \frac{\lambda}{\xi} \left(1 - (-\log(p))^\xi\right)$
	<i>Specification on $(\tilde{q}_1, \tilde{q}_2, \tilde{q}_3)$</i>
	$y_i = -\log(p_i)$
stationary	$\mathbf{J} = \begin{pmatrix} 1 & \frac{1-y_1^\xi}{\xi} & & \frac{\lambda}{\xi^2} (y_1^\xi (1-\log(y_1^\xi)) - 1) \\ 0 & \frac{y_1^\xi - y_2^\xi}{\xi} & & \frac{\lambda}{\xi^2} (y_1^\xi (\log(y_1^\xi) - 1) - y_2^\xi (\log(y_2^\xi) - 1)) \\ 0 & \frac{y_2^\xi - y_3^\xi}{\xi} & & \frac{\lambda}{\xi^2} (y_2^\xi (\log(y_2^\xi) - 1) - y_3^\xi (\log(y_3^\xi) - 1)) \end{pmatrix}$
	$\mathbf{M}_1^{\text{gev}}$
	$q_p^{(i)}(t) = \mu_i + \frac{\lambda_i}{\xi} \left(1 - (-\log(p))^\xi\right)$
	<i>Specification on $(\tilde{q}_1^{(1)}, \tilde{q}_2^{(1)}, \tilde{q}_3^{(1)}, \tilde{q}_1^{(2)}, \tilde{q}_2^{(2)})$</i>
Step change	$\mathbf{J} = \begin{pmatrix} 1 & 0 & \frac{1-y_1^\xi}{\xi} & 0 & \frac{\lambda_1}{\xi^2} (y_1^\xi (1-\log(y_1^\xi)) - 1) \\ 0 & 0 & \frac{y_1^\xi - y_2^\xi}{\xi} & 0 & \frac{\lambda_1}{\xi^2} (y_1^\xi (\log(y_1^\xi) - 1) - y_2^\xi (\log(y_2^\xi) - 1)) \\ 0 & 0 & \frac{y_2^\xi - y_3^\xi}{\xi} & 0 & \frac{\lambda_1}{\xi^2} (y_2^\xi (\log(y_2^\xi) - 1) - y_3^\xi (\log(y_3^\xi) - 1)) \\ 0 & 1 & 0 & \frac{1-y_1^\xi}{\xi} & \frac{\lambda_2}{\xi^2} (y_1^\xi (1-\log(y_1^\xi)) - 1) \\ 0 & 0 & 0 & \frac{y_1^\xi - y_2^\xi}{\xi} & \frac{\lambda_2}{\xi^2} (y_1^\xi (\log(y_1^\xi) - 1) - y_2^\xi (\log(y_2^\xi) - 1)) \end{pmatrix}$
	$\mathbf{M}_2^{\text{gev}}$
	$q_p(t) = \mu_0(1 + \mu_1 t) + \frac{\lambda_0(1 + \lambda_1 t)}{\xi} \left(1 - (-\log(p))^\xi\right)$
	<i>Specification on $(\tilde{q}_1(t_1), \tilde{q}_2(t_1), \tilde{q}_3(t_1), \tilde{q}_1(t_2), \tilde{q}_2(t_2))$</i>
linear trend	$\mathbf{J} = \begin{pmatrix} 1+\mu_1 t_1 & \mu_0 t_1 & \frac{(1+\lambda_1 t_1)(1-y_1^\xi)}{\xi} & \frac{\lambda_0 t_1 (1-y_1^\xi)}{\xi} & \frac{\lambda_0 (1+\lambda_1 t_1)}{\xi^2} (y_1^\xi (1-\log(y_1^\xi)) - 1) \\ 0 & 0 & \frac{(1+\lambda_1 t_1)(y_1^\xi - y_2^\xi)}{\xi} & \frac{\lambda_0 t_1 (y_1^\xi - y_2^\xi)}{\xi} & \frac{\lambda_0 (1+\lambda_1 t_1)}{\xi^2} (y_1^\xi (\log(y_1^\xi) - 1) - y_2^\xi (\log(y_2^\xi) - 1)) \\ 0 & 0 & \frac{(1+\lambda_1 t_1)(y_2^\xi - y_3^\xi)}{\xi} & \frac{\lambda_0 t_1 (y_2^\xi - y_3^\xi)}{\xi} & \frac{\lambda_0 (1+\lambda_1 t_1)}{\xi^2} (y_2^\xi (\log(y_2^\xi) - 1) - y_3^\xi (\log(y_3^\xi) - 1)) \\ 1+\mu_1 t_2 & \mu_0 t_2 & \frac{(1+\lambda_1 t_2)(1-y_1^\xi)}{\xi} & \frac{\lambda_0 t_2 (1-y_1^\xi)}{\xi} & \frac{\lambda_0 (1+\lambda_1 t_2)}{\xi^2} (y_1^\xi (1-\log(y_1^\xi)) - 1) \\ 0 & 0 & \frac{(1+\lambda_1 t_2)(y_1^\xi - y_2^\xi)}{\xi} & \frac{\lambda_0 t_2 (y_1^\xi - y_2^\xi)}{\xi} & \frac{\lambda_0 (1+\lambda_1 t_2)}{\xi^2} (y_1^\xi (\log(y_1^\xi) - 1) - y_2^\xi (\log(y_2^\xi) - 1)) \end{pmatrix}$

Table 3b

	Peaks				Inter-arrivals		
	Exponential	Generalized Pareto		Change point	Exponential	Change point	Trend coefficient
Parameters	$\hat{\lambda}$	\hat{q}_1	\hat{q}_2	$\hat{\tau}$	$\hat{\lambda}$	$\hat{\tau}$	$\hat{\lambda}_1$
	16.025	60 m ³ .s ⁻¹	42 m ³ .s ⁻¹	no prior information	1	no prior information	no prior information
Hyper-parameters	$IG(\lambda a, b)$ $a=2.5$ $b=24$	$G(\tilde{q}_1 a, b)$ $a=3.6$ $b=3.6/60$	$G(\tilde{q}_2 a, b)$ $a=3.6$ $b=3.6/42$	uniform on [1, ..., n-1]	$IG(\lambda a, b)$ $a=2.5$ $b=1.5$	uniform on [1, ..., n-1]	$N(\lambda_1 m, s)$ $m=0$ $s=0.05$

Table 4

	M_0^{exp}	M_1^{exp}	M_2^{exp}	M_0^{gp}	M_1^{gp}	M_2^{gp}	non-stationary vs. stationary	Generalized Pareto vs. exponential
Inter-arrivals	0.16	0.21	0.63	/	/	/	5.4	/
Peaks	0.15	0.04	0.01	0.28	0.5	0.02	1.3	4.1

Table 5

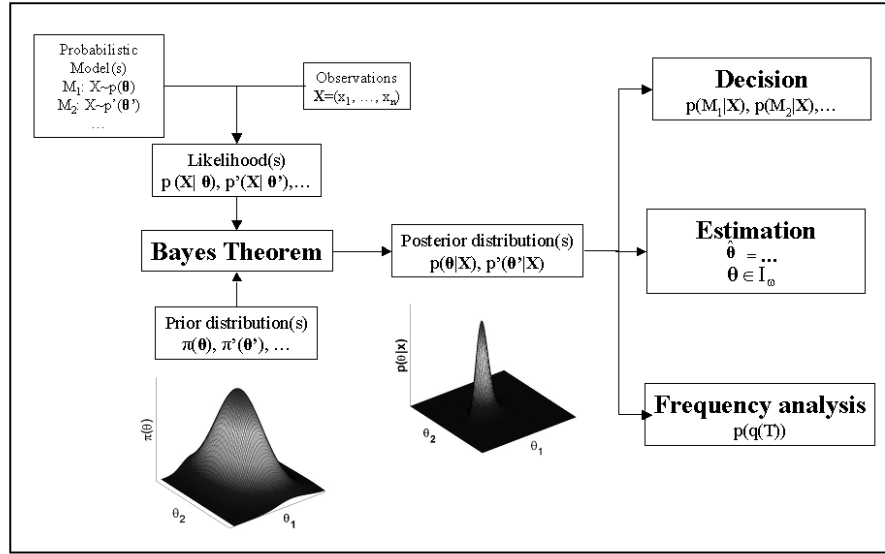


Figure 1

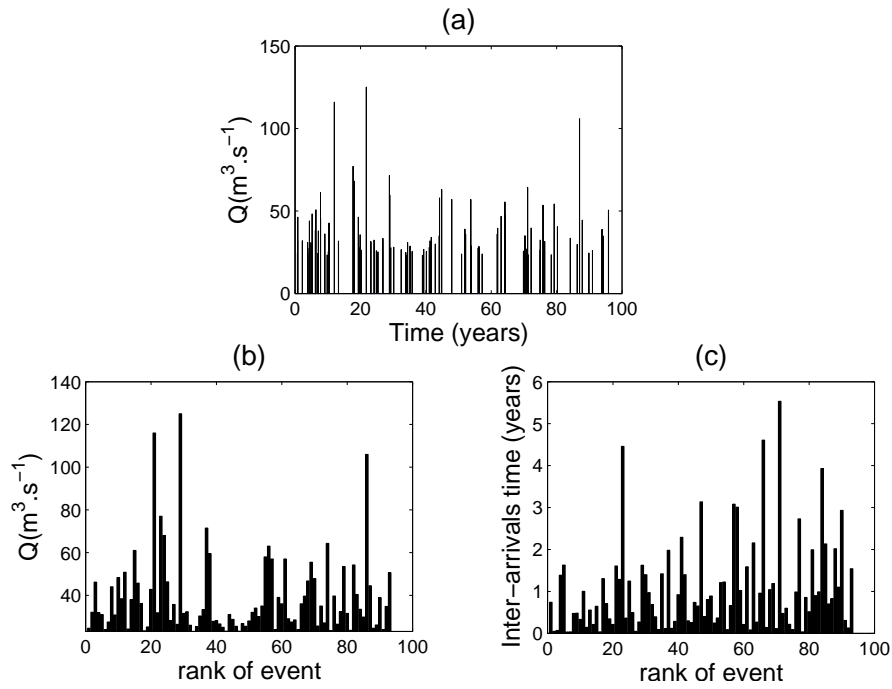


Figure 2

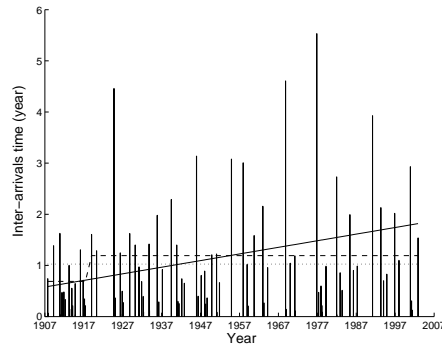


Figure 3

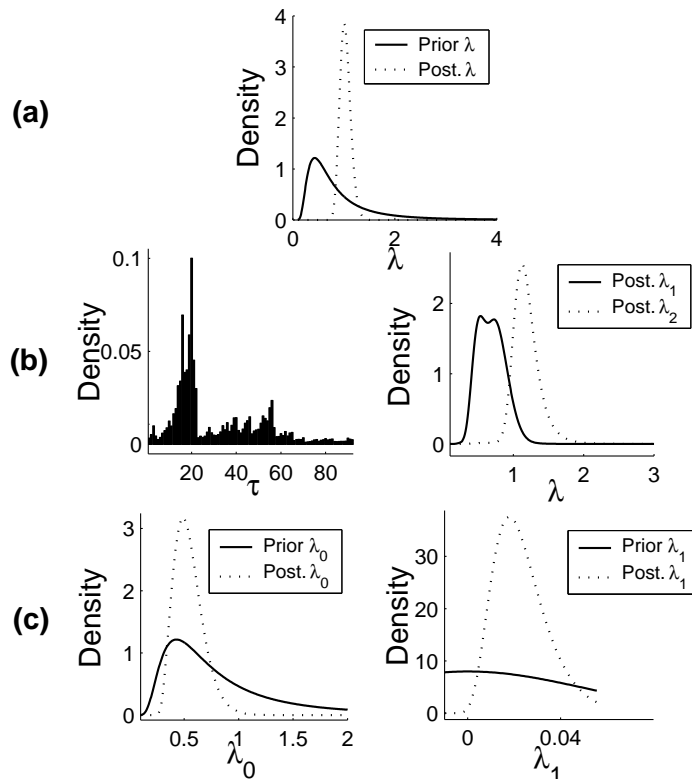


Figure 4

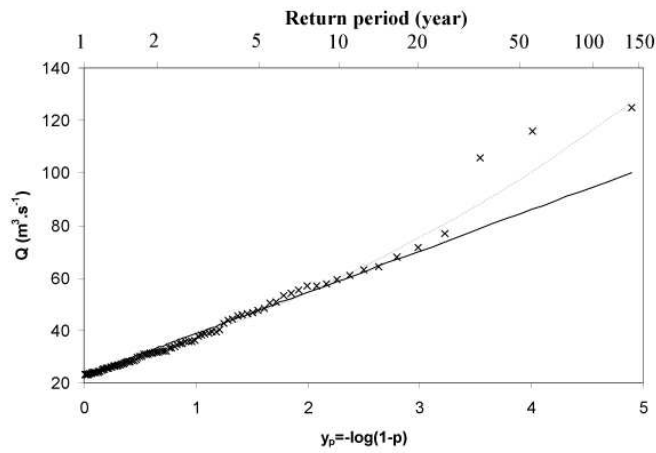


Figure 5

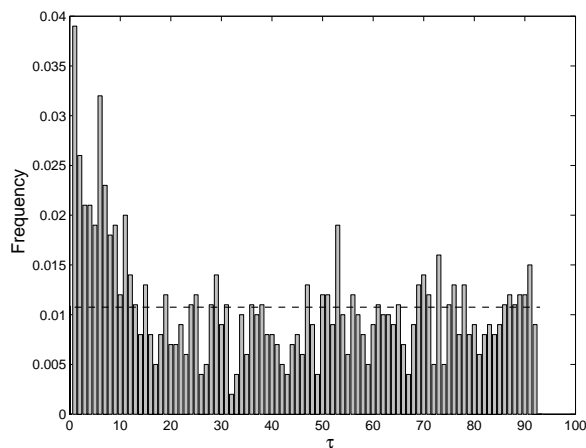


Figure 6

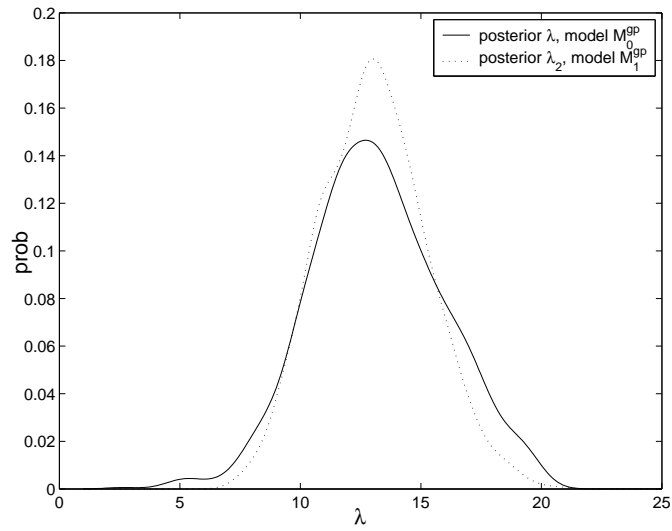


Figure 7

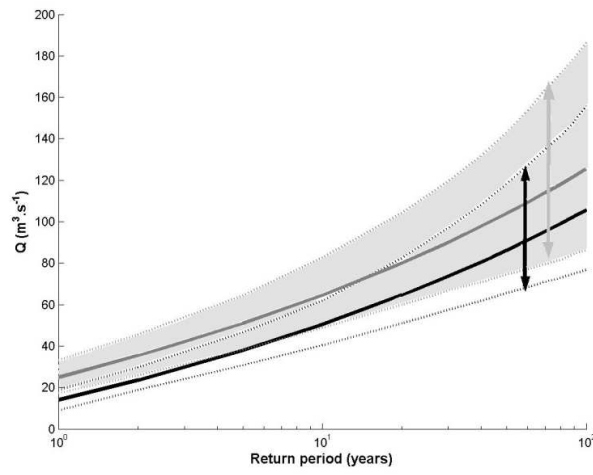


Figure 8

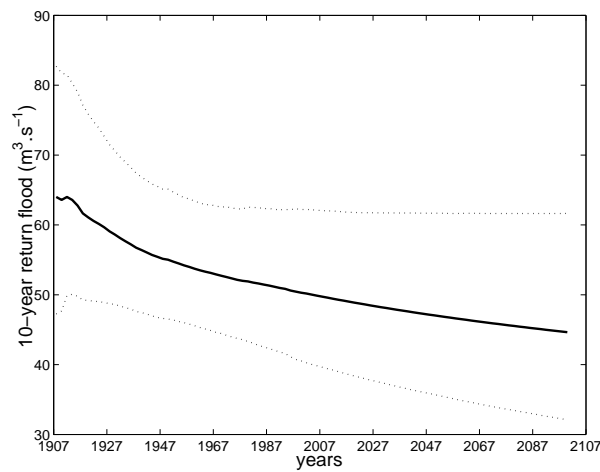


Figure 9

ARTICLE 2 : METHODES MCMC

Article soumis à *Water Resources Research*.

Accepté, sous presse.

AN APPLICATION OF BAYESIAN ANALYSIS AND MCMC METHODS TO THE ESTIMATION OF A REGIONAL TREND IN ANNUAL MAXIMA.

B. RENARD¹, V. GARRETA¹, M. LANG¹

(1) *Cemagref Centre de Lyon, U.R. Hydrologie-Hydraulique, 3 bis Quai Chauveau, CP 220, 69336 Lyon cedex 09, France.*

telephone:33 4 72 20 87 72

fax:33 4 78 47 78 75

e-mail : renard@lyon.cemagref.fr

Index terms:

0520 Data analysis: algorithms and implementation

0560 Numerical solutions

3275 Uncertainty quantification

1817 Extreme events

1860 Streamflow

Keywords : Regional analysis; Trend detection; MCMC ; Bayesian analysis ; Gibbs ; Metropolis ;

Abstract

Bayesian analysis is becoming increasingly popular in a number of fields, including hydrology. It appears to be a convenient framework for deriving complex models in agreement with both physical reality and statistical requirements. The aim of this paper is to present an application to the regional frequency analysis of extremes in a non-stationary context. A non-stationary regional model is thus proposed, together with the related hypotheses. The Bayesian inference of this model is then described. Markov chain Monte Carlo (MCMC) methods are needed for this purpose, because of the dimensionality of the model, and are described in this paper. The usefulness of such a model is then illustrated on a hydrological case study concerning annual maximum discharges of several sites. The advantage of regional analysis compared to at-site estimation is thus highlighted. Moreover, the Bayesian framework allows for a direct and comprehensive inference based on the posterior distribution, and is able to take into account modeling uncertainties, which is particularly useful when the stationarity of a series can neither be ensured nor be totally rejected.

I. Introduction

Bayesian analysis has become increasingly popular in recent years, and has shown its practical benefits in a number of applied problems. As an example, publications dealing with Bayesian inference are now usual in the field of hydrology (Coles and Tawn, 1996; Khodja et al., 1998; Lu and Berliner, 1999; Perreault et al., 1999, 2000a, b, c; Thyer and Kuczera, 2000, 2003a, b; Wang, 2001; Thyer et al., 2002; Coles and Pericchi, 2003; Coles et al., 2003; Parent and Bernier, 2003; Perreault and Fortin, 2003; Tapsoba et al., 2004; Moyeed and Clarke, 2005; Reis and Stedinger, 2005; Renard et al., 2006). This popularity can be primarily explained by the rediscovery of the usefulness of Markov chain Monte Carlo (MCMC) algorithms in the early 1990s (Gelfand and Smith, 1990). The aim of this article will be to describe an application of the Bayesian analysis and the MCMC methods to the estimation of a regional trend in the distribution of annual maximum discharges of several sites.

The basis of the Bayesian paradigm is to use the posterior distribution of parameters for inference. Suppose \mathbf{X} is a vector (or a matrix) of observed data, whose probability distribution is parameterized by $\boldsymbol{\theta}$. The parameters are here considered as random quantities. Let $\pi(\boldsymbol{\theta})$ denote the prior distribution of $\boldsymbol{\theta}$, which reflects the knowledge about the parameters before observing the data, and $p(\mathbf{X}|\boldsymbol{\theta})$ be the likelihood of the data. The posterior distribution is then defined by the Bayes relationship:

$$p(\boldsymbol{\theta}|\mathbf{X}) = \frac{p(\mathbf{X}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int p(\mathbf{X}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}} \quad (1)$$

Except for a small number of dimensions ($d \leq 3$), the posterior distribution is not easily tractable. First of all, a high dimensional distribution cannot directly be used to simply summarize the parameters' properties. Moreover, the denominator of equation (1), which is called the marginal likelihood of observations, can generally not be computed, which means that the posterior distribution is only known up to a constant of proportionality. In order to overcome these difficulties, one way is to use a multivariate sample from the posterior distribution. Marginal distributions will then be easily estimated using marginal samples, leading to classical point, interval or distribution estimates.

MCMC methods are tools for simulating data from an arbitrary distribution. The principle is to create a Markov chain ($\boldsymbol{\theta}^{(t)}$) which converges to a stationary distribution that is the desired distribution. Some general schemes exist to achieve this simulation, namely the Gibbs (Geman and Geman, 1984) and the Metropolis (Metropolis and Ulam, 1949; Metropolis et al., 1953) samplers. A more general version of the latter is known as the Metropolis-Hastings algorithm (Hastings, 1970). The interest of MCMC methods lies in the fact that the preceding samplers are usable for sampling from a very wide range of distributions. Despite this quasi-universality, some problems can arise during the algorithm implementation, in particular numerical overflows, slowness of the runs, or difficulties in convergence assessment. The books by Gelman et al. (1995) and Tanner (1996) provide some guidelines to improve the algorithms' performances. Recently, El Adlouni et al. (2006) compared different methods to check the convergence of the chains.

An interesting application of the Bayesian framework lies in the estimation of changes in hydrological series. For at-site Gaussian data, such an application has been developed by Lee and Heghinian (1977) and Bernier (1994). Perreault et al. (2000a; b) generalized this approach by dealing with both change detection and frequency analysis in a non-stationary context. Renard et al. (2006) adapted the latter approach in the case of extreme data. In some situations, it can be anticipated that the cause of change has an impact at a wide spatial scale

(e.g., climate-related change). In such cases, it would be relevant to search for consistent changes within a sample of several sites in a homogeneous region. Perreault et al. (2000c) therefore proposed a model for estimating a simultaneous single change in the mean of several Gaussian series. Unfortunately, the adaptation of this approach to the extreme values context is far from obvious, both for practical and theoretical reasons.

The aim of this article is to describe a Bayesian regional model for annual maximum discharges in a non-stationary context. It is organized as follows. Firstly, the Bayesian framework is described (section 2). Models used at local and regional scales are presented, together with the hypotheses we made. The use of the posterior distribution for parameter estimation, change detection and frequency analysis is also explained. In section 3, the strategy we adopted to sample from the posterior distribution is described, based on the combined use of two MCMC algorithms. A hydrological case study involving six gauging stations in France is finally presented in section 4, before drawing some conclusions and discussing perspectives (section 5).

All the algorithms used in this study are available at the following address:

<http://www.lyon.cemagref.fr/hh/PNRH-NS/>.

II. Bayesian inference

II.1. Non-stationary regional model

Let $X_t^{(i)}$ denote the annual maximum discharge of a station i ($i=1, \dots, p$) for a year t . Extreme value theory provides asymptotic arguments to the use of the generalized extreme value (GEV) distribution (Fisher and Tippett, 1928; Jenkinson, 1955), whose probability density function (pdf) can be written as follows:

$$GEV(x; \alpha, \beta, \xi) = \frac{1}{\alpha} \left[1 - \xi \left(\frac{x - \beta}{\alpha} \right) \right]^{\frac{1}{\xi} - 1} \exp \left\{ - \left[1 - \xi \left(\frac{x - \beta}{\alpha} \right) \right]^{\frac{1}{\xi}} \right\}, \quad (2)$$

$$\text{with } \alpha > 0 \text{ and } 1 - \xi \left(\frac{x - \beta}{\alpha} \right) > 0.$$

α , β and ξ are the scale, position and shape parameters. In the stationary case, the annual maxima series can thus be modeled with the GEV distribution,

$X_t^{(i)} \sim GEV(\alpha_i, \beta_i, \xi_i)$ ($M_0^{(i)}$ model). If the stationarity of the series can not be ensured, a pragmatic approach suggested by Coles (2001) consists in assuming that for a given year, the annual maximum still arises from a GEV distribution, but that the parameters of this distribution will be a function of time. In this paper, it will be assumed that the location parameter is a function of time, whereas the scale and shape parameters remain constant. This leads to the $M_1^{(i)}$ model, $X_t^{(i)} \sim GEV(\alpha_i, \beta_i (1 + \delta_i t), \xi_i)$. Alternative models may be used, for instance polynomial trends, step changes, trends on the scale or the shape parameter, etc. The methodology presented in this paper can be used for the inference of such models. However, because extreme discharges series are known to be affected by a strong natural variability together with large measurement uncertainties, we favored a parsimonious description of the temporal changes in the annual maxima distribution.

The approach applied at the regional scale consists in deriving a regional model comprising local parameters, which are related to a given site, and regional parameters, which are assumed to be equal for all sites within the region, as suggested by Katz et al. (2002). In the stationary case, it was therefore assumed that the shape parameter is identical for all sites,

which leads to the M_0 regional model $X_t^{(i)} \sim GEV(\alpha_i, \beta_i, \xi)$. A non-stationary model was derived by applying a regional trend to the location parameter, which leads to the M_I regional model $X_t^{(i)} \sim GEV(\alpha_i, \beta_i (1 + \delta t), \xi)$. It should be noticed that the assumptions we made about the data distribution, notably the hypothesis of identical shape and trend parameters for all sites within the region, may be irrelevant in heterogeneous regions. Methods for creating homogeneous regions have been extensively studied (e.g. Hosking and Wallis, 1997, or Ouarda et al., 1999), and will not be described in this paper. However, the adequacy between the data and the model will have to be properly checked.

The estimation of the parameters of the M_I model is the main objective of this paper. For a given year, data thus consist of a vector of annual maxima. Consequently, the multivariate distribution of this vector will be needed for estimation. Multivariate extreme value distributions constitute an area of active research for statisticians, but remain difficult to handle in practice, especially with more than two or three dimensions. Although some promising approaches have been proposed (Coles and Tawn, 1990; Bortot et al., 2000; Favre et al., 2004), this aspect of the problem is beyond the scope of this paper. We thus make the hypothesis of independence between stations. As noted by Katz et al. (2002), this should have little effect on point estimates, but could result in an underestimation of the parameter variances. The effect of ignoring spatial dependence will be discussed in more detail in a later section.

II.2. Bayesian inference

The first step in performing the Bayesian inference of the M_I model is to compute the likelihood. Let $\mathbf{X}^{(i)} = (x_t^{(i)})_{t=1, \dots, n}$ be the column vector of annual maxima for a particular site i , and $\mathbf{X} = (\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p)})$ be the data matrix. Let $\boldsymbol{\theta} = (\alpha_1, \beta_1, \dots, \alpha_p, \beta_p, \xi, \delta)$ denote the parameters vector of size $2p+2$ arising from the M_I model with p sites. Under the hypothesis of spatial independence between sites, the likelihood is computed as follows:

$$\begin{aligned} p(\mathbf{X} | \boldsymbol{\theta}) &= p(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p)} | (\alpha_1, \beta_1, \dots, \alpha_p, \beta_p, \xi, \delta)) \\ &= \prod_{t=1}^n p(x_t^{(1)}, \dots, x_t^{(p)} | (\alpha_1, \beta_1, \dots, \alpha_p, \beta_p, \xi, \delta)) \\ &= \prod_{t=1}^n \prod_{i=1}^p GEV(x_t^{(i)}; (\alpha_i, \beta_i (1 + \delta t), \xi)) \end{aligned} \quad (3)$$

The next step is to specify a prior distribution $\pi(\boldsymbol{\theta})$ for the parameters, without using the data. Several approaches can be considered for this purpose: prior knowledge of an expert (Coles and Tawn, 1996), additional data (Perreault et al., 2000c), empirical relationships with explanatory variables (Renard et al., 2006), regional information (Ribatet et al., 2006), etc. The lack of information about the phenomenon studied sometimes leads to the use of improper priors. However, such distributions can be problematic in the context of Bayesian model choice which will be described later. Consequently, it will be supposed in this paper that prior information is available and can be translated into a proper prior distribution. Given the likelihood of the data and the prior distribution of the parameters, the posterior distribution of the parameters can be computed from the Bayes' theorem (equation 1). Because of the limitations explained in the introduction, a sample simulated from this distribution using a MCMC algorithm is needed in practice. This sample is used to estimate the marginal posterior distribution of each parameter and related characteristics (mean, variance, probability intervals, etc.).

An attractive characteristic of Bayesian analysis is the possibility of computing model probabilities. Suppose that $\Lambda = \{M_1, \dots, M_q\}$ is a collection of models possibly suitable for describing the data. Assume that $\pi(\cdot)$ is a prior probability distribution on Λ , which means that a prior probability can be assigned to each model without using the data. The posterior probability of a model M_j can then be computed with the Bayes theorem:

$$p(M_j | \mathbf{X}) = \frac{p(\mathbf{X} | M_j)\pi(M_j)}{\sum_{i=1}^q p(\mathbf{X} | M_i)\pi(M_i)} \quad (4)$$

This equation requires the computation of the marginal likelihood of observations for each model, which is equal to the normalizing factor of equation (1):

$$p(\mathbf{X} | M) = \int p(\mathbf{X}, \boldsymbol{\theta} | M) d\boldsymbol{\theta} = \int p(\mathbf{X} | \boldsymbol{\theta}, M)\pi(\boldsymbol{\theta} | M) d\boldsymbol{\theta} \quad (5)$$

As stated previously, this integration is generally intractable with a large number of dimensions. Alternative methods therefore have to be used. The approach proposed by Chib (1995) was used in this paper, and is described in Appendix 1.

The posterior probabilities of the models considered can be helpful in a decision-making process. Bayes factors (Kass and Raftery, 1995) can be used for this purpose, as they provide a measurement of the pertinence of one model compared to another. The Bayes factor between model M_i and model M_j is defined as follows:

$$B_{i,j} = \left(\frac{p(M_i | \mathbf{X})}{p(M_j | \mathbf{X})} \right) / \left(\frac{\pi(M_i)}{\pi(M_j)} \right) = \frac{p(\mathbf{X} | M_i)}{p(\mathbf{X} | M_j)} \quad (6)$$

This quantity is then compared to 1, with stronger confidence in model M_i for high values. Kass and Raftery (1995) furnished guidelines for interpreting Bayes factors, together with additional developments. For instance, composite Bayes factors can be used to compare two sets of models (e.g. several stationary models and several trend models).

Hydrologists involved in flood risk assessment are interested in p -quantiles of the annual maxima distribution, that is the value which is exceeded with a probability of $1-p$. For a GEV distribution, the quantile q_p can be computed as:

$$q_p = \beta + \frac{\alpha}{\xi} \left(1 - (-\log(p))^{\xi} \right) \quad (7)$$

Given a model M_k , the posterior distribution of the quantile can thus easily be obtained from the posterior distribution of the parameters. Of course, in the case of a non-stationary model, the quantile posterior distribution will also depend on time. For instance, in the case of the M_l model, the p -quantile for a given site i at year t will be:

$$q_p(t) = \beta_i (1 + \delta t) + \frac{\alpha_i}{\xi} \left(1 - (-\log(p))^{\xi} \right) \quad (8)$$

An alternative to the selection of a model using Bayes factors or any other criterion is to average over all the models considered. This can be done using the following equation:

$$p(q_p(t) | \mathbf{X}) = \sum_{k=1}^q p(q_p(t) | \mathbf{X}, M_k) p(M_k | \mathbf{X}) \quad (9)$$

This approach is known as Bayesian model averaging. It allows the uncertainty related to the choice of a model to be taken into account. An underlying assumption is that the data arise from one of the models considered. Of course, as a model is a simplification of reality, this hypothesis is unlikely to be strictly fulfilled. However, by considering contrasting models, a more realistic quantification of the uncertainties can be obtained, compared to the approach

consisting in choosing a single model on the basis of a given criterion. The paper by Hoeting et al. (1999) provides a comprehensive review of Bayesian model averaging.

III. MCMC methods

The principle of MCMC methods is to create a random walk in the parameters space which converges to a stationary distribution that is the joint posterior distribution (Gelman et al., 1995). Several algorithms have been proposed for this purpose, and have been described and compared by Tierney (1994a; b). Whatever the method used, an important feature is to check that the algorithm has been run for long enough to ensure the simulated sample to be representative of the target distribution. Consequently, the first iterations are usually not used for inference, and are considered as burning time. Moreover, the simulated values are not independent: a large number of iterations may thus be needed to obtain an acceptable accuracy.

In the following section, the MCMC strategy used in this paper is exposed. $p(\boldsymbol{\theta} | \mathbf{X})$ denotes the target posterior distribution of a d -dimensional parameter vector $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d)$.

III.1. Use of a combined algorithm

Metropolis sampler

The Metropolis algorithm (Metropolis and Ulam, 1949; Metropolis et al., 1953) has been widely used in Bayesian applications, because of its simplicity and its efficiency. Its principle can be summarized as follows: starting from a vector generated at iteration $i-1$, a new candidate vector is generated thanks to a symmetric jump distribution. If this new vector leads to an increasing value of the target distribution, it is accepted as the generated value at iteration i . Otherwise, the ratio between the new and the previous value of the target distribution is computed, and used as the acceptance probability of the candidate vector. In case of rejection, the generated vector at iteration i remains equal to that of iteration $i-1$. The Metropolis algorithm was used in this paper with a Gaussian jump distribution with variance matrix $\boldsymbol{\Sigma}$. Appendix 2 describes the way the algorithm was implemented. To start a Metropolis simulation, a starting point and a jump variance are therefore needed. The choice of a suitable starting point can be made in different ways: as an example, if informative priors are used, the prior mode can be chosen. Other methods for deriving a rough estimate of the posterior mode are available, and are described by Gelman et al. (1995). The choice of a jump variance is more delicate and can be problematic in practice, whereas it is the most important task to derive an efficient algorithm. More accurately, if the jump variance is too high, most of the candidate vectors will be generated far from the high-density area of the posterior distribution, thus leading to rejection of the candidate vector. The algorithm will then generate a great number of tied values (*i.e.* successive identical values), and the number of iterations necessary to visit the whole parameters space will be unacceptable. Alternatively, with a too small jump variance, almost all candidate vectors will be accepted, because each new vector will be generated very close to the precedent, leading to a ratio of target distributions close to one. Once again, the parameters space will be visited too slowly.

Several solutions can be considered for deriving an adequate variance matrix. As an example, it may be possible to use the observed Fisher information matrix as a first approximation. This may be relevant for large samples, because of the predominant influence of the likelihood compared to the prior distribution. Alternatively, if a proper prior distribution is specified, it can also be used to approximate the variance matrix. The most efficient solution may be to use a Taylor series expansion of the posterior density. However, an analytical computation of

the derivatives will often be intractable, and numerical approximates will have to be computed.

Gibbs/Metropolis sampler

In order to overcome these difficulties, a possibility is to modify the characteristics of the algorithm during the iterations, by using the previously simulated values. For instance, it may be possible to periodically update the variance matrix of the jump distribution. This leads to the construction of adaptive algorithms. Gelman et al. (1995) provide some guidelines for this purpose. Some examples of adaptive Metropolis samplers can be found in the papers by Haario et al. (2001), Thyer et al. (2002) or Marshall et al. (2004).

The adaptive MCMC algorithm used in this paper is based on the Gibbs sampler (Geman and Geman, 1984). This sampling scheme is based on the principle of alternating conditional sampling: within an iteration i , each component of the parameters vector is updated successively, by sampling in the full conditional densities. More accurately, at iteration i , the component q is updated by sampling from the distribution

$p\left(\theta_q \mid \theta_1^{(i)}, \dots, \theta_{q-1}^{(i)}, \theta_{q+1}^{(i-1)}, \dots, \theta_d^{(i-1)}, \mathbf{X}\right)$. Unfortunately, this full conditional density does

generally not have the form of a standard distribution. It is thus necessary to add a new simulation step for each of the d conditional samplings. Several authors have proposed tools for this purpose. As an example, the Griddy-Gibbs sampling algorithm introduced by Ritter and Tanner (1992) is based on the inversion of the cumulative distribution function (cdf) of the full conditional distribution, this cdf being estimated on a discrete grid of points.

Alternatively, Gilks and Wild (1992) proposed using an acceptance-rejection (AR) method to sample from log-concave full conditional distributions, or the joint use of AR and Metropolis-Hastings algorithms if the log-concavity cannot be ensured (Gilks et al., 1995).

The approach applied in this paper is based on the suggestion of Gelman et al. (1996) and Coles and Tawn (1996) to use a Metropolis algorithm on the one-dimensional distribution

$p\left(\theta_q \mid \theta_1^{(i)}, \dots, \theta_{q-1}^{(i)}, \theta_{q+1}^{(i-1)}, \dots, \theta_d^{(i-1)}, \mathbf{X}\right)$. If a univariate Gaussian distribution is used as the

jump distribution, a jump variance is needed at each iteration of the Gibbs algorithm, and for each of the d conditional samplings. Choosing a constant jump variance for each conditional sampling may lead to a poor performance of the algorithm. Moreover, given the large number of Gibbs loops, it is not possible to manually check the convergence of the one-dimensional Metropolis runs. An automated procedure is therefore needed to compute an adequate jump variance for each conditional distribution.

To this end, it is possible to use the assumption that the distribution to sample from is roughly normal. This is not a drastic restriction in the Bayesian context, because the posterior distribution is asymptotically Gaussian under some regularity conditions. It is then possible to use the following property of the multivariate Gaussian distribution: the full conditional distribution is still Gaussian, and its variance does not depend on the values used for conditioning. This implies that, for each parameter, the variance can be transmitted between two iterations of the Gibbs algorithm. Moreover, this variance can be updated after each Metropolis run, in order to comply with an adequate acceptance rate. Gelman et al. (1995) recommended keeping the acceptance rate between 0.23 with a large number of dimensions ($d > 5$) and 0.44 ($d = 1$). Consequently, at the end of each one-dimensional Metropolis run, the acceptance rate τ is computed on the generated values. If $\tau > 0.44$, the jump variance is increased in order to decrease the acceptance rate. Similarly, if $\tau < 0.23$, the jump variance is decreased. This leads to the implementation of an adaptive Gibbs/Metropolis sampler, which is described in Appendix 3.

Compared to the Metropolis sampler, this adaptive algorithm has a larger complexity. In order to simulate a sample of length N , the Metropolis sampler will thus execute N loops including computations of the target distribution, while this number will be $dN_{Metro}N$ for the

Gibbs/Metropolis sampler, where N_{Metro} is the length of the one-dimensional Metropolis runs. However, it is anticipated that the adaptation rules will improve the efficiency of the algorithm, by decreasing the sensitivity to starting point or jump variance misspecification.

Combining samplers

Although very useful from a practical point of view, adaptive algorithms suffer from an important theoretical drawback: because the Markov chain properties of the simulated values are changed during the iterations, the stationarity of the target distribution is no longer ensured and has to be proven. Examples where an adaptive algorithm fails to generate values from the desired distribution have been provided by Tierney and Mira (1999) or Haario et al. (2001). At the opposite, the latter authors succeed in establishing the correct ergodic properties of their AM adaptive sampler. The adaptive Gibbs/Metropolis used in this paper was empirically tested on a number of cases, and gave so far acceptable results. However, we are not able to prove its ergodic properties.

Based on these considerations, should the use of adaptive algorithms be avoided? Actually, a number of such algorithms have been developed (Gilks and Wild, 1992; Roberts and Gilks, 1994; Gilks et al., 1995; Tierney and Mira, 1999; Haario et al., 2001, 2005) and have given acceptable results in most applications. As an illustration, Haario et al. (2001) noted concerning their adaptive AP sampler that “for many practical applications, the error produced by the AP algorithm is, however, ignorable”, although a counter-example can be constructed. Nevertheless, in order to avoid errors due to particular target distributions, the following two-step procedure was adopted: the adaptive Gibbs/Metropolis algorithm is first used for a limited number of iterations. The simulated values are then used to compute a mean vector $\boldsymbol{\mu}$ and a variance matrix $\boldsymbol{\Sigma}$, but will not be used for inference. As a second step, a traditional Metropolis algorithm without adaptation, whose starting point is equal to $\boldsymbol{\mu}$, is used in order to simulate the sample which will be used for inference. The jump distribution is a Gaussian distribution with variance matrix $c^2\boldsymbol{\Sigma}$, with $c = 2.4/\sqrt{d}$, as suggested by Gelman et al (1995). Because no adaptation is made on this latter sampler, the theoretical properties of the simulated Markov chain are ensured.

III.2. Influence of starting characteristics on algorithm performance

The aim of this section is to highlight some properties of the Metropolis, the adaptive Gibbs/Metropolis and the combined samplers previously described, based on simplified target distributions. This comparison does not aim at deriving general conclusions about the sampler performances, but just at illustrating some issues which can be encountered in practice.

Samples were generated from the three following target distributions:

- ✓ $f_1(x_1, \dots, x_5) = f_1(\mathbf{x}) = N(\mathbf{x} | \boldsymbol{\theta}; \boldsymbol{\Sigma})$, where the variance matrix has the form $\boldsymbol{\Sigma}(i, j) = 0.8^{|i-j|}$, for $i, j = 1, \dots, 5$.
- ✓ $f_2(x_1, \dots, x_5) = \prod_{i=1}^5 [1 + 0.2x_i]^{-6} \exp\{-[1 + 0.2x_i]^{-5}\}$, *i.e.* f_2 is a multivariate distribution with independent GEV marginal distributions, whose location parameters are 0, scale parameters are 1 and shape parameters are -0.2 .
- ✓ $f_3(x_1, \dots, x_5) = f_3(\mathbf{x}) = 0.5N(\mathbf{x} | \boldsymbol{\theta}; \mathbf{I}) + 0.5N(\mathbf{x} | \boldsymbol{\delta}; \mathbf{I})$, *i.e.* f_3 is a mix of two multivariate normal distributions with independent marginal distributions and mean vectors equal to $\boldsymbol{\theta}=(0;0;0;0;0)$ and $\boldsymbol{\delta}=(8;8;8;8;8)$.

Inadequate starting characteristics were specified in order to evaluate the samplers' robustness. Starting vectors are equal to $-\mathbf{10}$ for the f_1 and f_3 distributions, and -4 for the f_2

distribution (in order to ensure that $1 - \xi\left(\frac{x - \beta}{\alpha}\right) > 0$). The variance matrix Ω of the

Gaussian jump distribution used in the Metropolis algorithm is equal to:

$$\Omega = \begin{pmatrix} 0.01 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 0 & 100 \end{pmatrix} \quad (10)$$

In this way, some components will have an underestimated variance, and some others an overestimated one. Similarly, the starting variances used in the Gibbs/Metropolis sampler are equal to (0.01;0.1;1;10;100).

The Metropolis algorithm was used to simulate 50 000 values from the target distribution. The first 20 000 iterations were considered as burning time. The remaining 30 000 values were then used for inference. The Gibbs/Metropolis algorithm was used with $N_{Metro}=100$, and a sample of length 5000 was simulated, with 2000 iterations of burning time. The combination of these two samplers was made as follows: the Gibbs/Metropolis sampler was used during 1000 iterations, with the starting points and the jump variance previously described. The mean vector and the variance matrix were then computed on the last 500 values, and used to derive the starting point and the variance matrix of the jump distribution of a Metropolis algorithm. 50 000 values were thus simulated, and the last 30 000 were used for inference.

The results presented in Figure 1 are related to the first components of the target distributions, whose jump variances were strongly underestimated. In the Metropolis case, it can be observed that the simulated values are a very poor representation of the expected distributions. In fact, both the burning time and the number of iterations are too low to provide an adequate description of the target distribution, because of the small size of the jumps. Thereby, for the f_1 distribution, the chain is still influenced by the too low starting value. In the f_2 case, the algorithm seems unable to sample a value from the tail of the distribution. In the f_3 case, the algorithm is locked up in the first mode, which can be explained by the fact that the jump variance is too low to be able to reach the other mode with a significant probability. In contrast, the samples arising from the adaptive Gibbs/Metropolis algorithm provide an adequate representation of the three target distributions. Consequently, the combined algorithm appears useful for overcoming the issues observed with the Metropolis sampler for the f_1 and the f_2 distributions. The f_3 case is more mitigated: although the combined algorithm is able to sample in the two modes of the target distributions, the accuracy of the estimates is not satisfactory. This accuracy is however likely to be enhanced by increasing the number of iterations of the Metropolis part of the combined algorithm. Similarly, results for the last components of each target distribution, whose jump variances were strongly overestimated, are presented in Figure 2. In the Metropolis case, the influence of the starting value is still observable for the f_1 target distribution. The fact that the marginal distributions are strongly positively correlated can explain this phenomenon: while the first components remain in the left tail of the distribution because of a too low jump variance, the other components are unlikely to reach the center or the right tail of the distribution. Conversely, in the f_2 and the f_3 cases, the values simulated with the Metropolis algorithm provide an adequate representation of the target distributions. Once again, the Gibbs/Metropolis sampler behaves satisfactorily, whatever the target distribution considered, while the accuracy of the values simulated by the combined sampler is adequate in the f_1 and the f_2 cases, but is not satisfactory with the f_3 distribution.

These results show that the combined sampler is able to overcome some issues observed with a simple Metropolis algorithm, by deriving acceptable starting characteristics. However, because the inference will still be conducted using a Metropolis sample, the limitations related to the use of a Gaussian jump distribution may remain. As an illustration, the shape of the strongly multimodal distribution f_3 is too different from that of the jump distribution, which prevents the algorithm from behaving satisfactorily. A more subtle implementation, including a non-Gaussian instrumental distribution, is thus necessary to avoid these problems. Consequently, the Metropolis-Hastings algorithm is likely to be a more suitable tool in this case, because of the possibility of using an asymmetrical jump distribution. Moreover, this example also emphasizes the importance of monitoring convergence of the chain, preferably by several methods, as suggested by El Adlouni et al. (2006). In a real-life problem, as the target distribution would be unknown, a convergence assessment based on a single chain is likely to arrive at a false convergence in the case of a sampler locked up in one of the two modes. Conversely, by using several independent chains, obtained with overdispersed starting points (Gelman et al., 1995), the existence of several modes is more likely to be detected.

IV. Case study

IV.1. Data and prior elicitation

Six undisturbed hydrological stations located in the center of France were used (Figure 3). A description of the general characteristics of the watersheds is provided in Table 1. Daily discharges are available for various periods. Only years for which all series are free from missing values were considered, which leads to 34 years of data between 1961 and 2002, with 1981, 1982, 1983, 1988, 1992, 1996, 1997 and 2001 as missing years. For each station, the annual maximum value is extracted from the daily discharges series.

The M_I model detailed in section II.1 is used to describe these data, with a minor modification to account for the beginning of the study period: $X_t^{(i)} \sim GEV(\alpha_i, \beta_i (1 + \delta(t - 1960)), \xi)$. A prior distribution is specified as follows. Firstly, empirical knowledge is used for the regional parameters. The shape parameter is assumed to follow a Gaussian distribution with zero mean and standard deviation 0.3, which implies that the interval $[-0.6; 0.6]$ encompasses more than 95% of the density. This prior distribution can be compared with the Martins and Stedinger (2000) geophysical prior, which is less variable and is entirely included in the interval $[-0.5; 0.5]$. We also assume that the regional trend parameter follows a Gaussian distribution. Because no evidence of climate-related changes has been reported for high flows in France, the mean is set to zero. Moreover, we consider that it is very unlikely that a change of more than 100% could have affected the location parameter $\beta_{j,t} = \beta_j (1 + \delta(t - 1960))$ between 1960 and 2000. More accurately, the prior variance is computed in order to meet the following specification:

$$\begin{aligned} \Pr\left(\left|\frac{\beta_{j,2000} - \beta_{j,1960}}{\beta_{j,1960}}\right| \leq 100\%\right) &= 0.95 \\ \Leftrightarrow \Pr\left(\left|\frac{\beta_j (1 + \delta(2000 - 1960)) - \beta_j (1 + \delta(1960 - 1960))}{\beta_j (1 + \delta(1960 - 1960))}\right| \leq 100\%\right) &= 0.95 \quad (11) \\ \Leftrightarrow \Pr(|40\delta| \leq 100\%) &= 0.95 \\ \Leftrightarrow \Pr(|\delta| \leq 1/40) &= 0.95 \end{aligned}$$

This leads to the prior standard deviation for δ being set to $1/80$, so that the change in 40 years, expressed as a percentage, lies in the interval $[-100\%;100\%]$ with a probability larger than 0.95.

Empirical knowledge for the local location and scale parameters is harder to derive, because these parameters are more closely related to the specific characteristics of the watershed (area, topography, geology, etc.). Consequently, the data derived from non-shared years are used for this purpose. Maximum likelihood (ML) estimates of the local stationary GEV distributions are computed. Because of the limited sample size available for some sites, the use of the Fisher information matrix to derive the distribution of the estimates may lead to biased results. Consequently, bootstrap samples of length 10 000 were generated for each site, and a Gamma distribution was fitted to the sample of ML estimates obtained:

$$Ga(x; a, b) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx), \quad x > 0 \quad (12)$$

Table 2 gives details of the values obtained, and the number of data available for prior elicitation.

Finally, the prior independence of all parameters is assumed, which leads to the following prior distribution:

$$\begin{aligned} \pi(\boldsymbol{\theta}) &= \pi(\alpha_1, \beta_1, \dots, \alpha_6, \beta_6, \xi, \delta) \\ &= N(\xi; 0, 0.3^2) N(\delta; 0, (1/80)^2) \prod_{j=1}^6 Ga(\alpha_j; a_j^{(\alpha)}, b_j^{(\alpha)}) Ga(\beta_j; a_j^{(\beta)}, b_j^{(\beta)}) \end{aligned} \quad (13)$$

MCMC simulations were performed on the non-normalized posterior distribution $p(\mathbf{X} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta})$ using $N_{Gibbs}=1000$ iterations of the Gibbs/Metropolis sampler ($N_{Metro}=100$) followed by $N_{iter}=100\,000$ iterations of the Metropolis one, as explained in section III. For each parameter, the starting point was set to the prior mean and the starting variance was set to the prior variance, which allows the first 1000 Gibbs/Metropolis iterations to be performed. The Metropolis sampler was then used with a starting point and a starting variance matrix derived from the last 500 Gibbs/Metropolis iterations.

IV.2. Assessment of MCMC runs convergence

As a first step, convergence of the MCMC simulations toward the posterior distribution has to be assessed. Unlike section III.2, the target distribution is here unknown, thus preventing any conclusion to be made concerning convergence by simply examining a single output of the MCMC runs. Gelman et al. (1995) suggested using parallel simulations performed with different starting points. Ten Gibbs/Metropolis simulations of length 1000 were thus used with overdispersed starting points. Figure 4 shows the random walks of parallel sequences in the planes formed: (a) by scale and position parameters of the first station and (b) by shape and trend regional parameters. It can be observed that simulated points join the same area after a few iterations, which indicates that mean and variance characteristics computed from the last 500 Gibbs/Metropolis iterations will give rather good starting characteristics for the subsequent Metropolis iterations.

In order to quantify the minimum number of iterations needed to reach convergence, the approach suggested by Gelman et al. (1995) was applied. It is based on the computation of the $\sqrt{\hat{R}}$ statistic, where \hat{R} estimates the ratio of total variance by within-sequences variance. If convergence is reached, then the between-sequences variance should be negligible, thus leading to a ratio close to one. For a simulation of length n , this ratio is estimated thanks to the second half of simulated values. Figure 5a shows the results for the 4 preceding parameters. It confirms that the Gibbs/Metropolis algorithm is very fast in reaching

convergence, as the $\sqrt{\hat{R}}$ statistic declines to an acceptable value with only 50 iterations (values below 1.2 are usually considered as acceptable). Because only the last 500 generated values will be used for computing the mean and the variance matrix, it can be considered that these estimates will be adequate as starting characteristics of the subsequent Metropolis iterations.

Once this first step has been achieved by Gibbs/Metropolis sampling, additional Metropolis runs have to be performed. A high number of values is required here, firstly because of the tied values generated, secondly because this sample will be used to perform the inference on the posterior distribution. Once again, the approach of Gelman et al. (1995) was applied, with ten overdispersed starting points, and a starting variance matrix computed on the last 500

Gibbs/Metropolis iterations. Figure 5b shows the value of the $\sqrt{\hat{R}}$ statistic for different run lengths. After 9000 iterations, it can be considered that convergence is reached. As the total number of iterations is equal to 100 000, this will give a sufficient number of values to perform inference on quantities of interest. More accurately, the following computations will be performed with the last 50 000 generated values.

IV.3. Regional trend estimation

Figure 6 shows the estimated posterior marginal distributions of two local parameters, namely the scale and the location parameters of the first station. Vertical lines represent locally-estimated parameters using the maximum likelihood method on the shared 34 years, and the thin lines represent the prior distributions. It can be observed that the results are consistent for the two parameters, although the Bayesian estimate of α_l is larger than the corresponding ML-estimate, because of the influence of the prior distribution. Moreover, the posterior distribution of α_l shows a slight departure from normality, which can easily be taken into account with the Bayesian framework, because no Gaussian asymptotic assumption is made. Figure 7 deals with the two regional parameters. Local and regional estimates are consistent for the trend parameter δ (right panel), which indicates that the assumption of a shared regional trend is plausible. This is also illustrated by Figure 8, which shows for each of the six stations the 34 years of data used for inference, together with the estimated regional trend. Moreover, the variance of the regional parameter is small compared to the range of locally-estimated values. This indicates that such a regional model is likely to reduce the estimation uncertainty. Focusing on the shape parameter estimate ξ , the assumption of an identical shape parameter over the region may appear doubtful (Figure 7, left panel). However, only the point estimates are shown, and the uncertainty is known to be very large for this parameter, especially with only 34 data points.

In order to derive a more accurate comparison between local and regional estimates, the posterior distribution in the regional model is compared to the posterior distributions of the local models $M_1^{(i)}$. The prior distributions are specified as explained in section IV.1, which means that for a particular site i :

$$\begin{aligned} \pi(\boldsymbol{\theta}) &= \pi(\alpha_i, \beta_i, \xi_i, \delta_i) \\ &= N(\xi_i; 0, 0.3^2) N(\delta_i; 0, (1/80)^2) Ga(\alpha_i; a_i^{(\alpha)}, b_i^{(\alpha)}) Ga(\beta_i; a_i^{(\beta)}, b_i^{(\beta)}) \end{aligned} \quad (14)$$

For each site, MCMC simulations are performed on the non-normalized posterior distribution $p(\mathbf{X} | \boldsymbol{\theta})\pi(\boldsymbol{\theta})$, as explained in section IV.1: the starting point and the starting variances of the Gibbs/Metropolis sampler are derived from equation (14) as the prior means and variances of each parameter. After 1000 iterations, the empirical mean and the empirical variance matrix are computed from the last 500 generated values, and are used to derive the starting characteristics of the Metropolis sampler, which will be run for 100 000 iterations. Figure 9

shows the box plots of the local trend and shape parameters, compared with the regional ones. Boxes extend from the first up to the third quartile, and the median is denoted by a line inside the box. The whiskers denote the quantiles with probabilities 0.05 and 0.95. The median values obtained for the shape parameters display a variability which is comparable to that obtained with the ML-estimates (vertical lines of Figure 7). However, the great uncertainty associated with the estimation of this parameter implies that all the 90% probability intervals are overlapping. A more complete investigation is therefore needed in order to evaluate the plausibility of the hypothesis of a regional shape parameter: this will be done in the next section IV.4. By contrast, the estimated median values of the trend parameters seem more consistent for the six stations. However, the uncertainty associated with each of the six at-site estimates remains strong, which explains the difficulty of detecting a trend in extreme data. For both parameters, the posterior variability of the regional estimates is strongly reduced compared to the local estimates: this emphasizes the benefit of a regional model, which increases the estimates precision by combining data from different locations. This benefit may need to be mitigated by the assumption of independence between stations, which was made for convenience. This issue will be addressed in the discussion section.

IV.4. Model checking

In order to assess the plausibility of the hypotheses used to construct the M_1 model (GEV distribution, regional trend and shape parameters, independence between stations), the present section compares the posterior predictive distribution to the observed distribution of the data. If the model is appropriate, a new sample of data simulated from this model should be similar to the observed data. More accurately, a sample can be replicated from the posterior predictive density, which is defined as follows:

$$p(\mathbf{y}^* | \mathbf{X}) = \int p(\mathbf{y}^* | \boldsymbol{\theta}) p(\boldsymbol{\theta} | \mathbf{X}) d\boldsymbol{\theta}, \quad (15)$$

where \mathbf{y}^* is an unknown but observable future value. By simulating a large number of samples arising from the model, the observed data can thus be located within the overall distribution of data that could have been observed (Gelman et al., 1995). In order to quantify how different the observed and the replicated data are, both having an identical sample size, the characteristics of their distributions are summarized in a test statistic $T(\mathbf{y})$. It is then possible to compute the p -value corresponding to the observed statistic $T(\mathbf{X})$, once replaced in the sample of $\left(T(\mathbf{y}^{(r)})\right)_{r=1, \dots, N_{rep}}$ replicated from the posterior predictive distribution. More

accurately, this p -value is estimated as the proportion of $\left(T(\mathbf{y}^{(r)})\right)_{r=1, \dots, N_{rep}}$ which exceeds the observed statistic $T(\mathbf{X})$. An extension of this method involves discrepancy measures, *i.e.* statistics depending both on the data and parameters, rather than classical test statistics. In the present case study, the problem is multivariate. The computation of the p -values will thus be made for each station. The choice of the test statistic has to reflect the departures from the model that can be expected. In the present case study, the main hypotheses are as follows:

- ✓ The local data follow a GEV distribution: this is a standard hypothesis when dealing with block maxima, because this distribution arises from extreme value theory. Test quantities related to the first three empirical moments are computed, namely the empirical mean, standard deviation and skewness.
- ✓ All the shape parameters are identical: this implies that all sites should have a comparable tail behavior. The skewness, which only depends on the shape parameter for GEV data, will be a first indicator. The maximum value of the

sample is also computed, because it is very sensitive to the value of the shape parameter.

- ✓ All the stations are affected by a common trend: Kendall's tau coefficient between time and annual maxima is computed.
- ✓ The stations are independent: Kendall's tau coefficient between the annual maxima of two sites is derived.

The results are presented in Table 3. Values smaller than 0.05 or larger than 0.95 are denoted by an asterisk. The first three moments of the generated data are consistent with the observed ones. Moreover, the skewness and the maximum have acceptable p -values, which indicates that the regional shape parameter hypothesis is compatible with the observations. The regional trend hypothesis seems acceptable for five of the six stations. At the first station, the observed values lead to a larger trend than that simulated by the model. In contrast, the independence between stations is clearly doubtful: as an illustration, three of the five Kendall's tau coefficients between the first station and the others have a p -value smaller than 0.05. This is not surprising, as choosing a homogeneous region generally implies selecting dependent sites. The potential effect of this hypothesis, which is made for convenience, will be addressed in the discussion section.

IV.5. Frequency analysis

Although the model seems acceptable for describing the data, this does not imply that other models would not be able to do as well. The regional trend model M_1 is therefore compared with the corresponding stationary model M_0 , which means that at-site annual maxima are assumed to arise from stationary GEV distributions, $X_t^{(i)} \sim GEV(\alpha_i, \beta_i, \xi)$.

Marginal likelihoods were first computed for models M_0 and M_1 , thanks to the Chib approach presented in Appendix 1. The logarithm of the marginal likelihood was approximately equal to -863.2 for the stationary model and -859.8 for the regional trend model. Secondly, when equal prior probabilities for each model were assumed, the posterior probabilities were equal to (Equation (4)):

$$\hat{p}(M_0 | \mathbf{X}) = \frac{0.5 p(\mathbf{X} | M_0)}{0.5 p(\mathbf{X} | M_0) + 0.5 p(\mathbf{X} | M_1)} = \frac{1}{1 + \exp(\log p(\mathbf{X} | M_1) - \log p(\mathbf{X} | M_0))} \approx 0.03 \quad (16)$$

$$\text{and } \hat{p}(M_1 | \mathbf{X}) \approx 0.97$$

The model assuming a consistent regional trend has a high probability, which means that a stationary regional model is far less suitable for describing these data. The Bayes factor between the two models is thus approximately equal to 30, which suggests, according to the scale of Kass and Raftery (1995), a strong evidence for the regional trend model.

The posterior distribution of the 0.9-quantile was then derived for each of the six sites, by using the model averaging procedure described in section II.2. The temporal changes in the posterior means are shown in Figure 10, together with 90% probability intervals. The quantile has decreased by approximately 10% for these six stations between years 1960 and 2000.

Extrapolation to the future has to be considered with caution, because it implicitly assumes that the estimated trend will persist for ever. If the cause of the change is suspected, the use of a more suitable covariate than time may be able to deliver more realistic extrapolations.

The frequency analysis obtained with a regional model was finally compared with local analyses performed on each site. More accurately, for a given site i , the posterior distributions of the parameters in models $M_0^{(i)}$ and $M_1^{(i)}$ were estimated, leading to the posterior probabilities of each model and to the related Bayes factor. The results are presented in Table 4. It can be observed that the posterior probability of the stationary model remains far from zero or one, which indicates that at the at-site scale, there is no clear evidence for one model

or the other. Consequently, Bayes factors are close to one, except for site 6, where the evidence for stationarity can be considered as “positive”, according to the scale of Kass and Raftery (1995). This result emphasizes the advantage of regional analysis: by increasing the accuracy of regional estimates, several weak but consistent changes within a homogeneous region can result in a regional trend with high significance. The temporal changes of the 0.9-quantile were also studied at the local scale, and are shown in Figure 11. Bayesian model averaging is especially useful in this case: because of the difficulty of choosing between the two concurrent models, it leads to a more realistic quantification of the uncertainties, especially when extrapolating to the future. Compared to the quantiles obtained with the regional model, the probability intervals are thus wider, which reflects the stronger uncertainty related to local estimates. It can also be noticed that the posterior mean values of the quantiles obtained with the regional model remain inside the probability intervals obtained with the local models: regional and local estimates are thus consistent, which illustrates the adequacy of the regional model to describe these data.

IV.6. Discussion

The preceding case study provides an illustration of the usefulness of the Bayesian framework and the MCMC methods for the estimation of complex models. Estimation and detection of a trend in extreme data is generally acknowledged to be a difficult problem, because of the strong variability of such variables. The approach proposed here uses the paradigm of regional analysis, which maintains that the estimation can be improved by gathering data from several sites in a homogeneous region. The more basic regional estimation scheme is based on the well-known index flood method (Dalrymple, 1960): within a homogeneous hydrological region, all scaled annual maxima series are assumed to have the same distribution. In practice, scaling is performed by dividing data by the sample means, as an example. In a second step, a unique distribution is fitted to this sample of normalized data. Unfortunately, this method implies a strong underestimation of the uncertainties. The alternative method is to translate the index flood assumption onto the local distributions parameters, as suggested by Katz et al. (2002). An example of such a study can be found in Buishand (1991), who used a two-steps estimation procedure. A more efficient estimation scheme would be to simultaneously estimate all parameters, which can be done by means of MCMC methods.

Moreover, a common drawback of almost, if not all the regional methods for extremes in hydrology, is that the spatial dependence of the data is ignored. This issue has been addressed by several authors, namely by Stedinger (1983), Hosking and Wallis (1988; 1997), or Madsen and Rosbjerg (1997). All these studies aim at quantifying the effect of inter-site dependence on the estimates' accuracy, rather than including it in a formal statistical model. A shared conclusion is that the existence of spatial dependence increases the variance of the estimates, but leaves the bias unchanged. Moreover, the estimates accuracy, in terms of RMSE, remains larger in the case of a regional analysis with dependent sites than in the case of at-site analysis. This conclusion is even valid in moderately heterogeneous regions, *i.e.* regions where the regional distribution is mis-specified. However, these conclusions have to be mitigated in the present case study. First of all, these authors used a different estimation method, namely the probability-weighted moments or the *L*-Moments. Secondly, these results have been obtained using Monte Carlo simulations, and their validity needs to be verified in other simulation contexts. More accurately, dependent data have been simulated with a Gaussian copula: a multivariate Gaussian data set is first generated, and marginal data are transformed in order to fit an extreme value distribution. Although this dependence structure seems well-supported by some real-world data, it can not be considered as universal: for instance, Hosking and Wallis (1988) provide an example where very high flood events are more dependent than moderate ones. They explain this structure by the tendency of extreme

floods to be generated by meteorological conditions that affect large areas. Based on a simulation study which mimics this structure, they found that the loss of accuracy is larger than in the case of the constant correlation produced by the Gaussian copula. Notice that in some regions, the opposite structure might be observed (smaller dependence for very high flood events), because of the tendency of extreme floods to be generated by convective storms, which only affect a limited area. Finally, a major difference between these studies and the case study presented in this paper is the existence of a trend. This additional parameter has a strong variability, which may affect the influence of the spatial dependence. Consequently, additional studies involving contrasting dependence structures in a non-stationary context would be of great interest.

V. Conclusion and perspectives

The aim of this article was to explore a non-stationary regional model for describing annual maximum discharges of several gauging stations within a homogeneous region. The estimation of the parameters was achieved by means of the Bayesian framework and MCMC methods, which are convenient tools for overcoming the issues related to such complex models.

The regional approach adopted in this study consists in considering that some parameters should be identical for all sites within the region. The existence of such regional parameters can be viewed as a consequence of standard regional methods (e.g. the index flood method). A regional model has the advantage of avoiding a normalization step, which leads to an underestimation of the uncertainty. Moreover, in a non-stationary context, prior beliefs about the structure of a change are often available. For instance, if a change is related to a large-scale climatic phenomenon, it is natural to search for consistency within a homogeneous region. Finally, from a statistical point of view, regional analysis is a way of decreasing the estimation uncertainties, by combining data from different locations. This is particularly useful in the case of extreme value models, whose parameters are known to be particularly difficult to estimate. The case study described in this paper illustrates the benefit of a regional inference compared to at-site analyses, especially in terms of trend detection.

The Bayesian framework appears as a suitable method for the inference of this kind of model, because it presents several advantages compared to classical likelihood-based approaches. Firstly, the prior distribution is a convenient way to incorporate all the available data and knowledge in the analysis. Secondly, the posterior distribution allows a complete and comprehensive inference to be derived, without using any asymptotical assumption, which may be irrelevant for small samples. Moreover, in the case of extreme value distributions, the end-point of the density is a function of the parameters, because of the condition

$1 - \xi\left(\frac{x - \beta}{\alpha}\right) > 0$. Consequently, some parameter values are incompatible with the data. A

blind application of the asymptotic normality argument may thus lead in some cases to the assignment of a significant non-zero density to impossible parameters values. This drawback does not exist in the Bayesian framework, because the posterior density would be strictly equal to zero for such parameter values. Finally, an interesting characteristic of the Bayesian analysis is the possibility of taking into account modeling uncertainties thanks to the Bayesian model averaging approach.

In spite of these advantages, the development of Bayesian models used to be limited by numerical difficulties. Much progress on MCMC methods has been achieved since then, and has provided a number of tools for overcoming the numerical issues. Consequently, MCMC methods can find applications in a very wide range of situations, and are especially useful in

problems with a large number of dimensions, where standard optimization methods can fail. However, some care is needed in order to assess the algorithms' convergence or to prevent numerical errors. Accordingly, the use of MCMC methods requires familiarity with the algorithms' functioning, as using them as a "black box" tool can lead to misinterpreted results. In this paper, both an adaptive sampler and a Metropolis algorithm were used. The adaptive Gibbs/Metropolis algorithm might suffer from theoretical limitations, and should therefore be used with caution.

The usefulness of the regional model studied in this paper has been illustrated by the case study. However, several improvements may be achieved, in order to consider less restrictive hypotheses. Firstly, the equality of some parameters over the whole region is a strong assumption. A more plausible hypothesis consists in assuming that such parameters are the realization of a spatial process. In this way, these parameters do not need to be equal anymore, but should respect structural relationships which ensure smooth variations in the region. Once again, the Bayesian approach is a suitable tool for deriving such models, by means of a Bayesian hierarchical model. Perreault (2000) developed this approach in the Gaussian case in order to model consistent step-changes in several sites. The model proposed by Cooley (2005) is a hierarchical regional model for extreme precipitations, where the variation of the parameters in the region is viewed as the result of a spatial process, which includes covariates such as altitude. The application of this kind of model to a non-stationary context would be a very interesting development.

As discussed in the case study, the spatial independence hypothesis is also a restrictive assumption, whose impact is not obvious to anticipate, especially in a non-stationary context. A major improvement would thus be to model dependence explicitly. Unfortunately, a number of problems still remain from a theoretical point of view in the case of extreme data. Multivariate models for extremes are still challenging, as most of the existing approaches are based on a too poor description of dependences, or are too difficult to handle in real-life problems. Specific methods have thus been studied for modeling spatial extremes (Schalther and Tawn, 2003; Naveau et al., 2005; De Haan and Pereira, 2006). It has to be explored how these approaches can be included in regional models.

More generally, recent developments in statistics of extremes provide a number of tools for hydrological applications. Deriving adequate and explicit statistical models together with convenient inference methods is likely to comply with more statistical rigor and to greatly improve the description of hydrological phenomena.

VI. Bibliography

- Bernier, J. 1994. Statistical detection of changes in geophysical series. *Engineering Risk in Natural Resources Management* **275**:159 - 176.
- Bortot, P., S. Coles, and J. A. Tawn. 2000. The multivariate Gaussian tail model: an application to oceanographic data. *Applied Statistics* **49**:31-49.
- Buishand, T. A. 1991. Extreme Rainfall Estimation by Combining Data from Several Sites. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* **36**:345-365.
- Chib, S. 1995. Marginal Likelihood From the Gibbs Output. *Journal of American Statistical Association* **90**:1313-1321.
- Coles, S. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag. 210 p. Springer-Verlag, London.
- Coles, S., and L. Pericchi. 2003. Anticipating catastrophes through extreme value modelling. *Journal of the Royal Statistical Society Series C-Applied Statistics* **52**:405-416.

- Coles, S., L. R. Pericchi, and S. Sisson. 2003. A fully probabilistic approach to extreme rainfall modelling. *Journal of Hydrology* **273**:35-50.
- Coles, S., and J. A. Tawn. 1990. Statistics of coastal flood prevention. *Philosophical Transactions of the Royal Society of London Series a - Mathematical Physical and Engineering Sciences* **332**:457-476.
- Coles, S. G., and J. A. Tawn. 1996. A Bayesian analysis of extreme rainfall data. *Applied Statistics-Journal of the Royal Statistical Society Series C* **45**:463-478.
- Cooley, D. 2005. *Statistical Analysis of Extremes Motivated by Weather and Climate Studies: Applied and Theoretical Advances*. University of Colorado. 122 p.
- Dalrymple, T. 1960. Flood frequency analyses. *in* Water-supply paper 1553-A. US Geological Survey.
- De Haan, L., and T. T. Pereira. 2006. Spatial extremes: Models for the stationary case. *Annals of Statistics* **34**:146-168.
- El Adlouni, S., A. C. Favre, and B. Bobee. 2006. Comparison of methodologies to assess the convergence of Markov chain Monte Carlo methods. *Computational Statistics and Data Analysis*. In press.
- Favre, A. C., S. El Adlouni, L. Perreault, N. Thiémonge, and B. Bobee. 2004. Multivariate hydrological frequency analysis using copulas. *Water Resources Research* **40**.
- Fisher, R. A., and L. H. Tippett. 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Cambridge Phil. Soc.* **24**.
- Gelfand, A. E., and A. F. M. Smith. 1990. Sampling-based approaches to calculating marginal densities. *Journal of American Statistical Association* **85**:398-409.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. *Bayesian data analysis*. C. Hall. 526 p. Chapman & Hall.
- Gelman, A., G. O. Roberts, and W. R. Gilks. 1996. Efficient Metropolis Jumping Rules. Pages 599-607 *in* *Bayesian Statistics 5*. Oxford University Press.
- Geman, S., and D. Geman. 1984. Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**:721-741.
- Gilks, W. R., N. G. Best, and K. K. C. Tan. 1995. Adaptive Rejection Metropolis Sampling within Gibbs Sampling. *Applied Statistics-Journal of the Royal Statistical Society Series C* **44**:455-472.
- Gilks, W. R., and P. Wild. 1992. Adaptive Rejection Sampling for Gibbs Sampling. *Applied Statistics-Journal of the Royal Statistical Society Series C* **41**:337-348.
- Haario, H., E. Saksman, and J. Tamminen. 2001. An adaptive Metropolis algorithm. *Bernoulli* **7**:223-242.
- Haario, H., E. Saksman, and J. Tamminen. 2005. Componentwise adaptation for high dimensional MCMC. *Computational Statistics* **20**:265-273.
- Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**:97-109.
- Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky. 1999. Bayesian model averaging: A tutorial. *Statistical Science* **14**:382-401.
- Hosking, J. R. M., and J. R. Wallis. 1988. The effect of intersite dependence on regional flood frequency analysis. *Water Resources Research* **24**:588-600.
- Hosking, J. R. M., and J. R. Wallis. 1997. *Regional Frequency Analysis: an approach based on L-Moments*. 226 p. Cambridge University Press, Cambridge, UK.
- Jenkinson, A. F. 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society* **81**:158-171.

- Kass, R. E., and A. E. Raftery. 1995. Bayes Factors. *Journal of the American Statistical Association* **90**:773-795.
- Katz, R. W., M. B. Parlange, and P. Naveau. 2002. Statistics of extremes in hydrology. *Advances in Water Resources* **25**:1287-1304.
- Khodja, H., H. Lubès-Niel, J. M. Sabatier, E. Servat, and J. E. Paturel. 1998. Analyse spatio-temporelle de données pluviométriques en Afrique de l'Ouest. Recherche d'une rupture en moyenne. Une alternative intéressante: les tests de permutations. *Revue de Statistique Appliquée* **46**:95 -110.
- Lee, A. F. S., and S. M. Heghinian. 1977. A shift of the mean level in a sequence of independent normal random variables - a Bayesian approach. *Technometrics* **19**:503-506.
- Lu, Z. Q., and L. M. Berliner. 1999. Markov switching time series models with application to a daily runoff series. *Water Resources Research* **35**:523-534.
- Madsen, H., and D. Rosbjerg. 1997. The partial duration series method in regional index-flood modeling. *Water Resources Research* **33**:737-746.
- Marshall, L., D. Nott, and A. Sharma. 2004. A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling. *Water Resources Research* **40**.
- Martins, E. S., and J. R. Stedinger. 2000. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research* **36**:737-744.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *Journal of chemical physics* **21**:1087-1092.
- Metropolis, N., and S. Ulam. 1949. The Monte Carlo method. *Journal of the American Statistical Association* **44**:335-341.
- Moyeed, R. A., and R. T. Clarke. 2005. The use of Bayesian methods for fitting rating curves, with case studies. *Advances in Water Resources* **28**:807-818.
- Naveau, P., D. Cooley, and P. Poncet. 2005. Spatial extremes analysis in climate studies. *in*. *Extreme Value Analysis*, Gothenburg, Sweden.
- Ouarda, T. B. M. J., M. Lang, B. Bobee, J. Bernier, and P. Bois. 1999. Analysis of regional flood models utilized in France and Quebec. *Revue des Sciences de l'Eau* **12**:155-182.
- Parent, E., and J. Bernier. 2003. Encoding prior experts judgments to improve risk analysis of extreme hydrological events via POT modeling. *Journal of Hydrology* **283**:1-18.
- Perreault, L. 2000. Analyse bayésienne rétrospective d'une rupture dans les séquences de variables aléatoires hydrologiques. Thesis. ENGREF, INRS-Eau. 200 p.
- Perreault, L., J. Bernier, B. Bobee, and E. Parent. 2000a. Bayesian change-point analysis in hydrometeorological time series. Part 1. The normal model revisited. *Journal of Hydrology* **235**:221-241.
- Perreault, L., J. Bernier, B. Bobee, and E. Parent. 2000b. Bayesian change-point analysis in hydrometeorological time series. Part 2. Comparison of change-point models and forecasting. *Journal of Hydrology* **235**:242-263.
- Perreault, L., and V. Fortin. 2003. Mixture and Hidden Markov models for peak flow analysis. *in*. *Seizièmes entretiens du centre Jacques Cartier*, Lyon, France.
- Perreault, L., M. Hache, M. Slivitzky, and B. Bobee. 1999. Detection of changes in precipitation and runoff over eastern Canada and US using a Bayesian approach. *Stochastic Environmental Research and Risk Assessment* **13**:201-216.
- Perreault, L., E. Parent, J. Bernier, B. Bobee, and M. Slivitzky. 2000c. Retrospective multivariate Bayesian change-point analysis: A simultaneous single change in the mean of several hydrological sequences. *Stochastic Environmental Research and Risk Assessment* **14**:243-261.

- Reis, D. S., and J. R. Stedinger. 2005. Bayesian MCMC flood frequency analysis with historical information. *Journal of Hydrology* **313**:97-116.
- Renard, B., M. Lang, and P. Bois. 2006. Statistical analysis of extreme events in a non-stationary context via a Bayesian framework. *Stochastic Environmental Research and Risk Assessment*. Under press.
- Ribatet, M., E. Sauquet, J. M. Gresillon, and T. B. M. J. Ouarda. 2006. A regional Bayesian POT model for flood frequency analysis. *Stochastic Environmental Research and Risk Assessment*. accepted.
- Ritter, C., and M. A. Tanner. 1992. Facilitating the Gibbs Sampler - the Gibbs Stopper and the Griddy-Gibbs Sampler. *Journal of the American Statistical Association* **87**:861-868.
- Roberts, G. O., and W. R. Gilks. 1994. Convergence of Adaptive Direction Sampling. *Journal of Multivariate Analysis* **49**:287-298.
- Schalther, M., and J. A. Tawn. 2003. A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika* **90**:139-156.
- Stedinger, J. R. 1983. Estimating a regional flood frequency distribution. *Water Resources Research* **19**:503-510.
- Tanner, M. A. 1996. *Tools for Statistical Inference*. Springer-Verlag. 208 p. New York.
- Tapsoba, D., M. Hache, L. Perreault, and B. Bobee. 2004. Bayesian Rainfall Variability Analysis in West Africa along Cross Sections in Space-Time grid Boxes. *Journal of Climate* **17**:1069-1082.
- Thyer, M., and G. Kuczera. 2000. Modeling long-term persistence in hydroclimatic time series using a hidden state Markov model. *Water Resources Research* **36**:3301-3310.
- Thyer, M., and G. Kuczera. 2003a. A hidden Markov model for modelling long-term persistence in multi-site rainfall time series 1. Model calibration using a Bayesian approach. *Journal of Hydrology* **275**:12-26.
- Thyer, M., and G. Kuczera. 2003b. A hidden Markov model for modelling long-term persistence in multi-site rainfall time series. 2. Real data analysis. *Journal of Hydrology* **275**:27-48.
- Thyer, M., G. Kuczera, and Q. J. Wang. 2002. Quantifying parameter uncertainty in stochastic models using the Box-Cox transformation. *Journal of Hydrology* **265**:246-257.
- Tierney, L. 1994a. Markov-Chains for Exploring Posterior Distributions. *Annals of Statistics* **22**:1701-1728.
- Tierney, L. 1994b. Markov-Chains for Exploring Posterior Distributions - Rejoinder. *Annals of Statistics* **22**:1758-1762.
- Tierney, L., and A. Mira. 1999. Some adaptive Monte Carlo methods for Bayesian inference. *Statistics in Medicine* **18**:2507-2515.
- Wang, Q. J. 2001. A Bayesian joint probability approach for flood record augmentation. *Water Resources Research* **37**:1707-1712.

Appendix 1: the Chib method

Marginal likelihood estimation will be based on the following observation, which holds for every $\theta = (\theta_1, \dots, \theta_d)$ where π and p are defined:

$$p(\mathbf{X} | M) = \frac{\pi(\theta)p(\mathbf{X} | \theta)}{p(\theta | \mathbf{X})} = \frac{f(\theta)}{p(\theta | \mathbf{X})} \quad (17)$$

In this equation, f is the non-normalized target distribution, which can always be computed at an arbitrary θ . The problem thus consists in estimating the posterior density of θ , by using samples generated by MCMC methods.

Chib (1995) proposed using the following equation:

$$p(\theta_1^*, \dots, \theta_d^* | \mathbf{X}) = p(\theta_1^* | \mathbf{X}) p(\theta_2^* | \theta_1^*, \mathbf{X}) \dots p(\theta_q^* | \theta_1^*, \theta_2^*, \dots, \theta_{q-1}^*, \mathbf{X}) \dots \times p(\theta_d^* | \theta_1^*, \dots, \theta_{d-1}^*, \mathbf{X}), \quad (18)$$

where $\theta^* = (\theta_1^*, \dots, \theta_d^*)$ is a given parameter vector with dimension d , preferably lying in a high density area. Each of the terms in equation (18) is a marginal density: for instance,

$p(\theta_q^* | \theta_1^*, \theta_2^*, \dots, \theta_{q-1}^*, \mathbf{X})$ is the first marginal of the multivariate distribution

$p(\theta_q, \theta_{q+1}, \dots, \theta_d | \theta_1^*, \theta_2^*, \dots, \theta_{q-1}^*, \mathbf{X})$, evaluated at θ_q^* . MCMC methods can therefore be applied to the non-normalized posterior density with the first $q-1$ components being fixed, that is $f(\theta_1^*, \theta_2^*, \dots, \theta_{q-1}^*, \theta_q, \theta_{q+1}, \dots, \theta_d)$, in order to generate a sample from this distribution. The problem then reduces to univariate density estimation using a sample of values. In this paper, a Gaussian kernel is used for this purpose.

Appendix 2: the Metropolis sampler

The Metropolis sampler was implemented as follows:

- ✓ Choose a starting point $\theta^{(0)}$ and a variance matrix Σ .
- ✓ Do $i = 1, \dots, N_{iter}$
 - ✓ Generate a candidate vector $\theta^* \sim N(\cdot | \theta^{(i-1)}, \Sigma)$
 - ✓ If $p(\theta^* | \mathbf{X}) \geq p(\theta^{(i-1)} | \mathbf{X})$, set $\theta^{(i)} = \theta^*$, else accept the candidate vector $(\theta^{(i)} = \theta^*)$ with probability $r = \frac{p(\theta^* | \mathbf{X})}{p(\theta^{(i-1)} | \mathbf{X})}$ or reject it $(\theta^{(i)} = \theta^{(i-1)})$ with probability $1-r$.

In order to avoid numerical overflows, it is useful to consider the logarithm of the posterior distribution, and to compute the posterior ratio as

$r = \exp(\log(p(\theta^* | \mathbf{X})) - \log(p(\theta^{(i-1)} | \mathbf{X})))$. Moreover, this ratio is invariant by multiplying the posterior distribution by a constant, which implies that the Metropolis can be applied to a non-normalized target distribution.

Appendix 3. The Adaptive Gibbs/Metropolis sampler.

This algorithm is implemented as follows:

- ✓ Choose a starting point $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})$, and starting variances for each component $V^{(0)} = (V_1^{(0)}, \dots, V_d^{(0)})$
- ✓ Do $j=1, \dots, N_{Gibbs}$:
 - ✓ Do $q=1, \dots, d$
 - ✓ Let $\gamma^{(0)} = \theta_q^{(j-1)}$ and $f_q^{(j)}(x) = p(x | \theta_1^{(j)}, \dots, \theta_{q-1}^{(j)}, \theta_{q+1}^{(j-1)}, \dots, \theta_d^{(j-1)}, \mathbf{X})$
 - ✓ Do $i=1, \dots, N_{Metro}$
 - ✓ Generate a candidate value $\gamma^* \sim N(\cdot | \gamma^{(i-1)}, V_q^{(j-1)})$
 - ✓ If $f_q^{(j)}(\gamma^*) \geq f_q^{(j)}(\gamma^{(i-1)})$, set $\gamma^{(i)} = \gamma^*$, else accept the candidate vector $(\gamma^{(i)} = \gamma^*)$ with probability $r = \frac{f_q^{(j)}(\gamma^*)}{f_q^{(j)}(\gamma^{(i-1)})}$ or reject it $(\gamma^{(i)} = \gamma^{(i-1)})$ with probability $1-r$.
 - ✓ Set $\theta_q^{(j)} = \gamma^{(N_{Metro})}$
 - ✓ Compute the acceptance rate τ on the previous N_{Metro} iterations
 - ✓ If $\tau < 0.23$, set $V_q^{(j)} < V_q^{(j-1)}$
 - ✓ If $\tau > 0.44$, set $V_q^{(j)} > V_q^{(j-1)}$

LIST OF CAPTIONS

TABLES

Table 1: Hydrological properties of the six stations used in the case study.

Table 2: Prior specifications for the location and scale parameters. a and b are the parameters of the Gamma distributions used as prior distributions.

Table 3: Bayesian p -values obtained for various characteristics of the regional distribution.

Table 4: Posterior probabilities of local stationary models (compared to local trend models), and related Bayes factors.

FIGURES

Figure 1: Simulated values of the first component from three target distributions. Real marginal densities are denoted by a solid line. (a) Gaussian distribution (b) Independent GEV marginal distributions. (c) Independent mixed-Gaussian marginal distributions.

Figure 2: Simulated values of the last component from three target distributions. Real marginal densities are denoted by a solid line. (a) Gaussian distribution (b) Independent GEV marginal distributions. (c) Independent mixed-Gaussian marginal distributions.

Figure 3: Location of the six sites used in the case study.

Figure 4: Gibbs/Metropolis simulated random walks with ten different starting points (black squares): (a) scale and location parameters for the first site, (b) shape and trend regional parameters.

Figure 5: $\sqrt{\hat{R}}$ statistic for different run lengths. (a) Gibbs/Metropolis, (b) Metropolis.

Figure 6: Histograms of marginal simulated samples for two local parameters of site 1. Locally-estimated values are denoted by vertical lines and prior distributions by thin lines.

Figure 7: Histograms of marginal simulated samples for the two regional parameters. Locally-estimated values are denoted by vertical lines and prior distributions by thin lines.

Figure 8: Annual maxima series for the six stations studied. The line denotes the regional trend on the mean.

Figure 9: Comparison of at-site estimates (site number 1, ..., 6) and regional estimates for the shape (left) and the trend (right) parameters.

Figure 10: Posterior mean and 90% probability interval for the 0.9-quantile of the six stations studied. The posterior distributions arise from the regional models M_0 and M_I .

Figure 11: Posterior mean and 90% probability interval for the 0.9-quantile of the six stations studied. For each site i , the posterior distributions arise from the local models $M_0^{(i)}$ and $M_1^{(i)}$. For comparison purposes, the axes are identical to those of Figure 10.

Site Number	River / Station	Drainage area (km ²)	Mean annual flow (m ³ .s-1)	10-year daily discharge (m ³ .s-1)
1	Dordogne / Saint Sauves d'Auvergne	87	3.6	43
2	Santoire / Condat	172	4.71	74
3	Vézère / Bugeat	143	4.43	47
4	Vézère / Uzerche	601	15.1	136
5	Corrèze / Corrèze	168	5.6	61
6	Corrèze / Brive-la-Gaillarde	947	21.2	268

Table 1

Site number	Sample size used for prior specification	Location parameter		Scale parameter	
		<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>
1	39	20.03	2.46	41.88	2.68
2	20	21.09	1.72	84.11	2.42
3	18	28.39	2.20	65.12	2.26
4	51	66.96	2.10	244.19	3.19
5	26	21.51	1.32	51.40	1.47
6	48	72.73	1.39	346.00	2.22

Table 2

	Bayesian <i>p</i> -value					
	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
Mean	0.547	0.202	0.741	0.911	0.749	0.922
Standard deviation	0.400	0.154	0.711	0.936	0.919	0.939
Skewness	0.156	0.185	0.346	0.801	0.851	0.542
Maximum value	0.288	0.127	0.724	0.914	0.920	0.807
Kendall's tau with time	0.976 *	0.786	0.398	0.622	0.372	0.670
Kendall's tau with site 1	/	0.003 *	0.068	0.042 *	0.036 *	0.119

Table 3

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6
Posterior probability $p(M_0^{(i)} \mathbf{X})$	0.66	0.30	0.38	0.53	0.28	0.79
Bayes factor $B_{0,1}$	1.95	0.43	0.61	1.11	0.39	3.86

Table 4

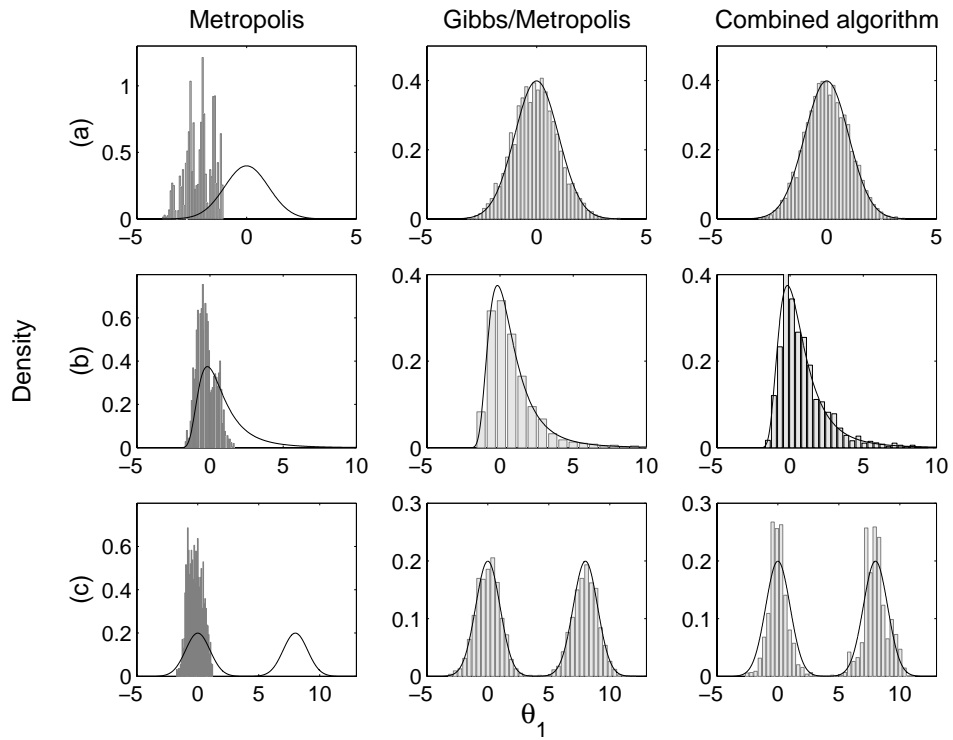


Figure 1

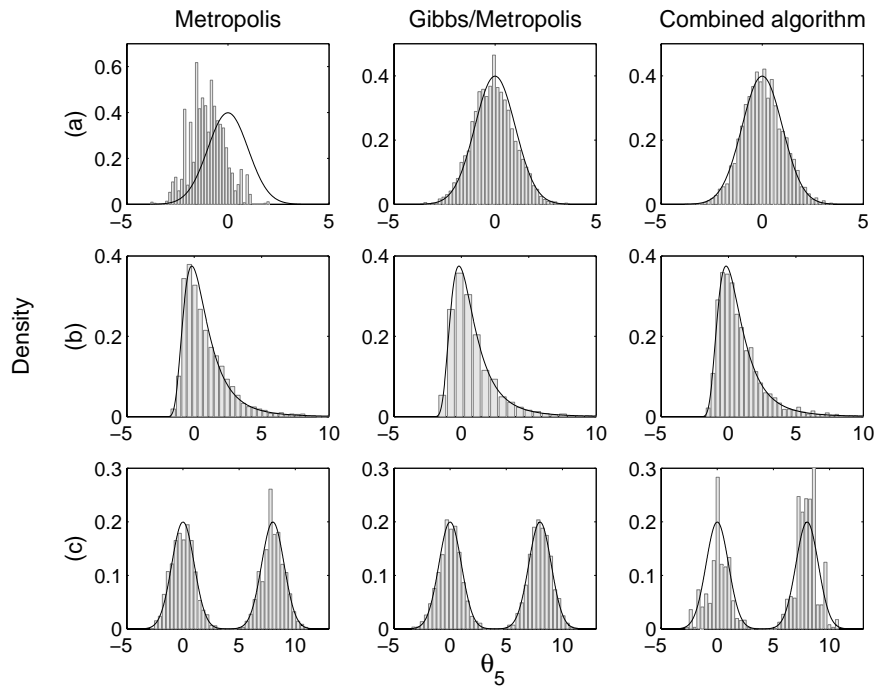


Figure 2

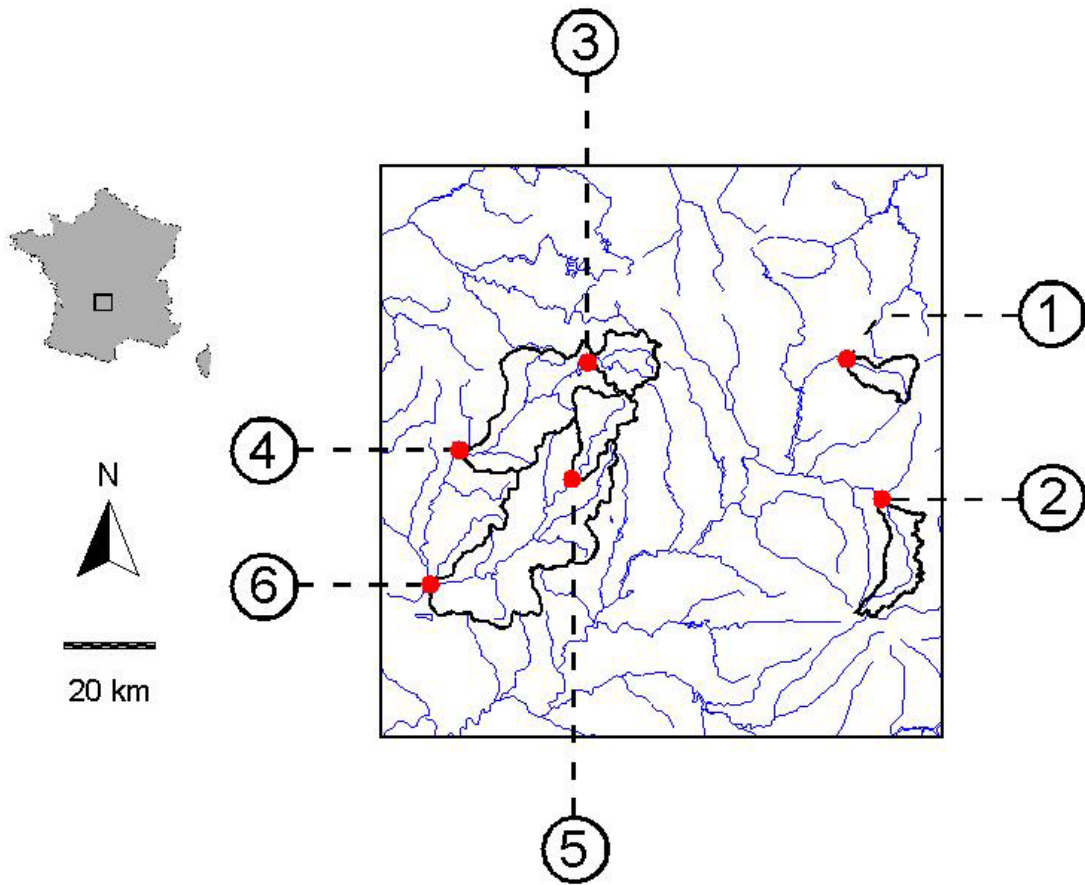


Figure 3

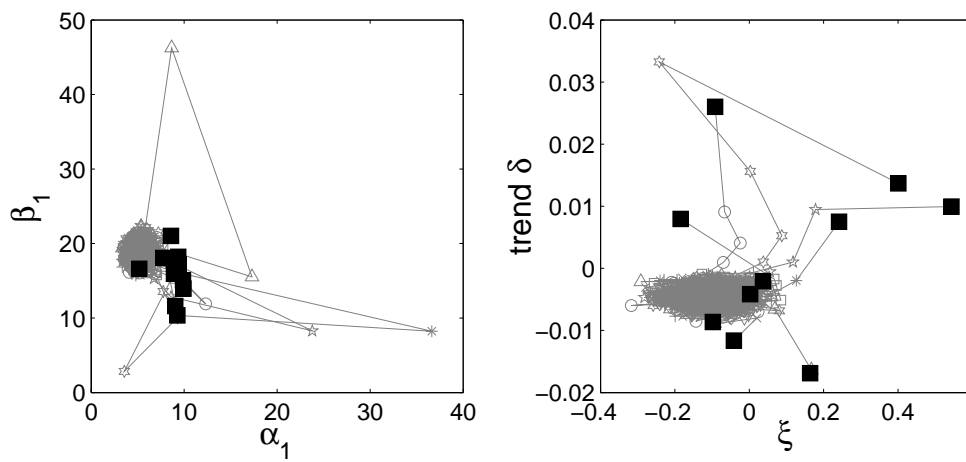


Figure 4

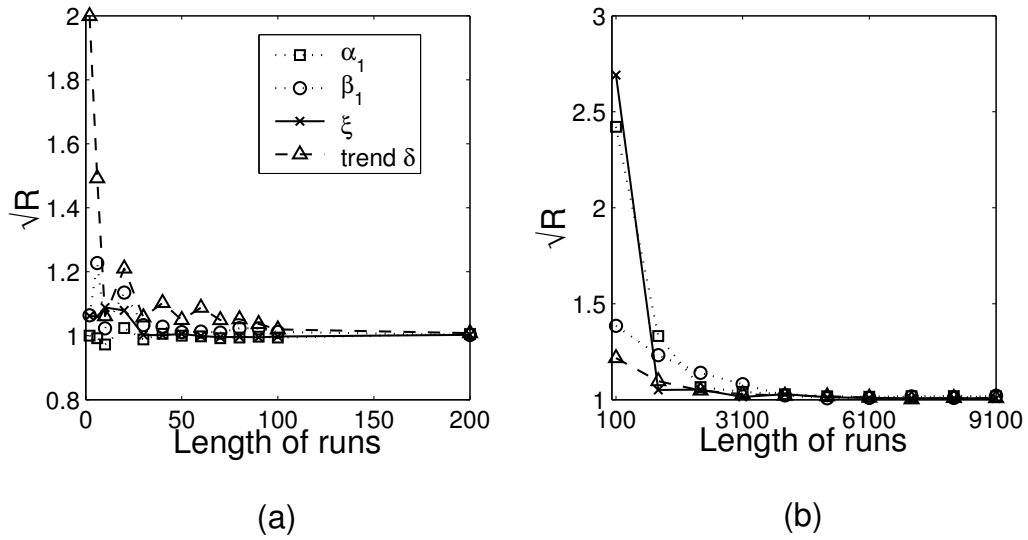


Figure 5

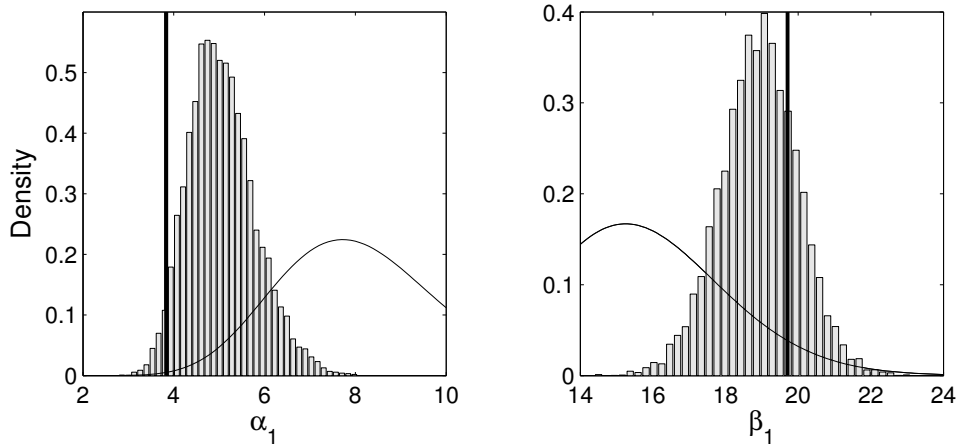


Figure 6

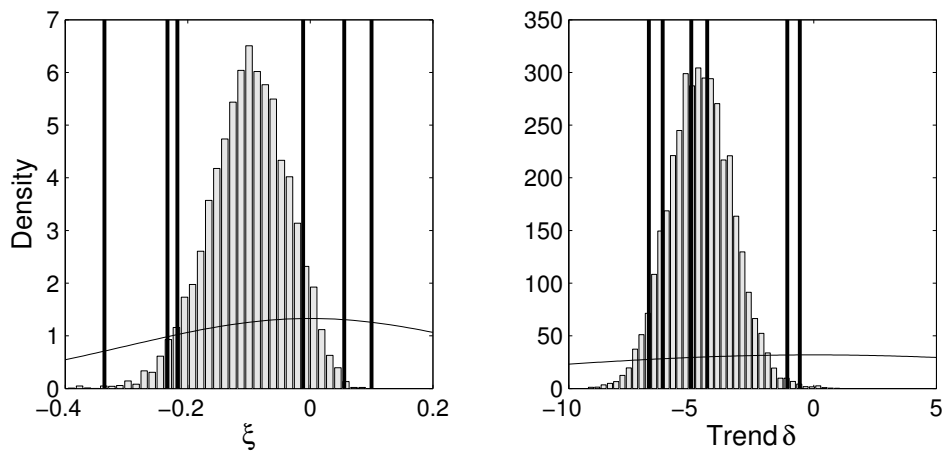


Figure 7

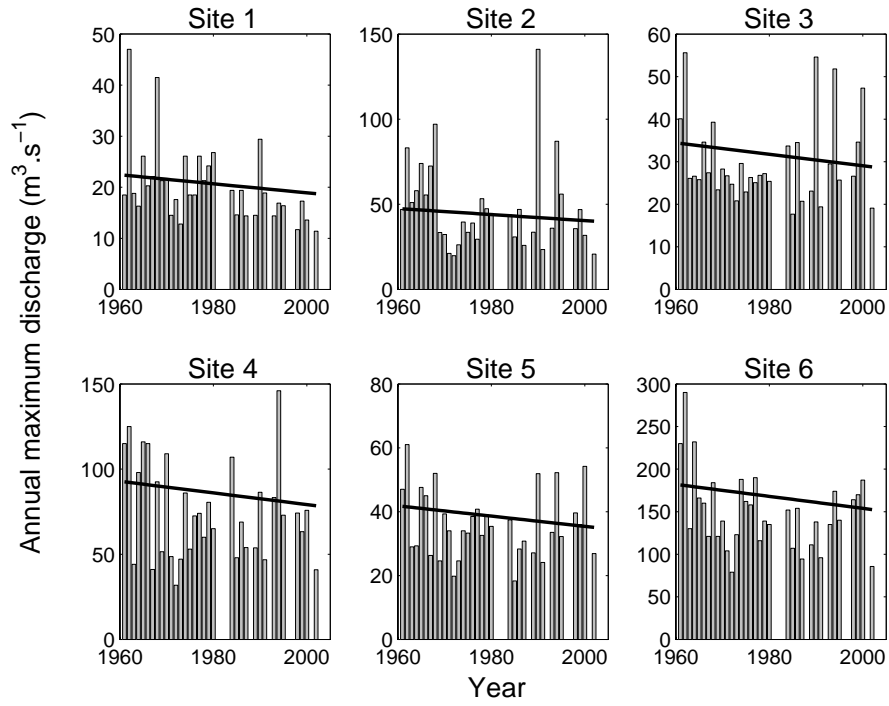


Figure 8

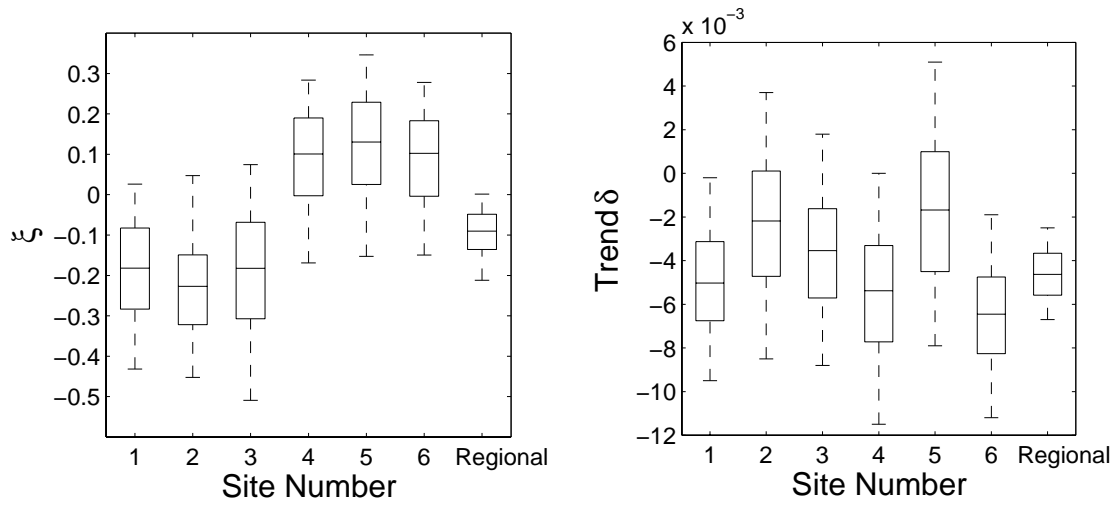


Figure 9

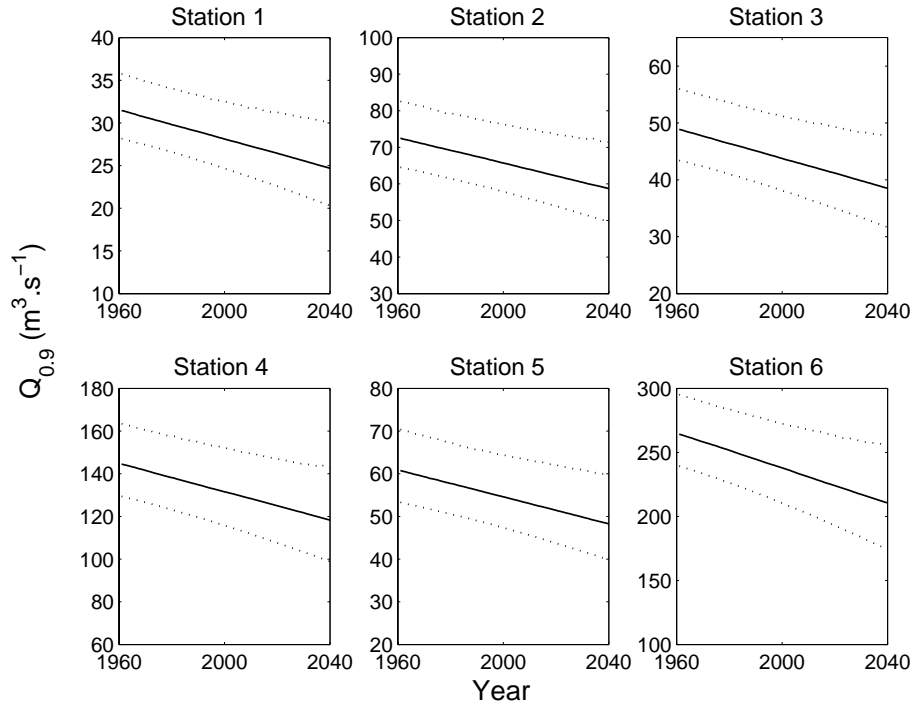


Figure 10

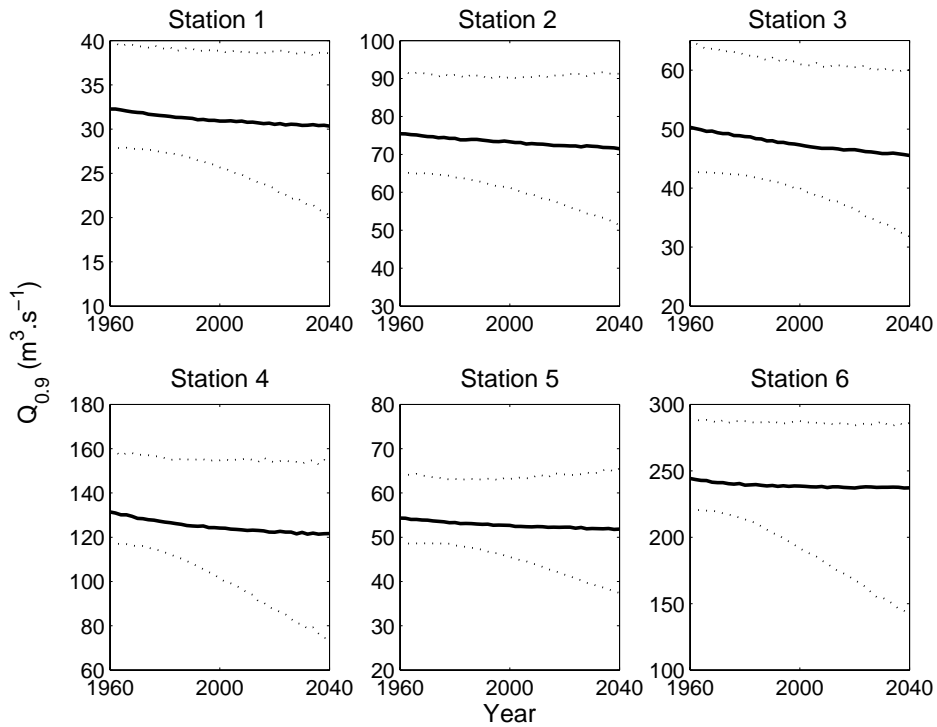


Figure 11

ARTICLE 3 : UTILISATION DE COPULES

Article soumis à *Advances in Water Resources*.

Accepté, sous presse.

USE OF A GAUSSIAN COPULA FOR MULTIVARIATE EXTREME VALUE ANALYSIS: SOME CASE STUDIES IN HYDROLOGY.

B. RENARD¹, M. LANG¹

(1) *Cemagref Centre de Lyon, U.R. Hydrologie-Hydraulique, 3 bis Quai Chauveau, CP 220, 69336 Lyon cedex 09, France.*

telephone:33 4 72 20 87 72

fax:33 4 78 47 78 75

e-mail : renard@lyon.cemagref.fr, lang@lyon.cemagref.fr

Abstract: Risk assessment requires a description of the probabilistic properties of hydrological variables. In a number of cases, this description is made on a single variable, whereas most hydrological events are intrinsically multivariate. In this context, copulas have recently received attention in order to derive a multivariate frequency analysis. After a reminder of the general results in the field of multivariate extreme value theory, the paper gives a description of a very simple copula, the Gaussian copula. Four case studies demonstrate its usefulness in the contexts of field significance determination, regional risk analysis, Discharge-Duration-Frequency (QdF) models with design hydrograph derivation and regional frequency analysis. The limitations and potential errors related to this statistical tool are also highlighted.

Keywords: multivariate analysis; extreme value analysis; extremes dependence; copula; field significance; regional frequency analysis; QdF models; design hydrographs; asymptotic properties; risk assessment.

I. Introduction

Extreme value theory (EVT) is widely used by hydrologists, especially for flood and drought mitigation. The severity of a hydrological extreme event is expressed as a non-exceedance probability, or equivalently, in terms of return period. This can be used to locate an observed event at a probabilistic scale and the theory can be used for extrapolation, i.e. for the determination of the probability of observing an event, even outside the range of observations. In some cases, hydrologists can be interested in several hydrological variables, which would lead to use of a multivariate statistical tool. Examples of such a situation in the flood mitigation field include the following:

- A flood event can be studied in different ways. The most commonly assessed trait is the peak flow of the event, but the volume or threshold exceedance duration can also be of interest. As an example, an event with a peak of a hundred-years return period could be less damaging than an event with a ten-years return period both in peak and volume. Unfortunately, these traits are not independent, thus preventing each variable from being studied separately. The multivariate distribution of the triplet (peak, volume, duration) is needed. Such an analysis has been undertaken by Adamson et al. [1] or Grimaldi and Serinaldi [23] as an example.
- A flood event at a river confluence can be the result of high discharges in only one or both of the upstream flows. If the discharges of these two streams are independent, the probabilistic behavior of the downstream flow, equal to the sum of the two upstream ones, can be obtained by convolution. Nevertheless, both tributaries can tend to have simultaneous high flows. In this case, the preceding calculation can lead to a strong underestimation of the risk. Once again, the multivariate distribution of the two upstream flows is needed. Some examples are provided by Mousavi [46], Le Clerc and Lang [35] and Favre et al. [17]. Coles [6] describes the underlying methodology.
- Water resources managers usually have to deal with several rivers in a given geographical area. A currently observed phenomenon is that a ten-year flood event is observed almost every year at a regional scale: the at-site non-exceedance probability is not suitable at the regional scale [62]. This phenomenon can easily be quantified in the case of M independent sites: the probability of observing at least one event with non-exceedance probability p is equal to $1 - p^M \geq 1 - p$. In contrast, if all sites are perfectly correlated, then the regional and the local probabilities are equal. Between these two extreme cases, which are never encountered in practice, the regional behavior of flood events is related to the dependence between stations, which has to be taken into account through a multivariate distribution.

Multivariate extreme value theory (MEVT) is studied by mathematicians (see Coles [6] for a review of the topic). Unfortunately, theoretical results are often difficult to use in practice, compared to the classical EVT. Additional hypotheses therefore have to be made to take into account dependence between variables. In this context, copulas have recently received particular attention [4, 13, 17, 21, 23, 54, 57], because they are relatively easy to handle and they can be used in a wide range of situations. In counterpart, they are based on a model of the dependence structure, which has to be checked for adequacy.

The aim of this article is to present some examples involving a particular copula, the Gaussian copula. In the first part, a review of some results of MEVT will be given, with a general description of copulas, and a more detailed description of the Gaussian copula (section II). The first application deals with the problem of the field significance determination in multi-

testing problems (section III.1). The second example deals with regional risk estimation, as presented in the introduction (section III.2). In section III.3, the Discharge-Duration-Frequency (QdF) methodology will be studied, and a copula will be used to compute the return period of a design hydrograph. The problem of intersite dependence and its implications for regional frequency analysis will be explored in section III.4. Examples where a Gaussian copula is not able to model properly the dependence structure of observations will also be presented (section III.5). Finally, the results obtained in this paper will be discussed (section IV), before giving some conclusions about the advantages and drawbacks in using a Gaussian copula, and proposing some perspectives for improving multivariate hydrological extreme events analysis (section V).

II. Multivariate extreme value theory

II.1. Some results of MEVT for componentwise maxima

Let (X_1, \dots, X_n) be independent and identically distributed variables. Let M_n denote the maximum of these variables, $M_n = \max_{i=1, \dots, n}(X_i)$. Under regularity conditions, it can be proven [18] that the only possible limit distribution of M_n , normalized by suitable values, is the general extreme value distribution (GEV), whose density is:

$$f(x; \alpha, \beta, \xi) = \frac{1}{\alpha} \left(1 - \frac{\xi(x - \beta)}{\alpha} \right)^{\frac{1}{\xi} - 1} \exp \left[- \left(1 - \frac{\xi(x - \beta)}{\alpha} \right)^{\frac{1}{\xi}} \right] \quad (1)$$

This theorem justifies the use of the GEV distribution to model annual maximum discharges, for example. In the multivariate case, such a theorem is also available. For clarity the bivariate case will be described. Let $(X_i, Y_i)_{i=1, \dots, n}$ be independent and identically distributed vectors.

Let M_n denote the 2-dimensional vector of the componentwise maxima,

$M_n = (M_n^x, M_n^y) = \left(\max_{i=1, \dots, n}(X_i), \max_{i=1, \dots, n}(Y_i) \right)$. As previously, the only possible limit distribution of M_n has the form [6]:

$$\begin{aligned} G(x, y) &= \Pr(M_n^x / n \leq x, M_n^y / n \leq y) \\ &= \exp \{ -V(\tilde{x}, \tilde{y}) \} \end{aligned}$$

where:

$$\begin{aligned} \tilde{x} &= \left(1 - \frac{\xi_x}{\alpha_x} (x - \beta_x) \right)^{-1/\xi_x}, \quad \tilde{y} = \left(1 - \frac{\xi_y}{\alpha_y} (y - \beta_y) \right)^{-1/\xi_y} \\ V(x, y) &= 2 \int_0^1 \max \left(\frac{w}{x}, \frac{1-w}{y} \right) dH(w) \end{aligned} \quad (2)$$

and H is a distribution on $[0; 1]$ with a mean of $1/2$.

Unfortunately, it appears that the limit distribution cannot be expressed in a classical parameterized way: an infinity of H distributions is able to validate the mean constraint. The choice of this function is the main difficulty of multivariate extreme value analysis. In practice, $G(x, y)$ can be chosen in a parametric family of distributions, which will be able to model a wide range of dependence structures, ranging from independence to total dependence. An example of such a family is the logistic family [6]. Up to this point, it is important to notice that this choice is a hypothesis which is not based on theoretical results, in

contrast to the use of the GEV in the univariate case. Consequently, it is necessary to check the adequacy of the dependence model and to keep in mind that the overall uncertainty is made up of the uncertainty related to the estimation of the parameters, which can be quantified by standard statistical methods, and of the modeling uncertainty, which is more difficult to assess.

II.2. Copulas

Copulas are alternative tools for dealing with multivariate extremes, and have become very popular in recent years. Let $(X^{(1)}, \dots, X^{(d)})$ be a d -dimensional random vector with a probability distribution of:

$$F(x_1, \dots, x_d) = \Pr(\{X^{(1)} \leq x_1\} \cap \dots \cap \{X^{(d)} \leq x_d\}) \quad (3)$$

A copula is a function c verifying:

$$c : [0;1]^d \mapsto [0;1]$$

and

$$(4)$$

$$F(x_1, \dots, x_d) = c(F_1(x_1), \dots, F_d(x_d))$$

The c function is used to model dependence between variables, while marginal distributions can be described in an usual way, with for example GEV distributions. Consequently, the copula function can be used in order to derive a multivariate distribution with the desired marginal distributions. Conversely, let F be a multivariate distribution with marginal distributions F_1, \dots, F_d . It can then be proven that c exists [56], which means that any multivariate distribution can be written in the form of equation (4). Moreover, if the marginal distributions F_i are continuous, then c is unique.

Nevertheless, as with the H distribution, no theoretical results are available to determine the copula: given marginal distributions, an infinity of multivariate joint distributions can be derived. Once again, a parametric family has to be chosen to model the dependence (see Favre et al. [17] for a review). As an illustration, Archimedean copulas can be constructed as follows. Let $\varphi : [0;1] \rightarrow [0; +\infty]$ be a continuous decreasing function such that $\varphi(1) = 0$ and $(-1)^k d^k \varphi^{-1}(t) / dt^k \geq 0$ for all t in $]0; +\infty[$ and $k=1, \dots, d$, where d is the number of dimensions. The following function is then a copula:

$$c : [0;1]^d \mapsto [0;1]$$

$$c(x_1, \dots, x_d) = \begin{cases} \varphi^{-1}\left(\sum_{i=1}^d \varphi(F_i(x_i))\right) & \text{if } \sum_{i=1}^d \varphi(F_i(x_i)) \leq \varphi(0) \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Classical choices for the φ generator function include $\varphi(t) = t^{-\alpha} - 1$ (Clayton copula),

$\varphi(t) = \log\left(\frac{e^{\alpha t} - 1}{e^\alpha - 1}\right)$ (Frank copula) or $\varphi(t) = (-\log t)^\alpha$ (Gumbel-Hougaard copula). The α

parameter thus summarizes the dependence. Although such copulas are very useful in the bivariate case, they are more problematic with a high number of dimensions, because a single parameter is not sufficient to describe a random vector with contrasted levels of dependence between marginal components. It is thus necessary to generalize the previous generator functions, by including additional parameters. Grimaldi and Serinaldi [23] provide an example of such a generalization. An alternative family may be easier to handle in highly dimensional cases and comprises the Elliptical copulas. These copulas use a symmetric and positive definite matrix in order to model dependence. The elements of this matrix can be interpreted as dependence measures between couples of variables, leading to an analogy with

the correlations used in the case of a multivariate Gaussian distributions. The most well known elliptical copulas are the Gaussian copula, which will be described in details later, and the Student's copula.

Whatever the family chosen, the number of parameters quickly grows with dimension. Estimation using standard methods such as maximum likelihood can be impossible because of numerical difficulties. A common estimation scheme consists in estimating marginal parameters separately as a first step, then estimating copula parameters knowing marginal estimations. On the other hand, the efficiency of estimators is not ensured.

Although appealing, the copula theory also suffers from a number of drawbacks, essentially because of its lack of connection with standard multivariate extreme value theory. Although specific copulas have been proposed to account for the theoretical properties of extremes [4, 22], additional developments are still needed, especially in the field of spatial extremes. This problem may be a limit for extrapolation, i.e. for the computation of very low probabilities, corresponding to multivariate events lying outside the observations range. Mikosh [45] thus lists a number of questions that remain unsolved, and provides a skeptical point of view about the use of copulas.

II.3. The Gaussian copula

The Gaussian copula is a member of the Elliptical copulas family. The dependence is modeled by means of a symmetric and definite positive matrix, whose elements are used to describe the dependence between couples of variables. This model is thus convenient when the number of dimensions is more than two or three, and has been widely studied in finance (see Cherubini et al. [5] and references therein). In the hydrology field, the meta-Gaussian density studied by Kelly and Krzysztofowicz [32] and Herr and Krzysztofowicz [25] can also be viewed as a Gaussian copula.

The Gaussian copula is defined as follows:

$$c(u_1, \dots, u_d) = \Phi_d \left(\phi^{-1}(u_1), \dots, \phi^{-1}(u_d) \right) \quad (6)$$

where ϕ is the cdf of the standard normal distribution $N(0, I)$ and Φ_d is the cdf of a multivariate normal distribution with mean θ and covariance matrix Σ .

In other words, the multivariate cumulative distribution of the data will have the following form:

$$\begin{aligned} F(x_1, \dots, x_d) &= c(F_1(x_1), \dots, F_d(x_d)) \\ &= \Phi_d \left(\phi^{-1}(F_1(x_1)), \dots, \phi^{-1}(F_d(x_d)) \right) \end{aligned} \quad (7)$$

or equivalently, the multivariate density can be written as:

$$f(x_1, \dots, x_d) = f_1(x_1) \times \dots \times f_d(x_d) \times |\Sigma|^{-1/2} \exp \left\{ - \frac{\left[\phi^{-1}(F_1(x_1)), \dots, \phi^{-1}(F_d(x_d)) \right] \left[\Sigma^{-1} - I \right] \left[\phi^{-1}(F_1(x_1)), \dots, \phi^{-1}(F_d(x_d)) \right]^T}{2} \right\} \quad (8)$$

The main advantage of this copula is its simplicity: once the data have been transformed by $\phi^{-1}(F_i(\cdot))$, the well known multivariate normal distribution is used to calculate probabilities. Figure 1 illustrates the principle of the copula, while Figure 2 shows the shape of the multivariate distribution obtained in a two-dimensional case. It also shows the limit of the Gaussian copula: with GEV marginal distributions, the distribution of the data has to be suitably described by this V-shaped dependence. In particular, a Gaussian copula will not be able to model properly more complex dependence structures. As an example, dependence between extreme variables may depend on marginal values. In other cases, it may be inappropriate to model the dependence of transformed variables by simple correlations. It is

therefore necessary to check the adequacy of the dependence structure implied by the Gaussian copula. Standard model diagnosis used in multivariate Gaussian analysis can be used for this purpose. As an example, Chi-square plots provide a graphical diagnosis for multivariate normality. The principle can be described as follows: let $(\tilde{X}^{(1)}, \dots, \tilde{X}^{(d)})$ denote the d -dimensional variable of interest, transformed toward standardized marginal normality thanks to $\phi^{-1}(F_i(\cdot))$. Under the assumption of multivariate normality $N(\boldsymbol{\theta}; \boldsymbol{\Sigma})$, the following quadratic form should fit a χ_d^2 distribution:

$$D = (\tilde{X}^{(1)}, \dots, \tilde{X}^{(d)}) \boldsymbol{\Sigma}^{-1} (\tilde{X}^{(1)}, \dots, \tilde{X}^{(d)})^T \quad (9)$$

In practice, the data are transformed by a normal-score transformation, in order to avoid departure from the expected distribution which may be due to marginal estimation errors, and the correlation matrix is estimated empirically on the transformed data. The transformed data are then used to construct a sample of d_i , by:

$$d_i = (\tilde{x}_i^{(1)}, \dots, \tilde{x}_i^{(d)}) \hat{\boldsymbol{\Sigma}}^{-1} (\tilde{x}_i^{(1)}, \dots, \tilde{x}_i^{(d)})^T \quad (10)$$

A QQ-plot can finally be constructed: if $d_{(i)}$ denotes the i^{th} sorted value of $(d_i)_{i=1, \dots, n}$ and $\chi_d^2(\alpha)$ the α -quantile of a Chi-square distribution with d degrees of freedom, then the points $\left(d_{(i)}; \chi_d^2 \left(\frac{i-0.5}{n} \right) \right)$ should remain close to the $y=x$ line. Alternative diagnostic tools may be

used, for example by using univariate goodness-of-fit tests based on linear combinations of the marginal values (see *e.g.* Mardia [40] for additional methods).

The next step consists in estimating parameters. The estimation scheme presented in the previous section can be used to estimate first the marginal parameters, then the $\boldsymbol{\Sigma}$ covariance matrix. Phoon et al. [49] proposed an alternative method, which allows the marginal and dependence parameters to be estimated independently:

1. Estimation of Spearman's rank correlations $r_{i,j}$ between pairs of variables. These correlations are invariant by monotonic transformation.
2. Transformation toward Pearson correlation coefficients by $\rho_{i,j} = 2 \sin\left(\frac{\pi}{6} r_{i,j}\right)$
3. Estimation of marginal parameters (Maximum likelihood for instance).

The drawback of this method is that the obtained $\boldsymbol{\Sigma}$ covariance matrix is not ensured to be non-negative definite. If this is not the case, $\boldsymbol{\Sigma}$ can be estimated by the classical observed covariance matrix between transformed data \tilde{X} by $\hat{\boldsymbol{\Sigma}} = \frac{1}{n} \tilde{X}^T \tilde{X}$.

To conclude with this presentation, notice that a Gaussian copula is an easy tool to simulate data with prescribed marginal distributions and correlations: as a first step, a multivariate normal data set is simulated from $\Phi_d(\boldsymbol{\theta}, \boldsymbol{\Sigma})$, where the elements of the diagonal of $\boldsymbol{\Sigma}$ are equal to one. Marginal samples are then transformed by $F_i^{-1}(\phi(\cdot))$.

III. Use of the Gaussian copula in Hydrology

III.1. Field significance

Field significance determination is a problem encountered in multi-testing studies. It has been studied by Livezey and Chen [37], Lettenmaier et al. [36], Douglas et al. [15], and Yue and Wang [63]. In the hydro-meteorological field, a major preoccupation is the determination of

the impacts of climatic change on various variables. The method normally consists in testing a number of stations for stationarity. As an example, suppose that 100 stations are tested with a risk equal to 0.05. Under the hypothesis that all stations are stationary, approximately five stations should be detected as non-stationary, because of the 5% risk. But what would be the conclusion with 6 significant results? In other words, what is the minimum number of locally significant results to conclude, with a risk α , that all these results cannot be due to chance? In order to answer this question, the distribution of N , the number of locally significant tests under the hypothesis that all series are stationary, is needed. This distribution can be calculated theoretically as follows: let $\mathbf{X}^{(i)}$ denote the data recorded at site i , $i=1, \dots, p$. Assume that all series are tested by comparing a test statistic $S_i = g(\mathbf{X}^{(i)})$ with a critical value c . Let Ω_k denote the set of all possible combinations of k elements among p . The distribution of N can then be derived as:

$$\begin{aligned} \Pr(N = k) &= \sum_{\Omega_k} \Pr\left(S_{n_1} > c, \dots, S_{n_k} > c, S_{n_{k+1}} \leq c, \dots, S_{n_p} \leq c\right) \\ &= \sum_{\Omega_k} \Pr\left(g(\mathbf{X}^{(n_1)}) > c, \dots, g(\mathbf{X}^{(n_k)}) > c, g(\mathbf{X}^{(n_{k+1})}) \leq c, \dots, g(\mathbf{X}^{(n_p)}) \leq c\right) \\ &= \sum_{\Omega_k} \Pr\left(\mathbf{X}^{(n_1)} \in g^{-1}(]c; +\infty[), \dots, \mathbf{X}^{(n_k)} \in g^{-1}(]c; +\infty[), \mathbf{X}^{(n_{k+1})} \in g^{-1}(]-\infty; c]), \dots, \mathbf{X}^{(n_p)} \in g^{-1}(]-\infty; c])\right) \\ &= \sum_{\Omega_k} \Pr\left(\mathbf{X} \in \mathfrak{R}_{\Omega_k}\right) \end{aligned} \quad (11)$$

where \mathfrak{R}_{Ω_k} is defined by:

$$\left\{ \mathbf{X} = (\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(p)}) \mid \mathbf{X}^{(n_1)} \in g^{-1}(]c; +\infty[), \dots, \mathbf{X}^{(n_k)} \in g^{-1}(]c; +\infty[), \mathbf{X}^{(n_{k+1})} \in g^{-1}(]-\infty; c]), \dots, \mathbf{X}^{(n_p)} \in g^{-1}(]-\infty; c]) \right\} \quad (12)$$

This equation shows that the multivariate distribution of the tested series is needed, except in the case of independence. In this case:

$$\begin{aligned} \Pr(N = k) &= \sum_{\Omega_k} \Pr\left(S_{n_1} > c, \dots, S_{n_k} > c, S_{n_{k+1}} \leq c, \dots, S_{n_p} \leq c\right) \\ &= \sum_{\Omega_k} \Pr\left(S_{n_1} > c\right) \dots \Pr\left(S_{n_k} > c\right) \Pr\left(S_{n_{k+1}} \leq c\right) \dots \Pr\left(S_{n_p} \leq c\right) \\ &= \sum_{\Omega_k} \alpha^k (1 - \alpha)^{p-k} \\ &= C_p^k \alpha^k (1 - \alpha)^{p-k} \end{aligned} \quad (13)$$

If all sites are independent, N follows a binomial distribution. With the preceding example, the minimum number of locally significant results to ensure a 5% field significance is thus equal to nine under the independence hypothesis, because $p(N \leq 8) = 0.937$ and $p(N \leq 9) = 0.972$.

If the tested series are dependent, a Gaussian copula can be used to take into account spatial correlations in field significance determination, in the following way:

- parameters estimation
- do $i=1, \dots, M$
 - simulation of a new multivariate data set thanks to the estimated copula
 - test of each series for stationarity
 - computation of N_i , the number of significant results

The simulated data series are all stationary, and agree with the observed correlations and marginal distributions. The sample of (N_i) can thus be used to approximate the distribution of N , and to compute a critical value.

This method was applied to a set of 13 hydrometric stations in the North-East of France. Between 39 and 84 years of daily discharges are available. Annual maxima were extracted and a Mann-Kendall test [33, 39] was applied to the 13 series obtained. Five stations showed a significant trend at risk 5% (Figure 3). The preceding procedure was then applied in order to evaluate the field significance of these five significant results. Marginal distributions are assumed to be Gumbel distributions, whose parameters are estimated with the method of moments on all available years. The Σ correlation matrix was directly estimated on transformed data by $\hat{\Sigma} = \frac{1}{n} \tilde{X}^T \tilde{X}$, because the Phoon method based on rank correlations led to a non-definite positive matrix. Only years shared by all stations were taken into account for this estimation (30 years). Estimated correlations ranged from 0.051 to 0.976. The adequacy of the Gaussian copula model was evaluated with the QQ-plot described in section II.3 (Figure 4), which shows no strong departure from the expected distribution. The distribution of N was finally approximated with $M=1000$ simulations, as shown by Figure 5. The decision depends on the combination of a local risk and a regional risk, which are not necessarily equal. Different combinations are summarized in Table 1. As an example, the five significant changes obtained with a 5% test were also regionally significant at a risk of 5%, but were not at risk 1%.

III.2. Regional risk estimation

In this section, the aim will be to estimate the probability of observing at least one event of given return period T , in a set of d stations, as explained in the introduction. The return period corresponding to this probability will be called the regional return period T_R . A similar analysis can be found in the paper by Troutman and Karlinger [62]. Let $X^{(i)}$ denote the annual maximum discharge of station i . If a Gaussian copula is suitable for modeling the multivariate data set, then this probability can be computed as follows:

$$\begin{aligned}
 P(\text{at least one event with return period } T) &= 1 - P(\text{no event with return period } T) \\
 &= 1 - P(\{X^{(1)} \leq q_T^{(1)}\} \cap \dots \cap \{X^{(d)} \leq q_T^{(d)}\}) \\
 &= 1 - F(q_T^{(1)}, \dots, q_T^{(d)}) \\
 &= 1 - c(F_1(q_T^{(1)}), \dots, F_d(q_T^{(d)})) \quad (14) \\
 &= 1 - c(1 - 1/T, \dots, 1 - 1/T) \\
 &= 1 - \Phi_d(\phi^{-1}(1 - 1/T), \dots, \phi^{-1}(1 - 1/T)) \\
 &= 1/T_R
 \end{aligned}$$

This computation is only based on the dependence between stations, and does not depend on marginal estimations. More precisely, if the probability to compute was

$P(\{X^{(1)} \leq x^{(1)}\} \cap \dots \cap \{X^{(d)} \leq x^{(d)}\})$, then the uncertainty would have been greater, as it would include correlations and marginal parameter estimation uncertainties.

The preceding equation was applied to the same data set as in section III.1, leading to the results presented in Figure 6. The 90% confidence interval was obtained by bootstrapping (i.e. years are bootstrapped, and the corresponding data of 13 sites are added to the new data set). As an example, the regional return period of the event “observing at least one event of return period ten years” is approximately two years. In the same manner, the regional return period is approximately 15 years for a local return period of 100 years. It can be noticed that the confidence interval for the regional return period does not encompass the two extreme cases

of independence ($T_R = \frac{1}{1 - (1 - 1/T)^{13}}$) and total dependence ($T_R = T$).

Moreover, from a manager's point of view, an interesting question may be: for the set of 13 stations, what are the discharges to be protected from in order to obtain a regional return period T_R , if it is assumed that all local protection levels are identical? Figure 6 can be used to compute the local return period T , and the marginal estimations lead to the local quantiles with return period T . As an example, with a regional protection level $T_R = 10$ years, the local protection level has to be around 70 years. This difference may appear quite large, but once the return period has been translated in terms of discharges thanks to the marginal estimations, this difference between regional and local risks leads to a difference ranging from 25 to 33% between the discharge with local return period 10 years and the discharge to be protected from in order to obtain a regional return period of 10 years. Finally, the preceding computations were processed by using point-estimates of the parameters. Consequently, sampling uncertainty has to be kept in mind when interpreting such results.

III.3. QdF Analysis

The Discharge-Duration-Frequency (QdF) model has been developed in order to generalize the well known Intensity-duration-Frequency (IdF) model used for precipitations analysis [19, 28-30, 44, 50]. It aims at describing the relationship between quantiles computed for mean discharges $V(d, T)$ over a range of durations. More accurately, the model assumes that quantiles decrease as a hyperbolic function of duration:

$$V(d, T) = \frac{V(0, T) - P}{1 + d / \Delta} + P \quad (15)$$

where P and Δ are parameters which have to be estimated.

One application of the QdF methodology deals with design hydrograph construction. In order to estimate the impact of floods on human activities, hydrologists are asked to provide a hydrograph with a given return period. Of course, this is an almost intractable problem, because a hydrograph can be described in a number of ways, such as peak flow, duration or volume. The QdF approach is used here to compute a Mono-Frequency Synthetic Hydrograph (MFSH), *i.e.* a hydrograph whose mean discharges over a given range of durations have the same return period. The method used to construct the MFSH can be found in Le Clerc [34]. It is based on the study of the hydrograph shapes. When the hypothesis of shape invariance can be accepted, a design hydrograph is computed using the mean of non-dimensional hydrographs, with the discharge divided by peak flow and a synchronization on peak flow. The recession is then corrected in order that each mean discharge over duration d matches with the various quantiles $V(d, T)$ from a QdF analysis. When the various hydrographs have different shapes, Le Clerc [34] proposed dealing with different subsets of hydrographs, related to the various flood origins, *e.g.* snow melt, thunderstorms or frontal rainfall, or various spatial flood extents within the catchment.

Another approach is based on the multivariate analysis of the mean discharges Vd over different durations d . The hydrograph shape invariance implies a high rank correlation between the mean discharges Vd (the event with strongest peak discharge has also the strongest discharges over durations d_1, d_2, \dots). On the other hand, if it assumed that the discharges over different durations are not so dependent, the constructed MFSH could then have a smaller probability than that suggested by the return period used for its construction. A Gaussian copula can once again be used in order to estimate this dependency, and to derive what we will call the "multivariate MSFH return period" T_M .

Such an evaluation is described below by two case studies. The first deals with the Zorn river at Waltenheim (688 km²). Mean discharges were computed over durations $d=1, 2, 3, 4$ and 5 days, and annual maxima were then extracted on 80 years of data. Figure 7a shows the scatterplot matrix of the data: it appears that mean discharges are very strongly correlated over this duration range, which indicates that the shape invariance hypothesis is acceptable.

The Gaussian copula was then estimated with the Phoon method, and its adequacy was checked with the QQ-plot described in section II.3. The multivariate MFSH return period T_M was finally calculated as follows:

$$\begin{aligned}
 \Pr(\text{all mean discharges have a return period greater than } T) &= \Pr(\{X^{(1)} > q_T^{(1)}\} \cap \dots \cap \{X^{(d)} > q_T^{(d)}\}) \\
 &= \Pr(\{F_1(X^{(1)}) > F_1(q_T^{(1)})\} \cap \dots \cap \{F_d(X^{(d)}) > F_d(q_T^{(d)})\}) \\
 &= \Pr(\{\phi^{-1}(F_1(X^{(1)})) > \phi^{-1}(1-1/T)\} \cap \dots \cap \{\phi^{-1}(F_d(X^{(d)})) > \phi^{-1}(1-1/T)\}) \\
 &= \bar{\Phi}_d(\phi^{-1}(1-1/T), \dots, \phi^{-1}(1-1/T)) \\
 &= \bar{c}(1-1/T, \dots, 1-1/T) \\
 &= 1/T_M
 \end{aligned} \tag{16}$$

where $\bar{\Phi}_d(y_1, \dots, y_d) = \Pr(\{Y_1 > y_1\} \cap \dots \cap \{Y_d > y_d\})$ with $\mathbf{Y} = (Y_1, \dots, Y_d) \sim N(0, \Sigma)$

and $\bar{c}(u_1, \dots, u_d) = \bar{\Phi}_d(\phi^{-1}(u_1), \dots, \phi^{-1}(u_d))$

As we are interested in the survival function $\Pr(X_i > x_i \forall i = 1, \dots, d)$, the notations \bar{c} and $\bar{\Phi}_d$ have been introduced in equation (16). However, it should be noticed that there is no simple relationship similar to the equation $\bar{c} = 1 - c$ which holds in the univariate case. Figure 7b illustrates the relationship between the return period used for MFSH construction and the multivariate MFSH return period. It shows that a hydrograph constructed with a 100-years return period has in fact a return period $T_M \approx 160$ years. Conversely, a hydrograph of return period $T_M = 100$ years would be obtained from mean discharges quantiles $V(d, T)$ with a return period T around 65 years. Translated in terms of discharge, this leads to a difference of about 7% between the mean discharge $V(d, T=100 \text{ years})$ and the mean discharge which would lead to a hydrograph return period of 100 years, for the range of durations d considered. The constructed MFSH can thus be considered as suitable for flood scenarios.

The same study was conducted on the Ubaye river at the Lauzet (946 km²), with 45 years of available data. Mean discharges are here computed over durations $d = 1, 10, 20$ and 30 days, because most floods are related to snow melt, leading to very slow events. The data present a weaker dependence than in the preceding case, especially between the daily duration and the others. Consequently, the ratio between marginal and multivariate return periods T and T_M becomes larger: a hydrograph constructed with $T=100$ years has in fact a return period T_M about 1000 years (Figure 8a). This result emphasizes the lack of shape invariance for snow-related events: moderate rainfalls can be superimposed on the high baseflow created by snow melt, thus leading to a great variety of shapes at a daily resolution. By contrast, Figure 8b shows the results of the QdF analysis conducted on the same river, but only for events occurring between September and February (rainfall-related floods), with durations $d=1, 2, 3$ and 4 days. The results are now similar to those obtained with the Zorn River. The construction of a MFSH is thus possible for such flood events.

These two examples provide an illustration of the link between the shape variability of the hydrographs and the dependence between the mean discharges Vd over different durations d . The Gaussian copula provides a quantification of the relationship between the return period T of each mean discharge and the multivariate return period T_M , which can be very different from one catchment to another one. However, the meaning of the multivariate return period T_M is far from obvious, as the sum of the probabilities ($p_1 + p_2$) can be far from one, with $p_1 = \Pr(\text{hydrograph} > \text{MFSH})$ and $p_2 = \Pr(\text{hydrograph} < \text{MFSH})$. A number of intermediate cases can occur, with a cross-over between an observed hydrograph and a MFSH: some mean discharges Vd have a larger return period, and others a smaller, than the prescribed return period used for the MFSH construction. Consequently, such multivariate analysis is more recommended for the study of the hydrograph shape (likelihood of various shape hypotheses) than for an exact assessment of the probability of a design hydrograph.

III.4. Regional Frequency Analysis

Regional frequency analysis is used by hydrologists in order to improve the estimation of high quantiles. The principle is to collect data originating from different locations and to derive a regional distribution of extreme streamflows or rainfalls. Since the precursory work of Dalrymple [11], a number of methodological improvements have been proposed (e.g. [27]). However, a common drawback of almost all regional methods for extremes is that they ignore the spatial dependence of data. This problem has been addressed by Stedinger [58], Hosking and Wallis [26, 27] or Madsen and Rosbjerg [38]. Roughly speaking, these authors found that ignoring intersite dependence led to underestimation of the variance of the estimates, but did not lead to any bias. In the following example, dependence will be explicitly taken into account in the probabilistic model, by means of a Gaussian copula.

Six rainfall stations located around Paris, France, were used (Figure 9). 58 annual maximum values of daily rainfalls were extracted from the years between 1922 and 2003, with 1926, 1936, 1939-1944, 1948-1949, 1952, 1955-1958, 1981-1983, 1991 and 1997-2001 as missing years. The index-flood procedure of Dalrymple [11] comprises an homogenization step, by dividing the at-site samples by the at-site means or medians. For the six study stations, the at-site median values ranged from 28.5 mm to 31.75 mm. Moreover, from a meteorological point of view, the study area can be considered as homogeneous, with little altitudinal range. Consequently, we assume that the six rainfall series arise from an identical GEV distribution. The parameters of this regional distribution are estimated using two models:

In model 1 the intersite dependence is ignored. The multivariate density of a vector of six annual maxima $(x_t^{(1)}, \dots, x_t^{(6)})$ at year t can be written as:

$$f_1(x_t^{(1)}, \dots, x_t^{(6)}) = \prod_{i=1}^6 GEV(x_t^{(i)}; \alpha, \beta, \xi) \quad (17)$$

In model 2 the intersite dependence is modeled with a Gaussian copula, whose adequacy was checked with the Chi-square plot of section II.3. Using Equation (8), the multivariate density of a vector of six annual maxima $(x_t^{(1)}, \dots, x_t^{(6)})$ at year t can be written as:

$$f_2(x_t^{(1)}, \dots, x_t^{(6)}; \alpha, \beta, \xi, \Sigma) = \left(\prod_{i=1}^6 GEV(x_t^{(i)}; \alpha, \beta, \xi) \right) |\Sigma|^{-1/2} \times \exp \left\{ - \frac{[\mathbf{V}(x_t^{(1)}, \dots, x_t^{(6)}; \alpha, \beta, \xi)] [\Sigma^{-1} - \mathbf{I}] [\mathbf{V}(x_t^{(1)}, \dots, x_t^{(6)}; \alpha, \beta, \xi)]^T}{2} \right\}, \quad (18)$$

$$\text{with } \mathbf{V}(x_t^{(1)}, \dots, x_t^{(6)}; \alpha, \beta, \xi) = \left[\phi^{-1}(F(x_t^{(1)}; \alpha, \beta, \xi)), \dots, \phi^{-1}(F(x_t^{(6)}; \alpha, \beta, \xi)) \right]$$

and $F(x; \alpha, \beta, \xi)$ the cdf of the GEV distribution.

The dependence is thus summarized in a correlation matrix Σ . In order to obtain an acceptable number of parameters, the dependence $\rho_{i,j}$ between two stations is assumed to decrease as a function of the distance $d_{i,j}$:

$$\begin{cases} \rho_{i,j} = \gamma_0 \exp(-\gamma_1 d_{i,j}) & \text{if } i \neq j \\ \rho_{i,i} = 1 \end{cases} \quad (19)$$

Such a parameterization can be compared to classical variograms used in geostatistics. Because of the complexity of this model, a Bayesian estimation scheme using MCMC algorithms is applied. Some details of these methods in a hydrological context can be found for instance in the papers by Perreault et al. [47, 48], Thyer et al. [61], Marshall et al. [41] and

Renard et al. [53]. The likelihood of the data is simply computed from the product of the multivariate densities of observations, *i.e.* for one of the models $k=1,2$ described above:

$$p_k(\mathbf{X} | \boldsymbol{\theta}_k) = \prod_{t=1}^n f_k(x_t^{(1)}, \dots, x_t^{(6)}; \boldsymbol{\theta}_k) \quad (20)$$

with $\boldsymbol{\theta}_1 = (\alpha, \beta, \xi)$ and $\boldsymbol{\theta}_2 = (\alpha, \beta, \xi, \boldsymbol{\Sigma}) = (\alpha, \beta, \xi, \gamma_0, \gamma_1)$.

The prior distribution of the shape parameter ξ is chosen to be a Gaussian distribution with zero mean and standard deviation 0.3, which implies that the interval $[-0.6;0.6]$ encompasses more than 95% of the density. This prior distribution can be compared with the Martins and Stedinger [42] geophysical prior, which is less variable and is entirely included in the interval $[-0.5;0.5]$. For other parameters, almost non-informative priors are set, by using uniform distributions with large variances, namely $\alpha \sim U[0,1000]$, $\beta \sim U[-10000,10000]$,

$\gamma_0 \sim U[0,1]$ and $\gamma_1 \sim U[0,10]$. Finally, the independence between these prior marginal

distributions is assumed, in order to derive the multivariate prior distribution $\pi_k(\boldsymbol{\theta}_k)$ of the parameters of the model k .

The posterior distribution is finally obtained up to a constant of proportionality by:

$$p_k(\boldsymbol{\theta}_k | \mathbf{X}) \propto p_k(\mathbf{X} | \boldsymbol{\theta}_k) \pi_k(\boldsymbol{\theta}_k) \quad (21)$$

Samples arising from the posterior distribution of the parameters are generated with MCMC algorithms in the following way. First of all, at-site maximum likelihood estimations are used to derive a rough estimate of the posterior means and variances of the parameters α , β and ξ .

Because almost non-informative priors are used, the posterior distribution will be mostly influenced by the likelihood, thus making the ML-estimates relevant. For the two remaining parameters of model 2, the exponential model described in equation (19) is fitted to the empirical correlations estimated by the method of Phoon, using a classical least-square approach. As a second step, these estimations are used as starting parameters of a Metropolis algorithm (see Renard et al. [53] for a detailed description), which was run for 100 000 iterations, and whose convergence was checked using the approach suggested by Gelman et al. [20]. Finally, the last 50 000 iterations are used to perform the inference.

Figure 10 describes the relationship between distance and dependence. The solid line represents the median exponential decrease described in Equation (19), and the dashed lines denote a 90% posterior confidence interval. The model of dependence/distance relationship seems acceptable for the six series studied. Figure 11 compares the posterior distributions of parameters α, β and ξ obtained with the two models. For the first two parameters, the results are almost identical for the mean posterior values, but the estimates of model 2, which takes into account the intersite dependence, have a larger variance. This is consistent with the conclusions of Stedinger [58], Hosking and Wallis [26, 27] and Madsen and Rosbjerg [38] described earlier. Conversely, the posterior variances of the shape parameter are almost identical for the two models, with a slight shift between the two distributions. This result shows that the effect of intersite dependence is not necessarily the same for all parameters. The implications for quantiles estimates are shown in Figure 12: the difference between the two models, in terms of posterior variance, is stronger between moderate quantiles (*i.e.* 10-years return period) than between high quantiles (*i.e.* 100-years return period).

These results also have implications for regional methods of extrapolation. As an illustration, the FORGEX method [16, 51, 59] uses the series of the annual maximum of the standardized values observed over a measurement network (netmax series) in order to extrapolate the distribution of a target site. This method uses the following observation [52]: if the N stations of the network are independent, then the netmax variable Y has the following distribution:

$$\begin{aligned}
F_Y(y) &= \Pr(Y \leq y) \\
&= \Pr\left(\text{Max}(X_1, \dots, X_N) \leq y\right) \\
&= \Pr\left(\{X_1 \leq y\} \cap \dots \cap \{X_N \leq y\}\right) \\
&= \prod_{i=1}^N \Pr(X_i \leq y) \\
&= \left(F_X(y)\right)^N
\end{aligned} \tag{22}$$

In other terms:

$$\begin{aligned}
-\log(-\log(F_Y(y))) &= -\log\left(-\log\left(\left(F_X(y)\right)^N\right)\right) \\
&= -\log(N) - \log(-\log(F_X(y)))
\end{aligned} \tag{23}$$

so that in a Gumbel repair, the distribution of the netmax lies at a distance of $\log(N)$ to the left of the population growth curve (*i.e.* the distribution of the standardized values). In order to take into account the effect of intersite dependence, this distance is in fact replaced by $\log(N_e)$, where N_e is an equivalent number of stations. The latter concept has been used for a long time [43] in order to take into account the information redundancy caused by dependence. However, in the present situation, this approach suffers from two drawbacks, whatever the method used to estimate the equivalent number of stations. First of all, the preceding results show that the effect of intersite dependence is not identical for all parameters, or equivalently for the whole quantile curve. Consequently, using a unique value as N_e to summarize the effect of dependence is questionable. Secondly, the parallelism between at-site and netmax growth curves is only true under the hypothesis of independence. As an illustration, a simple bidimensional case is explored: for a network of two sites, whose dependence is described with a Gaussian copula with correlation 0.8, the at-site distribution (a GEV(1,0,-0.5)) is plotted in a Gumbel repair, together with the netmax distribution under the independence hypothesis and the “true” Gaussian copula hypothesis. Figure 13 (left panel) shows that the at-site and the netmax distributions are parallel assuming independence, but not with the prescribed model of dependence. In the right panel, the gap between these two distributions is shown: in the independence case, the gap is constant and equal to $\log(2)$, but this not the case in the Gaussian copula case, where this gap depends on marginal values, and can therefore not be equal to $\log(N_e)$, whatever the N_e used. The same computation is made for the preceding case study (Figure 14), with identical conclusions. In this case, the distributions are plotted using point-estimates of the parameters. Consequently, they are affected by sampling uncertainty, which is not shown in the figure for clarity, but which is far from negligible, thus complicating the conclusions of the comparison.

The case study could be enhanced in several ways. First of all, the assumption of a single shared distribution may be relaxed. Rather than dividing the at-site samples by an index variable, the spatial variability of the parameters could be modeled with covariables such as altitude or distance from the sea. Such an approach has been proposed by Diggle et al. [14], Cooley [9] and Cooley et al. [10]. Moreover, from a Bayesian perspective, the use of a proper prior distribution could improve the accuracy of estimates. Bayesian model checking or Bayesian comparison of models could also be useful for deriving a more complete validation of the model assumptions. Finally, it is clear that the preceding conclusions are dependent on the Gaussian copula hypothesis, which remains questionable. Further investigations are needed to refine regional frequency analysis, using more complex dependence structures.

III.5. The Gaussian copula can fail

The need to check the adequacy of the copula before using it for any computation has been pointed out in section II. The following case study illustrates the fact that a Gaussian copula is far from being a universal tool. 55 flood events were selected from the daily discharge series of the Ubaye river at Barcelonette (549 km²; 1904-2001). The scatterplot of the peak discharge of the hydrograph versus the volume computed above a threshold equal to the half of the peak discharge is shown in Figure 15a. At first sight, the dependence structure of the data seems complex. Nevertheless, the shape of the scatterplot may be due to the marginal structure of the observations: on this river, flood events can result from snow melt or heavy rainfalls. Both marginal distributions are therefore likely to derive from a population mixture model. In order to erase the influence of marginal distributions, a normal score transformation is applied to the data, leading to the scatterplot shown in Figure 15c. Snow-related events are denoted by crosses, and rainfall-related ones by triangles. Despite this transformation, the two types of events still lead to different dependence structures: for snow-related events, peak flows and volumes are more correlated than for rainfall-related ones. It is thus clear that modeling the dependence structure with a single correlation parameter will lead to very poor results. As an illustration, the contour of the estimated bivariate Gaussian density is also plotted in the Figure 15c. It clearly appears that the distribution of the points is not consistent with a Gaussian description. This departure from normality can also be viewed in the Chi-square plot of Figure 15d, where high quantiles have a tendency to deviate from the theoretical line $y=x$.

A more subtle problem arises from the asymptotic properties of the copula, also known as the tail dependence properties. More accurately, the Gaussian copula implies the asymptotic independence of marginal values, which means that $\Pr(F_X(X) > p \mid F_Y(Y) > p) \xrightarrow{p \rightarrow 1} 0$. In other words, the dependence weakens for very extreme events. This property has a strong influence for the computation of very low probabilities, concerning events lying outside the range of observations. More explicitly, the risk can be strongly underestimated if an asymptotically independent model is used with asymptotically dependent data. For this reason, empirical or physical evidence of tail independence is necessary before using such a copula for extrapolation. Coles et al. [7] provide a number of tools for exploring the asymptotic properties of a multivariate data set. This phenomenon will be illustrated with a simulated case study. The following variables are considered:

$$\begin{aligned} X &\sim N(0;1) \\ Y &= X + Z, \text{ where } Z \mid X \sim N\left(0; \frac{\exp(-X^2/4)}{4}\right). \end{aligned} \quad (24)$$

The idea behind this construction is to create a couple of variables whose dependence increases with marginal values. A sample of (X, Y) with size 10^6 was thus generated. In order to deal with known marginal distributions, data were finally transformed to fit a GEV distribution with parameters $(1, 0, -0.5)$:

$$(\tilde{X}, \tilde{Y}) = \left(G^{-1} \{ \hat{F}_X(X) \}, G^{-1} \{ \hat{F}_Y(Y) \} \right) \quad (25)$$

where G is the cdf of the GEV distribution and \hat{F}_X is the empirical cdf of X .

The first 100 couples were used for the copula estimation. The scatterplot of these data is shown in Figure 16a. In order to focus on the problems caused by the dependence structure, the marginal parameters were not estimated but fixed at the true value $(1, 0, -0.5)$. Figure 16b shows the scatterplot of the normalized data, and the contour of the bivariate density

estimated with the Phoon method. At first sight, the Gaussian copula seems to be a reasonable description of the dependence of the data.

The Gaussian copula is now used to extrapolate far outside the observations range. More accurately, the probability $p = \Pr(\tilde{X} > u, \tilde{Y} > u)$ is computed, with u possibly lying outside the observations range. This probability is estimated with the Gaussian copula as

$p = \bar{\Phi}_d\left(\phi^{-1}(G(u)), \phi^{-1}(G(u))\right)$, and is compared to the theoretical value, empirically estimated with the total sample of length 10^6 . The results are shown in Figure 17, where the probability described above is expressed in terms of return period ($T=1/p$). On the left-hand side, it can be observed that the Gaussian copula is adequate in the observation range. Conversely, the quantile curves differ strongly in the extrapolation range: the curve associated with the Gaussian copula grows faster than the theoretical one, which leads to a severe underestimation of risk. As an example, the quantile is under-estimated by about 38% for a return period of 10^3 years, and by about 58% for a return period of 10^4 years. This can be explained by the fact that the simulated data are not asymptotically independent. The probability of observing an extreme value in both components is thus higher than that suggested by the Gaussian copula model.

IV. Discussion

The case studies presented in sections III.1 to III.4 demonstrate the usefulness and the relative simplicity of the Gaussian copula for dealing with multivariate events. Conversely, section III.5 highlights its limitations and emphasizes the error risk related to extrapolation. As no theoretical argument can be employed to justify the use of this copula, this raises the question of how to decide, in practical cases, if this model is acceptable or should be avoided. Two different aspects have to be explored for this purpose. The first is to check that the Gaussian copula is consistent with the observed data. This implies that both the marginal distributions and the dependence structure have to be in agreement with the observations. For the marginal distributions, a number of standard univariate tools can be used (e.g. goodness-of-fit tests or QQ-plots). Moreover, univariate extreme value theory provides a strong theoretical justification for the use of extreme value distributions. After marginal transformations, the distribution is determined by a simple multivariate Gaussian model. Diagnoses for multivariate normality have also been extensively studied, thus providing a number of tools for model checking [40].

The second aspect is more difficult to handle, and is related to the asymptotic dependence properties of the data. From a theoretical point of view, it can be shown that the limit distribution of the componentwise maxima (Equation (2)) has the property of asymptotic dependence, except in the case of exact independence. Consequently, the Gaussian copula might be considered at first sight to be inadequate for describing these data, because it is asymptotically independent. Alternative multivariate distributions may be used [8], e.g. using the logistic family. Such a development has been undertaken by Tawn [60] and Schalter and Tawn [55] for an arbitrary number of dimensions, and may be applied with additional hypotheses in order to reduce the number of parameters (2^d-1). However, this limitation has to be mitigated, because the limit distribution of Equation (2) is valid when the block size tends toward infinity. In most hydrological applications, the block size is typically of 365 days, with non-independent daily data, and a number of zero values in the case of rainfall. It is thus possible that the asymptotic arguments used to derive the limit distribution of componentwise maxima do not hold in real-life hydrological studies. Although the latter argument can also be argued for univariate extreme value distributions, it seems that the problem is more critical in

the multidimensional case. For instance, Bortot et al. [3] reported empirical evidence of asymptotically independent data in the oceanographic field. Alternatively, Hosking and Wallis [26] argued that a Gaussian copula model was fairly well supported by British annual flood series, although in some cases, stronger dependence seemed to occur at higher levels. From a practical point of view, a given data set of componentwise maxima will generally not include enough data to provide empirical evidence of asymptotic dependence or independence. Physical arguments may be used to discriminate between these two hypotheses, by using for example meteorological model simulations, or by exploiting the knowledge of historical extreme events.

Consequently, our point of view is that the Gaussian copula should not be used for the computation of very low probabilities, unless strong arguments can be derived to justify the asymptotic properties of the data. Conversely, the Gaussian copula can be very useful within the observations range. As an illustration, in the case study in section III.4, the copula can be used to take into account intersite dependence in regional frequency analysis, and to derive a more realistic quantification of the uncertainties than with the independence hypothesis, which is in any case a very crude model for describing dependent data.

V. Conclusion and perspectives

The aim of this article was to present some possible applications of the Gaussian copula in the flood mitigation field. The main advantage of this method is its simplicity, even with more than two or three dimensions. Rough estimates can thus be obtained in problems involving an evaluation of the dependence of variables. Moreover, the Gaussian copula can be used to simulate values with prescribed correlations and marginal distributions. Four case studies were used to demonstrate its usefulness in the contexts of field significance determination, regional risk analysis, QdF models with design hydrograph derivation and regional frequency analysis. Nevertheless, this tool is far from universal, as it is based on no theoretical justification. The suitability of the Gaussian dependence model therefore has to be properly checked. Moreover, even if the copula seems well fitted to observed values, caution is needed before using it for extrapolation.

Alternative copula families can be used in order to improve the dependence model arising from the Gaussian copula. For instance, the Student copula may be more appropriate if tail dependence is observed in the data. Nevertheless, the choice of the more suitable copula for modeling a data set is not straightforward. Various criteria may be used for this purpose, but they will only reflect a particular feature of the data. For instance, two models may lead to an adequate fit to the data on the basis of a given criterion, but to contrasted results in extrapolation. The choice of a copula in such a case is problematic, especially if the data set is not informative enough to provide relevant indications about the asymptotic dependence properties. Modeling uncertainty can thus be an important part of the overall uncertainty. Another problem concerns the parameter estimation uncertainties, because the maximum likelihood method can usually not be used, due to the great number of parameters to be estimated. Computations based on the Fisher information matrix are thus impossible. Finally, whatever the copula used, the lack of theoretical justification remains a limitation for extrapolation.

The preceding approaches only consist in modeling the extreme part of the variables distribution. An alternative consists in studying the whole distribution, and evaluating how dependence evolves for high values. For instance, Bortot et al. [3] applied a model where the tail dependence is described using a multivariate Gaussian distribution, after suitable marginal transformations. In spirit, this model is similar to that of the Gaussian copula. The major difference is that the data are not block maxima, and that the Gaussian tail model is only

applied above a multivariate threshold (u_1, \dots, u_d) , which has to be high enough to ensure that the asymptotic assumptions hold, both for marginal and joint distributions. Moreover, given that the distribution belongs to the domain of attraction of a multivariate extreme value distribution, some theoretical results related to the theory of regular variation [2] can be used in order to compute the probability of extreme sets. An example of such study is provided by De Haan and De Ronde [12]. Such an approach is appealing, because it is physically more convincing: couples of values are describing the same physical event, while componentwise maxima lead to the use of couples of values which may describe two distinct events. Nevertheless, some difficulties still hold. Firstly, obtaining a sample of independent values is not obvious. As an example, if the data consist of daily discharges at two locations, a sampling strategy has to be found which ensures the independence of successive values and preserve enough data for estimation. Secondly, this approach can only be used to compute the probability of sets for which both components are extreme. This can be a strong limitation: in the confluence problem, as an example, an extreme downstream discharge can result from a high value in only one of the two upstream flows. Heffernan and Tawn [24] recently proposed a semi-parametric conditional approach which overcomes this difficulty. More generally, we believe that multivariate extreme events analysis is likely to improve hydrological risks assessment, as emphasized by Katz et al. [31]. In a number of situations, multivariate events are used without a formal statistical model to account for dependence. As an example, the well-known index flood methodology [11] is intended to improve quantiles estimations by using data from several sites. Unfortunately, among other problems, ignoring spatial dependence leads to an underestimate of the quantiles uncertainty. Alternatively, most hydrological extreme events (floods or droughts) are intrinsically multivariate, as noted by Adamson et al. [1]. Potential damage is thus likely to be a function of several random variables. Hydrologists involved in risk assessment should thus be attentive with the progress achieved by statisticians in multivariate extremes theory.

Acknowledgements

The financial support provided by Cemagref and EDF for the PhD research of B. Renard is gratefully acknowledged. Discharge data series are derived from the national HYDRO database from the French Ministry of Environment, and the rainfall series have been reviewed by Météo-France within the European IMFREX project. Both organisms are acknowledged. We also thank two anonymous reviewers for their helpful comments.

VI. Bibliography

- [1] Adamson PT, Metcalfe AV, Parmentier B. Bivariate extreme value distributions: an application of the Gibbs sampler to the analysis of floods. *Water Resources Research* 1999; 35:2825-2832.
- [2] Bingham NH, Goldie CM, Teugels JL. *Regular variation*. New York: Cambridge University Press, 1987.
- [3] Bortot P, Coles S, Tawn JA. The multivariate Gaussian tail model: an application to oceanographic data. *Appl. Stat.* 2000; 49:31-49.
- [4] Caperaa P, Fougères AL, Genest C. Bivariate distributions with given extreme value attractor. *J. Multivar. Anal.* 2000; 72:30-49.
- [5] Cherubini U, Luciano E, Vecchiato W. *Copula Methods in Finance*: Wiley, 2004.
- [6] Coles S. *An Introduction to Statistical Modeling of Extreme Values*. London: Springer-Verlag, 2001.
- [7] Coles S, Heffernan JE, Tawn JA. Dependence measures for extreme value analyses. *Extremes* 1999; 2:339-365.
- [8] Coles S, Tawn JA. Modelling extreme multivariate events. *J. R. Stat. Soc. Ser. B-Methodol.* 1991; 53:377-392.
- [9] Cooley D. *Statistical Analysis of Extremes Motivated by Weather and Climate Studies: Applied and Theoretical Advances*, PhD thesis, University of Colorado, 2005
- [10] Cooley D, Nychka D, Naveau P. A Spatial Bayesian Hierarchical Model for a Precipitation Return Levels Map. In: *Extreme Value Analysis*. Gothenburg, Sweden, 2005.
- [11] Dalrymple T. Flood frequency analyses. In: *Water-supply paper 1543-A: US Geological Survey*, 1960.
- [12] De Haan L, De Ronde J. Sea and Wind: Multivariate Extremes at Work. *Extremes* 1998; 1:7-45.
- [13] De Michele C, Salvadori G, Canossi M, Petaccia A, Rosso R. Bivariate statistical approach to check adequacy of dam spillway. *J. Hydrol. Eng.* 2005; 10:50-57.
- [14] Diggle PJ, Tawn JA, Moyeed RA. Model-based geostatistics. *J. R. Stat. Soc. Ser. C- Appl. Stat.* 1998; 47:299-326.
- [15] Douglas EM, Vogel RM, Kroll CN. Trends in floods and low flows in the United States: impact of spatial correlation. *J. Hydrol.* 2000; 240:90-105.
- [16] Faulkner DS, Jones DA. The FORGEX method of rainfall growth estimation - III: Examples and confidence intervals. *Hydrol. Earth Syst. Sci.* 1999; 3:205-212.
- [17] Favre AC, El Adlouni S, Perreault L, Thiémonge N, Bobee B. Multivariate hydrological frequency analysis using copulas. *Water Resources Research* 2004; 40.
- [18] Fisher RA, Tippett LH. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Cambridge Phil. Soc.* 1928; 24.
- [19] Galéa G, Prudhomme C. Notions de base et concepts utiles pour la compréhension de la modélisation synthétique des régimes de crue des bassins versants au sens des modèles QdF. *Rev. Sci. Eau* 1997; 1:83-101.
- [20] Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian data analysis*: Chapman & Hall, 1995.
- [21] Genest C, McKay J. The joy of copulas: bivariate distributions with uniform marginals. *The American Statistician* 1986; 40:280-283.

- [22] Genest C, Rivest L-P. A characterization of Gumbel's family of extreme value distributions. *Statistics and Probability Letters* 1989; 8:207-211.
- [23] Grimaldi S, Serinaldi F. Asymmetric copula in multivariate flood frequency analysis. *Adv. Water Resour.* 2006; 29:1155-1167.
- [24] Heffernan JE, Tawn JA. A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society* 2004; 66:497-546.
- [25] Herr HD, Krzysztofowicz R. Generic probability distribution of rainfall in space: the bivariate model. *J. Hydrol.* 2005; 306:234-263.
- [26] Hosking JRM, Wallis JR. The effect of intersite dependence on regional flood frequency analysis. *Water Resources Research* 1988; 24:588-600.
- [27] Hosking JRM, Wallis JR. *Regional Frequency Analysis: an approach based on L-Moments.* Cambridge, UK: Cambridge University Press, 1997.
- [28] Javelle P. Caractérisation du régime des crues: le modèle débit-durée-fréquence convergent. Approche locale et régionale, PhD thesis, INP Grenoble, Cemagref Lyon, 2001
- [29] Javelle P, Grésillon JM, Galéa G. Discharge-duration-Frequency curves modeling for floods and scale invariance. *Comptes Rendus de l'Académie des Sciences, Sciences de la terre et des planètes.* 1999; 329:39-44.
- [30] Javelle P, Ouarda T, Lang M, Bobee B, Galéa G, Grésillon JM. Development of regional flood-duration-frequency curves based on the index-flood method. *J. Hydrol.* 2002; 258:249-259.
- [31] Katz RW, Parlange MB, Naveau P. Statistics of extremes in hydrology. *Adv. Water Resour.* 2002; 25:1287-1304.
- [32] Kelly KS, Krzysztofowicz R. A bivariate meta-Gaussian density for use in hydrology. *Stoch. Hydrol. Hydraul.* 1997; 11:17-31.
- [33] Kendall MG. *Rank correlation methods.* London: Griffin, 1975.
- [34] Le Clerc S. Revisiter la notion de scénario hydrologique de référence pour la caractérisation des inondations, PhD thesis, Université Joseph Fourier Grenoble, Cemagref Lyon, 2004
- [35] Le Clerc S, Lang M. Flood frequency analysis downstream confluences. Comparison between bivariate densities and experimental data. In: *International conference on flood estimation.* Berne, Switzerland, 2002; 295-304.
- [36] Lettenmaier DP, Wood EF, Wallis JR. Hydro-Climatological Trends in the Continental United-States, 1948-88. *J. Climate* 1994; 7:586-607.
- [37] Livezey RE, Chen WY. Statistical field significance and its determination by Monte Carlo techniques. *Monthly Weather Review* 1983; 111:46-59.
- [38] Madsen H, Rosbjerg D. The partial duration series method in regional index-flood modeling. *Water Resources Research* 1997; 33:737-746.
- [39] Mann HB. Nonparametric tests against trend. *Econometrica* 1945; 13:245-259.
- [40] Mardia KV. Tests of univariate and multivariate normality. In: Krishnaiah PR, ed. *Handbook of Statistics 1: analysis of variance.* Amsterdam, Holland, 1980; 279-320.
- [41] Marshall L, Nott D, Sharma A. A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling. *Water Resources Research* 2004; 40.
- [42] Martins ES, Stedinger JR. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research* 2000; 36:737-744.
- [43] Matalas NC, Langbein WB. Information content on the mean. *Journal of Geophysical Research* 1962; 67:3441-3448.
- [44] Meunier M. Regional flow-duration-frequency model for tropical island of Martinique. *J. Hydrol.* 2001; 247:31-53.

- [45] Mikosch T. How to model multivariate extremes if one must? *Stat. Neerl.* 2005; 59:324-338.
- [46] Mousavi NS. Composition des lois élémentaires en hydrologie régionale : application à l'étude des régimes de crue, PhD thesis, Université Joseph Fourier Grenoble, Cemagref Lyon, 1997
- [47] Perreault L, Bernier J, Bobee B, Parent E. Bayesian change-point analysis in hydrometeorological time series. Part 1. The normal model revisited. *J. Hydrol.* 2000; 235:221-241.
- [48] Perreault L, Bernier J, Bobee B, Parent E. Bayesian change-point analysis in hydrometeorological time series. Part 2. Comparison of change-point models and forecasting. *J. Hydrol.* 2000; 235:242-263.
- [49] Phoon KK, Quek ST, Huang HW. Simulation of non-Gaussian processes using fractile correlation. *Probab. Eng. Eng. Mech.* 2004; 19:287-292.
- [50] Prudhomme C. Modèles synthétiques des connaissances en hydrologie, PhD thesis, Université Montpellier II, Cemagref Lyon, 1995
- [51] Reed DW, Faulkner DS, Stewart EJ. The FORGEX method of rainfall growth estimation - II: Description. *Hydrol. Earth Syst. Sci.* 1999; 3:197-203.
- [52] Reed DW, Stewart EJ. Inter-site and inter-duration dependence in rainfall extremes. In: Turkman VBaKF, ed. *Statistics for the Environment 2 : Water related issues.* Chichester, UK: Wiley, 1994; 125-143.
- [53] Renard B, Garreta V, Lang M. An application of Bayesian analysis and MCMC methods to the estimation of a regional trend in annual maxima. submitted to *Water Resources Research.* 2006.
- [54] Salvadori G, De Michele C. Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. *Water Resources Research* 2004; 40.
- [55] Schalther M, Tawn JA. A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika* 2003; 90:139-156.
- [56] Sklar A. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* 1959; 8:229-231.
- [57] Song P XK. Multivariate dispersion models generated from Gaussian copula. *Scand. J. Stat.* 2000; 27:305-320.
- [58] Stedinger JR. Estimating a regional flood frequency distribution. *Water Resources Research* 1983; 19:503-510.
- [59] Stewart EJ, Reed DW, Faulkner DS, Reynard NS. The FORGEX method of rainfall growth estimation - I: Review of requirement. *Hydrol. Earth Syst. Sci.* 1999; 3:187-195.
- [60] Tawn JA. Modelling multivariate extreme value distributions. *Biometrika* 1990; 77:245-253.
- [61] Thyer M, Kuczera G, Wang QJ. Quantifying parameter uncertainty in stochastic models using the Box-Cox transformation. *J. Hydrol.* 2002; 265:246-257.
- [62] Troutman BM, Karlinger MR. Regional flood probabilities. *Water Resources Research* 2003; 39.
- [63] Yue S, Wang CY. Regional streamflow trend detection with consideration of both temporal and spatial correlation. *Int. J. Climatol.* 2002; 22:933-946.

LIST OF CAPTIONS

TABLES

Table 1: Comparison of observed and critical number of significant tests for various local and regional risks. A star denotes a regionally significant result.

FIGURES

Figure 1. Principle of the Gaussian copula

Figure 2. Isolines of a bivariate distribution arising from a Gaussian copula with marginal distributions $GEV(1, 0, -0.2)$ and correlation coefficient equal to 0.8.

Figure 3. Annual maxima of 13 stations in the north-east of France. Stations with significant trend at risk 5% are denoted with a star in their figure box.

Figure 4. Chi-square plot (see text for explanation) for the 13 stations studied.

Figure 5. Estimated distribution of N , the number of locally significant results, under the assumption that all series are stationary.

Figure 6. Local versus regional return period (see text for definition). 90% confidence intervals are obtained by Bootstrap.

Figure 7. QdF Analysis of the Zorn River. (a) Scatterplot matrix of annual maxima, (b) return period of the mean discharges Vd versus multivariate hydrograph return period. Total dependence and independence cases are respectively denoted by dash and dash-dot lines.

Figure 8. QdF Analysis of the Ubaye River at Lauzet. (a) all events; (b) events occurring between September and February. Total dependence and independence cases are respectively denoted by dash and dash-dot lines.

Figure 9. Location of the six rainfall stations.

Figure 10. Relationship between intersite distance and intersite dependence. Crosses denote Phoon's estimates of correlations, the median line with 90% confidence intervals are obtained by means of the posterior distribution of parameters.

Figure 11. Posterior marginal distributions of scale, location and shape parameters. Dashed line: independence hypothesis, solid line: Gaussian copula.

Figure 12: Posterior distribution of quantiles with probabilities 0.9 and 0.99. Dashed line: independence hypothesis, solid line: Gaussian copula.

Figure 13. Bidimensional case: Comparison between at-site distribution (triangles) and netmax distribution, in the cases of independence (crosses) and with a Gaussian copula (circles). Left: distributions in a Gumbel repair, right: difference with at-site distribution.

Figure 14. Network of six rainfall series. Comparison between at-site distribution (triangles) and netmax distribution, in the cases of independence (crosses) and with a Gaussian copula (circles).

Figure 15. (a) Scatterplot of peak daily discharge versus volume of selected hydrographs on the Ubaye River at Barcelonnette. (b) Scatterplot of peak discharge versus volume, transformed

by the estimated empirical cdf. (c) Normalized scatterplot, with estimated Gaussian distribution. Snow-related events are denoted by crosses, and rainfall-related ones by triangles. (d) Chi-square plot.

Figure 16. Scatterplot of the first 100 simulated values. (a): raw values, (b): normalized values and isolines of the estimated bivariate Gaussian density.

Figure 17. Return period associated with the probability $P(\tilde{X} > u, \tilde{Y} > u)$. Solid line represents the theoretical probability, and dotted line the value estimated with the Gaussian copula. (a) u is in the observations range, (b) extrapolation.

Local α	Observed number of significant tests	Critical number of significant tests at regional risk		
		1%	5%	10%
0.01	3	3 *	2 *	2 *
0.05	5	6	4 *	3 *
0.1	5	8	5 *	4 *

Table 1

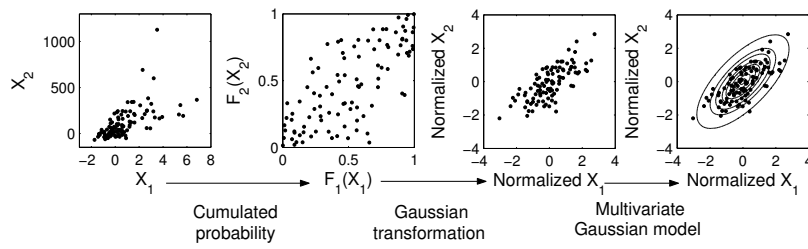


Figure 1

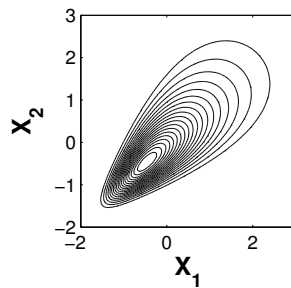


Figure 2

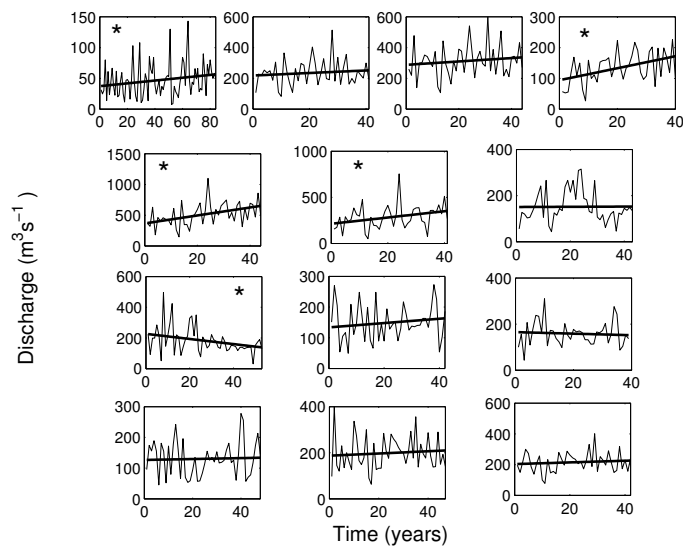


Figure 3

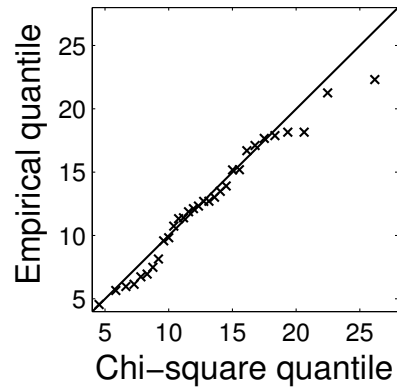


Figure 4

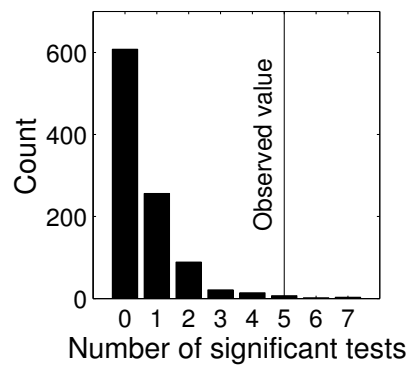


Figure 5

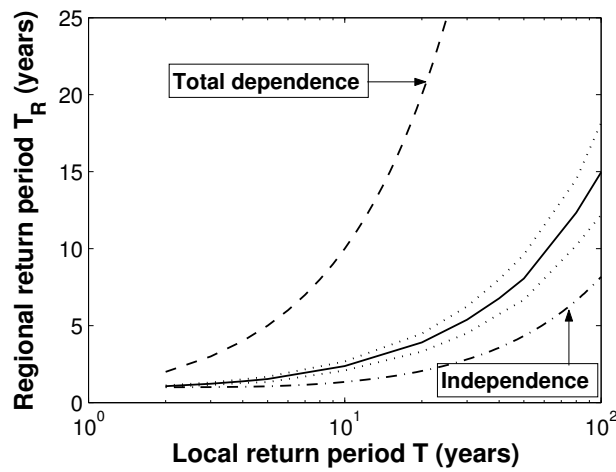


Figure 6

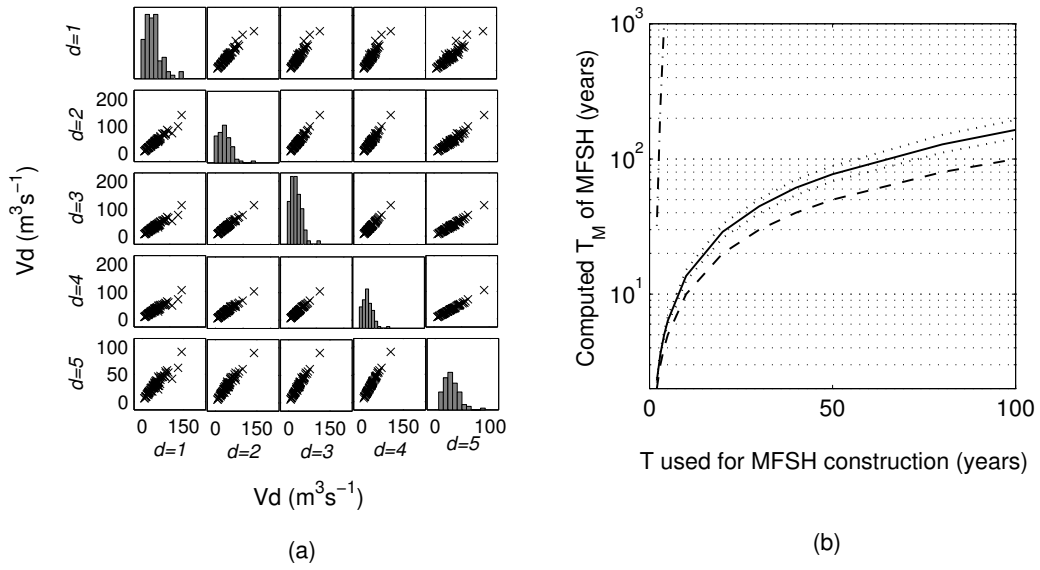


Figure 7

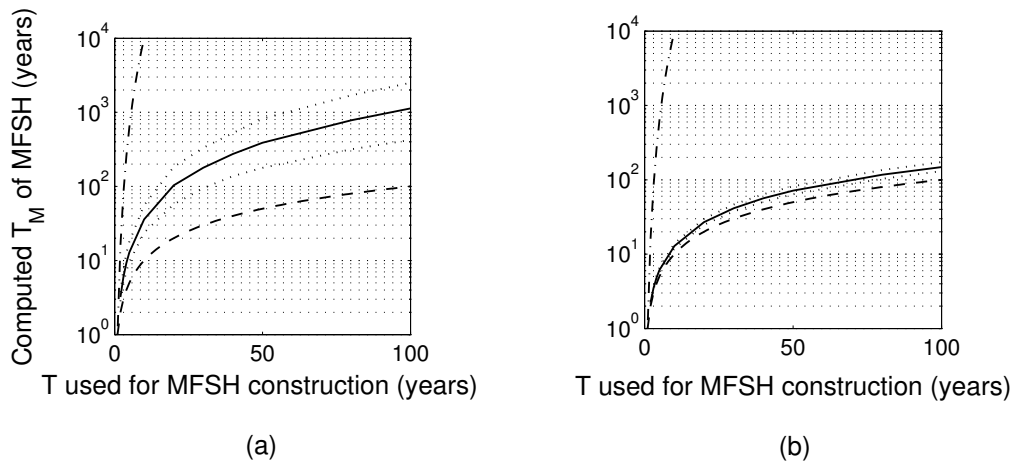


Figure 8

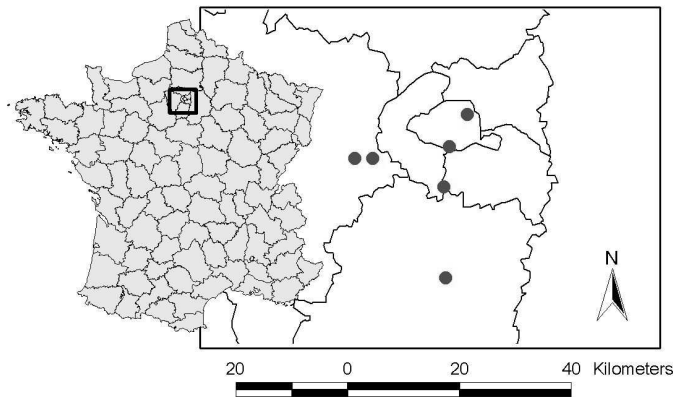


Figure 9

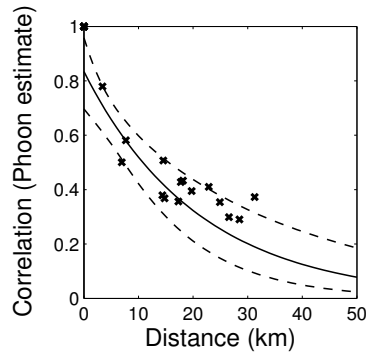


Figure 10

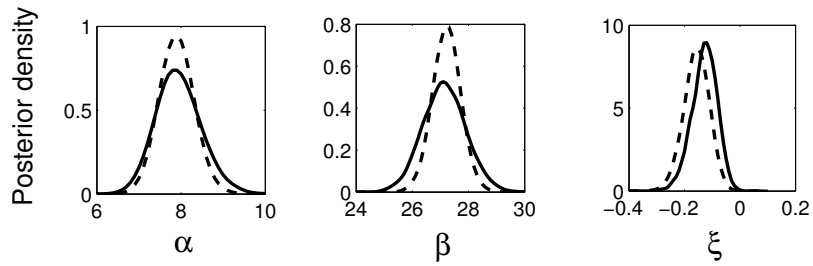


Figure 11

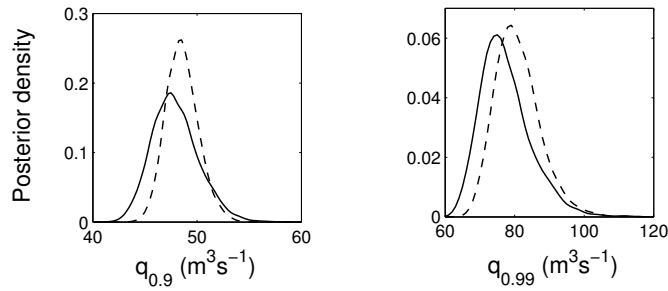


Figure 12

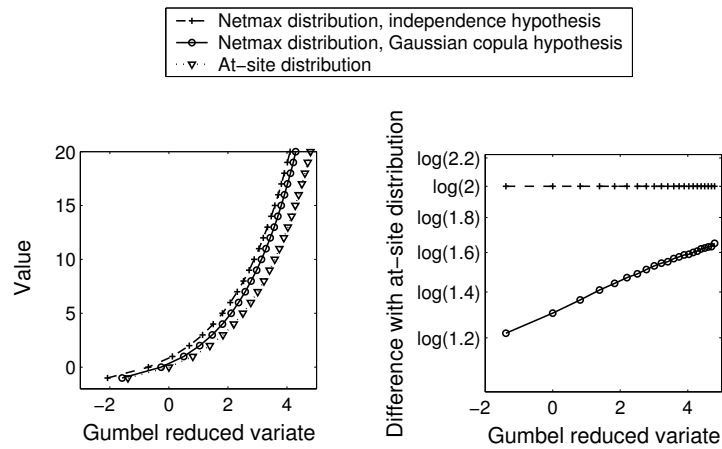


Figure 13

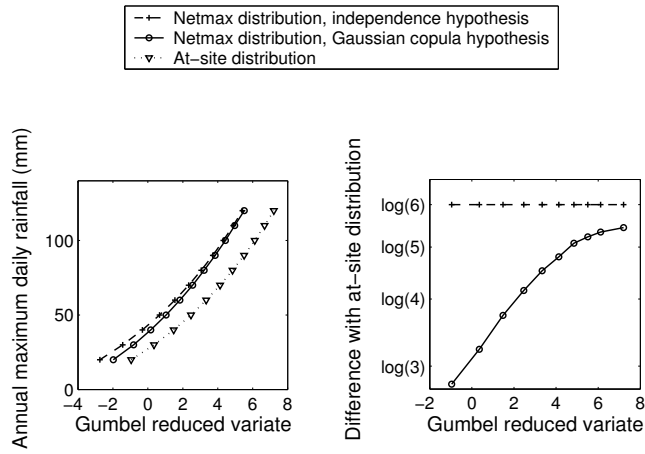


Figure 14

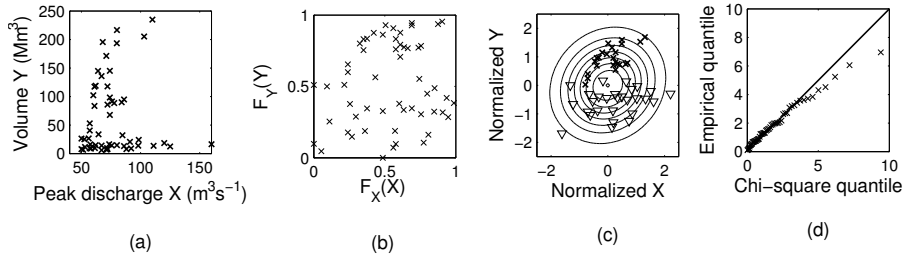


Figure 15

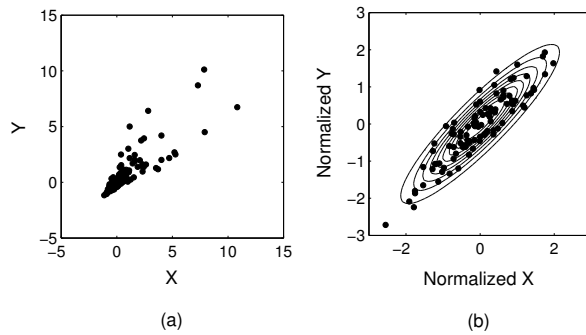


Figure 16

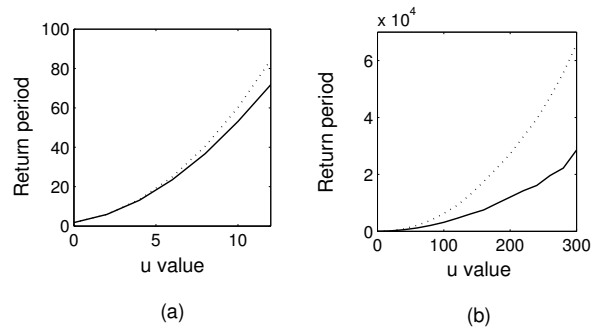


Figure 17

RESUME

Le changement climatique est une réalité largement reconnue aujourd'hui dans la communauté scientifique. Néanmoins, son impact sur les extrêmes hydrologiques est délicat à observer et à prévoir. Dans ce contexte, l'objectif de cette thèse est d'apporter des éléments de réponse aux trois questions suivantes : Comment détecter des changements dans les séries hydro-climatiques ? Quelles sont les évolutions observées sur les régimes extrêmes des cours d'eau en France métropolitaine ? Comment prendre en compte ces éventuels changements dans les méthodes de gestion des risques ? Chacun de ces trois objectifs sera traité à l'échelle locale et à l'échelle régionale.

Au niveau local, une analyse par simulation a permis de comparer un grand nombre de tests statistiques dédiés à la détection de tendances ou de ruptures, et de réaliser un schéma de sélection du test le plus approprié à la variable d'étude. Au niveau régional, deux aspects ont été pris en compte : la significativité et la cohérence régionale. Dans les deux cas, des méthodes sont proposées et comparées aux outils existant. L'analyse d'environ 200 stations hydrométriques en France ne montre pas de changement généralisé, ni en crue, ni en étiage, ni en régime. A plus petite échelle, certaines évolutions sont cependant détectées, mais les résultats ne suffisent pas à en certifier la cause : variabilité naturelle, évolution du bassin versant ou changement climatique. Enfin, les méthodes d'analyse fréquentielle ont été adaptées pour prendre en compte d'éventuelles évolutions temporelles. A cet effet, plusieurs outils statistiques se sont révélés particulièrement intéressants. Tout d'abord, l'analyse Bayésienne permet de considérer l'incertitude liée à la stationnarité. A l'échelle régionale, l'estimation des paramètres de modèles régionaux non stationnaires est possible grâce aux méthodes MCMC. Enfin, nous proposons une première piste, basée sur l'utilisation d'une copule Gaussienne, pour prendre en compte formellement la dépendance spatiale.

ABSTRACT

Climate change is widely considered as a reality by scientists. Nevertheless, impacts on hydrological extremes are more difficult to observe and to predict. The aim of this thesis is to answer the following questions: How to detect changes in hydro-climatic series? What are the observed changes for extreme discharges in France? How to take into account possible changes in frequency analysis? These objectives refer to both the local and the regional scales.

A number of at-site tests for change have been compared by simulations, leading to the selection of the more suitable method for a given variable. At the regional scale, both field significance and regional consistence are studied. New approaches are proposed and compared with the available methods. The study of almost 200 hydrometric stations does not show global patterns of change for high flows, low flows or regime. At a smaller scale, some consistent trends are detected, but without conclusion about their cause: natural variability, watershed evolution or climate change. Lastly, frequency analysis methods are modified in order to deal with possible non-stationarity. Several statistical tools were found to be suitable for this purpose. The Bayesian approach allows incorporating the uncertainty related to stationarity. At a regional scale, non-stationary models can be dealt with thanks to MCMC methods. Finally, a first approach for dealing with intersite dependence is proposed, based on a model for dependence arising from a Gaussian copula.