



Doctorat ParisTech

École Doctorale *Géosciences, Ressources Naturelles et Environnement*

THÈSE

pour obtenir le grade de docteur délivré par

**L'Institut des Sciences et Industries
du Vivant et de l'Environnement
(AgroParisTech)**

Spécialité : Hydrologie

présentée et soutenue publiquement par

Laure LEBECHEREL

le 5 mai 2015

Sensibilité des calculs hydrologiques à la densité des réseaux de mesure hydrométrique et pluviométrique

Directeur de thèse : **Vazken ANDRÉASSIAN**

Jury

M. Christophe CUDENNEC

M. François ANCTIL

M. Gilles DROGUE

M. Denis RUELLAND

Mme Bénédicte AUGÉARD

Mme Rachel PUECHBERTY

Agrocampus Ouest (Rennes, France)

Université Laval (Québec, Canada)

Université de Lorraine (Metz, France)

HydroSciences (Montpellier, France)

ONEMA (Vincennes, France)

SCHAPI (Toulouse, France)

Président

Rapporteur

Rapporteur

Examinateur

Invité

Invité

AgroParisTech

Irstea, Unité de recherche Hydrosystèmes et Bioprocédés

1 rue Pierre-Gilles de Gennes, CS 10030, 92761 Antony cedex

Remerciements

L'une des premières choses dont on se rend compte lorsque l'on commence une thèse, c'est qu'on ne peut parcourir ce long chemin seul. C'est pourquoi il est temps de remercier toutes les personnes m'ayant apporté le soutien scientifique et personnel nécessaire au bon déroulement de cette thèse.

Je souhaite tout d'abord remercier chaleureusement mon directeur de thèse, Vazken Andréassian, de m'avoir donné l'opportunité de mener ces travaux de recherche. Je le remercie également pour sa disponibilité, sa patience et ses conseils avisés. Je souhaite remercier également Charles Perrin d'avoir toujours été disponible pour suivre mes travaux de recherche et d'avoir participé attentivement à la relecture de ce mémoire. Ses remarques et conseils m'ont été très précieux tout au long de la thèse.

J'adresse mes sincères remerciements aux membres de mon jury de thèse, qui ont accepté d'évaluer le travail réalisé au cours de ces trois années : François Anctil et Gilles Drogue pour avoir accepté le rôle de rapporteurs et pour avoir participé à l'amélioration de ce manuscrit, ainsi que Christophe Cudennec, Denis Ruelland, Rachel Puechberty et Bénédicte Augeard pour leurs rôles d'examineurs.

Je tiens à remercier également les membres de mes comités de suivi pour leurs regards critiques et leurs précieux conseils, qui m'ont permis d'avancer dans ce travail de recherche : Rémy Garçon, Pierre Javelle, Nicolas Le Moine, Pascal Maugis et Jean-Michel Soubeyroux.

Je souhaite remercier l'ONEMA pour le soutien financier apporté à cette thèse, et une nouvelle fois Bénédicte Augeard pour avoir suivi ces travaux de recherche et y avoir apporté sa vision opérationnelle.

Il est également important pour moi de remercier les enseignants hydrologues que j'ai pu rencontrer lors des 5 années d'étude effectuées à l'UPMC, en amont de cette thèse, et qui m'ont donné goût à l'Hydrologie : Pierre Ribstein, Ludovic Oudin, Laetitia Moulin, Hocine Bendjoudi...

Je remercie bien sûr l'équipe jeune et dynamique Hydrologie d'Irstea Antony, qui m'a accueillie durant ces trois ans de thèse (et même un peu plus) et auprès de qui j'ai beaucoup appris : Helena, Guillaume, Carina, Carine, Pierre N., Olivier, Andréa, Philippe, Alban, ... Merci aussi à ceux qui nous ont quittés en route : Marine, Julien, Damien, Mathilde, David, ainsi que tous les stagiaires passés par l'équipe. Merci aux anciens doctorants de nous avoir montré la marche à suivre (Annie, Ioanna, Laurent, Pierre B., Dimitri, Florent et François). Un merci particulier à Ioanna, ma première co-bureau, pour les quelques notions de Grec qu'elle a essayé de m'inculquer. Je la remercie sincèrement pour son amitié et son soutien permanent encore aujourd'hui. Un grand merci aussi à Louise pour nos nombreux échanges sur nos travaux de recherche respectifs et pour avoir toujours su trouver les mots justes quand le doute s'installait. Merci à Julie pour sa joie de vivre et son petit

grain de folie. Et enfin, merci à tous les membres de l'Hydrologie Sociale pour les week-ends socialement hydrologiques et les fameux Mamanes à la Butte aux Cailles.

Merci aussi à toute l'unité HBAN pour leur chaleureux accueil et la bonne ambiance générale qui y règne au quotidien. En particulier, merci à l'équipe de choc : Louise, Julie, Violaine, Nass, Laetitia, Carolina, Carina, ... pour leur précieuse amitié et nos nombreuses soirées détentes. Même s'il était rare de me voir sur le terrain de frisbee, je remercie aussi tous ceux avec qui j'ai pu jouer ces quelques parties. Je tiens également à remercier l'équipe d'appui à la recherche (Nathalie, Laurence et Elisabeth) pour leur précieuse aide dans les démarches administratives, ainsi que Roger, notre super informaticien, pour sa disponibilité et sa bonne humeur permanente.

En dernier lieu, j'adresse mes sincères remerciements à ma famille et mes amis, qui m'ont toujours soutenue et qui ont su m'accompagner et m'encourager durant ces trois ans.

Enfin, un grand merci à Romain pour son soutien permanent et sa présence au quotidien.

Résumé

Les données de pluie et de débit sont d'une importance capitale pour réaliser des calculs hydrologiques. La pluie est un élément essentiel pour les études de bilan hydrique ainsi que pour la prévision et la simulation des débits, puisqu'elle est utilisée en entrée des modèles hydrologiques. Les données de débit sont également essentielles pour caler et valider les modèles : elles informent sur les régimes et les extrêmes, les tendances passées, et sur le comportement hydrologique du bassin versant. La pluie et le débit étant des éléments variables dans le temps et l'espace, une bonne représentativité spatio-temporelle des informations de débit et de pluie est capitale afin de limiter les incertitudes des calculs hydrologiques. L'existence de réseaux de mesure hydrométéorologiques suffisamment denses pour rendre compte de cette variabilité est, de ce fait, essentielle.

Ces réseaux peuvent cependant paraître onéreux pour leurs gestionnaires, entraînant des réflexions sur leur rationalisation. Toutefois, cette rationalisation, qui se traduit souvent par une baisse de la densité du réseau de mesure, rend notre connaissance du cycle hydrologique plus incertaine, et peut, par conséquent, augmenter les incertitudes des calculs hydrologiques. Quantifier cette augmentation présente un certain nombre de difficultés, car elle dépend de l'objectif hydrologique, des outils utilisés et des caractéristiques des bassins versants étudiés.

Le principal objectif de la thèse était d'étudier l'impact de la densité spatio-temporelle des réseaux hydrométriques et pluviométriques sur les performances de divers calculs hydrologiques (simulation de débit au pas de temps journalier, estimation du module, de débit de crue extrême et de caractéristiques d'étiage). Afin de produire des conclusions générales, les recherches se sont appuyées sur un large échantillon de bassins versants français.

La première partie a porté sur l'impact de la densité des réseaux hydrométriques pour des bassins peu ou non jaugés. Pour les bassins non jaugés, la robustesse de la méthode de régionalisation a dans un premier temps été analysée selon deux méthodes de réduction de la densité du réseau voisin (désert hydrométrique et réduction aléatoire). La méthode du désert hydrométrique a été par la suite retenue pour évaluer la sensibilité des calculs hydrologiques à la disponibilité spatiale des informations de débit. Nos résultats suggèrent que pour tous les calculs envisagés, les performances du processus de régionalisation diminuent lorsque le réseau de bassins voisins devient moins dense, mais que cette chute de performances est moindre en comparaison de celle attribuée à la méthode de régionalisation elle-même (rien ne vaut les observations sur le site d'étude). Dans un second temps, nous avons confirmé l'intérêt d'utiliser quelques mesures ponctuelles de débit sur ces bassins non jaugés, en combinant cette information à une information régionale. Nous avons poussé plus loin l'analyse en nous intéressant à la différence entre mesures redondantes et mesures aléatoires et en proposant des équivalences.

La deuxième partie a porté sur l'impact de la densité spatiale des réseaux pluviométriques sur divers calculs hydrologiques. Les résultats sont moins généralisables que pour le réseau hydrométrique, révélant des tendances variées au sein de l'échantillon de bassins et entre les calculs hydrologiques ciblés. Toutefois, les baisses des performances du modèle GR4J lorsque la densité du réseau pluviométrique diminue semblent être liées à la variabilité spatiale de la pluie du bassin versant.

Abstract

Rainfall and runoff data are very important to make hydrological computations for various objectives. Precipitation is essential information for water balance studies and for hydrological simulation and forecasting, since it is used as input of hydrological models. Runoff data are also essential to calibrate and validate models: they provide information on hydrological regimes, streamflow extremes, past trends and catchment hydrological behavior. Since precipitation and runoff are spatially and temporally variable, a good spatiotemporal representativeness of runoff and rainfall data is crucial to limit the uncertainty in hydrological computations. Therefore the presence of sufficiently dense hydrometeorological monitoring networks to account for this variability is essential.

However, these networks may be expensive for their managers, leading to reflections on their rationalization. This rationalization, which often means a reduction of network density, can limit our knowledge of the hydrological cycle, and can significantly increase the uncertainties in hydrological computation. Quantifying this increase raises several difficulties, since it depends on the hydrological objectives, tools used and catchment characteristics.

The main objective of this research was to study the impact of the spatiotemporal density of runoff and rainfall networks on the performances of several hydrological computations (namely simulation of daily streamflow, estimation of long term average streamflow and extreme streamflows). To get general results, studies were based on a large set of French catchments.

The first part of this thesis focused on the impact of the hydrometric network density for ungauged or poorly gauged catchments. For ungauged catchments, the robustness of regionalization approaches was first analyzed by two methods of reduction of neighboring network density (hydrometrical desert and random reduction). Then the hydrometrical desert method was used to evaluate the sensitivity of hydrological computations on the spatial availability of runoff data. Our results suggest that for all the computations of streamflow considered here, the efficiency of the regionalization process decreases when the flow gauging network density is reduced. Results also show that this drop in efficiency is lower than the drop due to the regionalization method itself (i.e. going from gauged to ungauged situation; nothing beats the observed data on the studied sites). In a second step, we confirmed the value of using a few flow measurements on ungauged catchments, by combining this information with regional information. We pushed the analysis further by focusing on the differences between redundant measurements and random measurements, and by proposing equivalences.

The second part focused on the impact of the rainfall network density on several hydrological computations. Results are less generalizable than those on the hydrometric network, revealing various trends within the set of catchments and between the hydrological computations considered. However, the decrease of GR4J model performances when the density of rainfall network is reduced seems to be related to the spatial variability of rainfall in the catchment.

Table des matières

Remerciements	1
Résumé	3
Abstract	5
Introduction générale	13
Chapitre 1. Optimisation de réseaux et valeur de l'information hydrométéorologique: état de l'art, contexte de la thèse	19
1.1. Introduction	21
1.2. Echelles temporelle et spatiale des variables hydrométéorologiques	21
1.2.1 Variabilité spatio-temporelle des précipitations.....	21
1.2.2 Variabilité spatio-temporelle des débits	22
1.3. Coût des réseaux de mesure	24
1.3.1 Coût des réseaux hydrométéorologiques d'Electricité de France	24
1.3.2 Coût du réseau pluviométrique de Météo-France	24
1.3.3 Coût du réseau hydrométrique national.....	25
1.4. L'idée de rationalisation des réseaux de mesure	25
1.5. La valeur économique des réseaux de mesure	27
1.5.1 La valeur économique des réseaux de mesure et des outils hydrométéorologiques en général	27
1.5.2 Optimisation du réseau hydrométrique.....	29
1.6. Synthèse	30
Partie I Impact de la densité du réseau hydrométrique sur les calculs hydrologiques	33
Chapitre 2. Impact de la densité du réseau hydrométrique en hydrologie : état de l'art	35
2.1. Introduction	37
2.2. Impact de la densité spatiale du réseau hydrométrique en hydrologie	37

2.2.1	Efficacité de différentes approches d'estimation de paramètres sur bassins versants non jaugés	37
2.2.2	Etudes sur le nombre de bassins voisins utilisés pour le transfert d'information à des bassins non jaugés.....	39
2.3.	Impact de la longueur des chroniques de débit en hydrologie	43
2.3.1	Longueur des chroniques de débit pour le calage des paramètres	43
2.3.2	Longueur des chroniques de débit pour le calcul des débits caractéristiques	49
2.4.	Synthèse	50
Chapitre 3.	Matériel et méthodes.....	51
3.1.	Introduction.....	53
3.2.	Présentation de la base de données de la partie I.....	53
3.2.1	Les données hydro-climatiques.....	53
3.2.2	La sélection d'un échantillon de bassins versants.....	54
3.3.	Le modèle hydrologique pluie-débit GR4J	57
3.3.1	Description du modèle	57
3.3.2	Utilisation du module neige CemaNeige.....	58
3.3.3	Calage des paramètres du modèle GR4J et du module CemaNeige	58
3.4.	La méthode de régionalisation basée sur la proximité spatiale	60
3.5.	Etude de l'impact de la distance hydrologique utilisée pour la modélisation de bassins versants non jaugés	62
3.5.1	Pourquoi est-il important d'évaluer la sensibilité des simulations régionalisées au choix de la distance hydrologique ?.....	62
3.5.2	Les différentes mesures de distance étudiées	63
3.5.3	Comparaison des différentes distances hydrologiques dans une perspective d'une application sur un bassin versant non jaugé.....	65
3.5.4	Conclusion	66
3.6.	Comparaison de deux méthodes d'évaluation de la robustesse de méthodes de régionalisation.....	68
3.6.1	Pourquoi est-il important d'évaluer la sensibilité d'une méthode de régionalisation à la densité du réseau hydrométrique ?.....	68
3.6.2	Deux méthodes alternatives pour l'analyse de sensibilité d'une méthode de régionalisation basée sur la proximité spatiale	68
3.6.3	Comparaison des deux méthodes d'analyse de sensibilité.....	71
3.6.4	Conclusion	76
3.7.	Synthèse.....	77

Chapitre 4. Une régionalisation de la formule de bilan en eau de Turc-Mezentsev	79
4.1. Introduction.....	82
4.1.1 The need for a reliable assessment of long-term average streamflow.....	82
4.1.2 Scope of the paper	82
4.2. Origins and applications of the Turc-Mezentsev formula	83
4.2.1 Origins of the Turc-Mezentsev formula	83
4.2.2 Physical and mathematical interpretation of the Turc-Mezentsev formula.....	84
4.2.3 Transposing the Turc-Mezentsev formula for streamflow.....	85
4.2.4 Recent studies based on the Turc-Mezentsev formula.....	86
4.3. Catchments and method	88
4.3.1 Data set.....	88
4.3.2 Testing geographical (spatial proximity-based) regionalization.....	90
4.3.3 Measuring regionalization performance	92
4.3.4 Assessing regionalization robustness using the hydrometrical desert method.....	92
4.4. Results.....	93
4.4.1 Comparison of the efficiencies for the four regionalization options	93
4.4.2 Robustness assessment.....	96
4.4.3 Discussion	97
4.4.4 Does catchment size have an impact on the results?	98
4.5. Conclusion and perspectives.....	98
4.5.1 Conclusion	98
4.5.2 Limits	99
4.5.3 Perspectives.....	99
4.6. Acknowledgements.....	99
4.7. Annex 1 – partial differential equations for streamflow as a function of P and E_0.....	100
Chapitre 5. Analyse de la sensibilité des calculs hydrologiques à la densité spatiale des réseaux hydrométriques	101
5.1. Introduction.....	103
5.2. Les calculs hydrologiques considérés	103
5.2.1 Simulation d’une chronique de débit journalier en mode non-jaugé avec le modèle GR4J	103
5.2.2 Calcul du module en mode non-jaugé.....	103
5.2.3 Calcul des débits de crues décennale et centennale en mode non-jaugé	103
5.2.4 Calcul des caractéristiques d’étiage en mode non-jaugé.....	104

5.3.	Critère de calage des paramètres et d'évaluation du modèle.....	104
5.4.	Evaluation préalable suivant les gammes de débit	106
5.5.	Evaluation de la robustesse de la régionalisation	106
5.6.	Analyse détaillée des performances obtenues en régionalisation	115
5.7.	Synthèse	127
Chapitre 6.	Stratégie de récolte de jaugeages ponctuels sur des bassins versants non-jaugés : au sujet de la redondance de l'information hydrologique.....	129
6.1.	Introduction.....	131
6.2.	Sélection de jaugeages aléatoires (SA) - Résumé de l'article présenté en annexe B	131
6.2.1	La méthode.....	131
6.2.2	Résultats	132
6.3.	Sélection du débit observé selon la gamme de débit classé (sélection redondante (SR))	134
6.3.1	Introduction.....	134
6.3.2	Méthodologie	135
6.3.3	Sensibilité par rapport au critère d'évaluation et au paramètre α	135
6.4.	Doit-on privilégier la quantité et la redondance ou la diversité des mesures?	138
6.5.	Synthèse	141
Partie II	Impact de la densité du réseau pluviométrique sur les calculs hydrologiques	143
Chapitre 7.	Impact de la densité du réseau pluviométrique en hydrologie : état de l'art.....	145
7.1.	Introduction.....	147
7.2.	Impact de la densité spatiale du réseau pluviométrique en hydrologie	147
7.2.1	Impact sur la qualité des estimations de la pluie	147
7.2.2	Impact sur la qualité des estimations de débit	149
7.3.	Impact de la longueur des enregistrements des données de pluie en hydrologie.....	153
7.4.	Synthèse	154
Chapitre 8.	Présentation de la base de données de la partie II.....	155
8.1.	Introduction.....	157
8.2.	Les données hydro-climatiques.....	157
8.2.1	Origine des données pluviométriques.....	157
8.2.2	Interpolation de la pluie de bassin	157

8.2.3	Les données de débit et d'évapotranspiration potentielle	158
8.3.	La sélection d'un échantillon de bassins versants	158
8.3.1	Critères de sélection	159
8.3.2	Sélection de l'échantillon	159
8.4.	Les calculs hydrologiques, le calage du modèle et leur évaluation	162
8.4.1	Les calculs hydrologiques considérés	162
8.4.2	Le calage des paramètres et évaluation du modèle.....	162
8.5.	Synthèse.....	162
Chapitre 9.	Analyse de la sensibilité des calculs hydrologiques à la densité spatiale des réseaux pluviométriques.....	165
9.1.	Introduction.....	167
9.2.	Quelques éléments de méthodologie	167
9.2.1	Mise en place des différentes classes de densité spatiale du réseau pluviométrique.....	167
9.2.2	Présentation des bassins-tests	168
9.3.	Impact de la densité spatiale du réseau pluviométrique sur divers calculs hydrologiques	170
9.3.1	Sur les deux bassins-tests	170
9.3.2	Analyse pour l'ensemble des 30 bassins versants de l'échantillon.....	174
9.4.	Tentatives d'interprétation des résultats	176
9.4.1	Relation de la courbe de décroissance des performances lorsque la densité du réseau pluviométrique diminue, avec la surface pluviométrique du bassin versant.....	176
9.4.2	Relation de la décroissance des performances lorsque la densité du réseau pluviométrique diminue, avec la variabilité spatiale de la pluie	177
9.4.3	Interprétation de la courbe de décroissance des performances en fonction d'autres caractéristiques	182
9.4.4	Sensibilité des performances par rapport à la diminution de la densité du réseau pluviométrique.....	184
9.5.	Synthèse.....	184
Conclusion générale	187
Bibliographie	195
Annexes	211

Introduction générale

Contexte de la thèse

L'hydrologue est spécialisé dans l'étude du cycle de l'eau, en particulier la transformation des précipitations en écoulement dans les rivières. Il est sollicité pour diverses missions utiles à la société, telles que l'évaluation de la ressource en eau, le dimensionnement d'ouvrages hydrotechniques et la prévision des événements extrêmes (crue et étiage). Ces missions ont pour enjeux de garantir la sécurité et la rentabilité des ouvrages, d'alerter la population et de réduire les pertes humaines et matérielles, ainsi que de contribuer au maintien du « bon état écologique » du cours d'eau et de sa capacité à fournir de l'eau potable, de l'eau d'irrigation et du liquide de refroidissement pour les centrales thermiques.

Toutefois, pour réaliser les calculs qui lui permettront de répondre à ces objectifs, l'hydrologue a besoin de données hydrométéorologiques de qualité et en quantité suffisante. Les données de débit sont par exemple essentielles pour caler et valider les modèles hydrologiques. Elles informent sur les régimes et les extrêmes, les tendances passées, et sur le comportement hydrologique du bassin versant. Par ailleurs, les séries météorologiques (pluie, température, rayonnement,...) permettent d'alimenter les modèles hydrologiques et de comprendre la genèse des écoulements. Elles sont également indispensables pour le calcul du bilan hydrique.

Dans ce contexte, on peut facilement admettre que la qualité et la quantité de ces données hydrométéorologiques peuvent avoir un impact non négligeable sur l'exactitude des calculs hydrologiques.

Problématique opérationnelle

La mise en place et l'entretien des réseaux de mesure hydrométéorologiques représentent un coût important pour leurs gestionnaires. C'est pourquoi, ces derniers réfléchissent de plus en plus à la manière dont ils peuvent rationaliser ces réseaux, c'est-à-dire optimiser au mieux l'emplacement et le nombre des instruments de mesure, de façon à réduire leurs coûts d'investissement et de fonctionnement tout en assurant un bon niveau d'information. A titre d'exemple, dans les décennies 1980 et 1990, pour des raisons principalement économiques, de considérables réductions de densité des réseaux hydrologiques ont été constatées dans de nombreux pays : Canada (21%), Finlande (7%), Nouvelle-Zélande (20%) [Pearson, 1998] et Etats-Unis (6%) ; et ces réductions se poursuivent [WMO, 2008]. Mais, si la réduction des réseaux de mesure hydrométéorologiques peut sembler un moyen pour les gestionnaires de limiter les dépenses à court terme, cela peut aussi se traduire par une perte économique à long terme, en raison des incertitudes additionnelles qui affecteront les calculs hydrologiques qui utilisent les données ainsi produites.

Quantifier les incertitudes des calculs hydrologiques liées à la densité spatio-temporelle des réseaux hydrométéorologiques représente ainsi un enjeu très important. Toutefois, cela n'est pas simple, car ces évaluations d'incertitude vont dépendre d'un certain nombre de facteurs : l'objectif hydrologique, les outils utilisés, les caractéristiques des bassins versants étudiés, etc. Par ailleurs, les besoins en matière de conception de réseaux peuvent varier selon l'objectif hydrologique fixé : par exemple, les besoins en termes de densité spatio-temporelle des réseaux hydrométéorologiques pour l'estimation des débits de crues ne sont pas nécessairement les mêmes que ceux destinés à

l'estimation des caractéristiques d'étiage. Il semble ainsi important de couvrir une certaine diversité d'objectifs hydrologiques pour concevoir les réseaux d'observation.

Problématique scientifique

De nombreuses études ont montré que la variabilité spatiale et temporelle des événements de pluie peut être importante [Berne et al., 2004; Emmanuel et al., 2012; Krajewski et al., 2003]. Cette pluie est classiquement mesurée par des pluviomètres, qui mesurent la quantité d'eau précipitée en un point. Pour des raisons économiques et pratiques, il est impensable d'installer des stations pluviométriques en chaque point du bassin versant. Une méthode d'interpolation est alors nécessaire afin d'estimer une pluie globale à l'échelle du bassin. A l'évidence, l'efficacité de ces méthodes d'interpolation va être conditionnée par l'échantillonnage de cette variable dans le temps et dans l'espace. Dans cette recherche, nous nous essaierons en particulier de répondre aux questions suivantes :

- *Dans quelle mesure la baisse de la densité du réseau pluviométrique va-t-elle affecter les performances des calculs hydrologiques ? Quel sera le « coût hydrologique » de cette réduction de densité ?*
- *Si l'on accepte une perte de performance des estimations hydrologiques de 10%, comment cela peut-il se traduire en termes de densité pluviométrique ?*

Pour le réseau hydrométrique, la question se pose différemment de celle du réseau pluviométrique. Pour une application en un point donné d'un cours d'eau, la situation idéale est bien entendu d'avoir une longue série de mesures issue d'une station pérenne. Il va de soi que rien ne peut être meilleur que les observations. Malheureusement, cette situation idéale n'est pas toujours rencontrée et les incertitudes liées aux méthodes alternatives permettant l'estimation de ces débits en des points non ou peu jaugés, peuvent être relativement importantes et dépendantes de la disponibilité spatio-temporelle des données hydrométriques avoisinantes. Au cours de cette thèse, nous allons donc essayer de répondre aux questions suivantes :

- *Selon la densité spatiale du réseau hydrométrique voisin, à quelles performances des calculs hydrologiques pouvons-nous nous attendre sur un bassin versant non jaugé et selon quelle gamme d'incertitude ? Comment les calculs hydrologiques sont-ils dégradés lorsque la densité spatiale du réseau hydrométrique voisin diminue ?*
- *Quel gain de performances l'information hydrométrique ponctuelle peut-elle apporter sur des bassins versants non jaugés ? Quelle gamme de débit est-il préférable d'échantillonner de façon à être le plus efficace pour la détermination des paramètres du modèle à transférer ? Doit-on privilégier la quantité et la redondance des mesures ponctuelles ou leurs diversités ?*

Au cours de cette thèse, nous souhaitons pouvoir couvrir une certaine diversité de calculs hydrologiques. Nous nous intéresserons en particulier à la simulation de débit au pas de temps journalier, l'estimation du module, de débits de crue extrême et de caractéristiques d'étiage.

Organisation du mémoire

Ce mémoire comprend neuf chapitres.

Le premier chapitre expose le contexte général de nos travaux. Il introduit l'intérêt d'un bon échantillonnage spatial et temporel des données de pluie et de débit, et expose quelques chiffres concernant le coût des réseaux pluviométriques et hydrométriques en France.

Les huit chapitres suivants sont répartis en deux grandes parties.

La première partie porte sur l'étude de l'impact de la densité spatio-temporelle du réseau hydrométrique sur les calculs hydrologiques.

Le deuxième chapitre propose une synthèse bibliographique concernant l'impact de la densité du réseau hydrométrique sur les calculs hydrologiques, en soulignant l'importance d'un réseau de mesure relativement dense pour la modélisation des bassins versants non ou peu jaugés. Puis, une synthèse des études menées sur l'impact de la durée des chroniques de débit est présentée.

Le troisième chapitre présente la base de données ainsi que les méthodologies adoptées au cours de cette première partie. En particulier, nous présentons l'échantillon de bassins versants, le modèle hydrologique et la méthode de régionalisation retenus pour les différents tests concernant l'impact de la densité du réseau hydrométrique. De plus, différentes distances séparant bassin versant non jaugé et bassin voisin sont évaluées selon leur efficacité. Enfin, deux méthodes visant à évaluer la robustesse de la méthode de régionalisation sont également proposées et comparées.

Le quatrième chapitre est un article visant à évaluer les possibilités de régionalisation de la formule de Turc-Mezentsev, afin d'améliorer son efficacité sur des bassins non jaugés. Cette formule relie le module aux précipitations et à l'évapotranspiration potentielle à long-terme. Dans ce chapitre, plusieurs options de régionalisation de cette formule sont ainsi comparées et la méthode du désert hydrométrique est également appliquée pour évaluer la robustesse des différentes possibilités de régionalisation.

Le cinquième chapitre présente une étude de sensibilité concernant l'impact de la densité spatiale du réseau hydrométrique sur divers calculs hydrologiques. Dans un premier temps, ces calculs hydrologiques sont exposés. Puis, la méthode du désert hydrométrique est appliquée pour l'analyse de sensibilité.

Le sixième chapitre expose la méthodologie et les principaux résultats de précédents travaux réalisés au sein de l'équipe, concernant l'utilisation de données ponctuelles sur des bassins non jaugés, en les couplant avec une information régionale. Dans ce chapitre, nous proposons d'aller plus loin dans l'analyse en nous intéressant à la différence entre mesures redondantes et mesures aléatoires et en proposant des équivalences.

La seconde partie de thèse porte sur l'étude de l'impact de la densité spatiale du réseau pluviométrique sur les calculs hydrologiques.

Le septième chapitre dresse un état de l'art de l'impact de la densité spatiale du réseau pluviométrique sur les calculs hydrologiques et de la question de l'impact de la longueur des séries disponibles.

Le huitième chapitre présente les données hydrométéorologiques, ainsi que le nouvel échantillon de bassins versants utilisés pour cette seconde partie.

Le neuvième chapitre expose la méthodologie suivie pour tester l'impact de la densité spatiale du réseau pluviométrique sur divers calculs hydrologiques. Les résultats associés à ces tests sont par la suite présentés et analysés.

Enfin, nous présentons les principales conclusions qui découlent de ces travaux de thèse, d'un point de vue scientifique et opérationnel, et évoquons quelques perspectives de recherche.

La Figure 1 illustre synthétiquement les tests détaillés ci-dessus afin de répondre aux objectifs de la thèse.

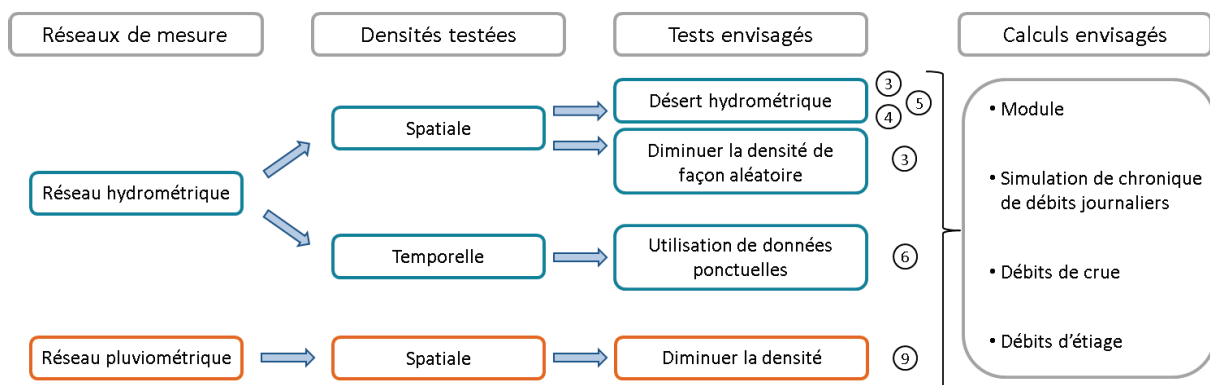


Figure 1. Illustration synthétique de la méthode utilisée afin de répondre aux objectifs de la thèse. La couleur bleue correspond aux tests concernant l'impact de la densité spatiotemporelle du réseau hydrométrique et la couleur orange concerne les tests envisagés sur l'impact de la densité spatiale du réseau pluviométrique sur les divers calculs hydrologiques. Les chiffres correspondent au numéro du chapitre où se trouvent la méthodologie et les résultats du test envisagé.

Chapitre 1.
**Optimisation de réseaux et valeur de
l'information hydrométéorologique: état de
l'art, contexte de la thèse**

1.1. Introduction

Dans ce chapitre, nous allons dans un premier temps montrer en quoi un bon échantillonnage spatial et temporel des données de pluie et de débit est important, puis nous donnerons quelques chiffres concernant le coût des réseaux pluviométriques et hydrométriques en France. Par la suite, nous parlerons de la rationalisation des réseaux de mesure ainsi que de la valeur économique et hydrologique de ces réseaux, et des outils hydrométéorologiques qui en dépendent. Enfin, des études d'optimisation de réseaux seront passées en revue.

1.2. Echelles temporelle et spatiale des variables hydrométéorologiques

La pluie et le débit sont des éléments très importants en hydrologie mais ceux-ci peuvent être très variables spatialement et temporellement.

1.2.1 Variabilité spatio-temporelle des précipitations

La pluie est classiquement mesurée à l'aide d'un réseau de postes d'observation et de mesure au sol, et plus précisément à l'aide de pluviomètres. Ces pluviomètres mesurent la quantité d'eau tombée en un point. Etant donné qu'on ne peut mesurer la pluie en tout point de l'espace, l'estimation de la pluie sur de grandes zones géographiques, comme par exemple les bassins versants, nécessite d'interpoler ces mesures de pluies ponctuelles à l'aide de méthodes d'interpolation, qui peuvent être variées.

La pluie étant un phénomène variable temporellement et spatialement, les erreurs d'estimation liées à son échantillonnage peuvent être non négligeables [Krajewski *et al.*, 2003; Roche, 1963; Storm *et al.*, 1988]. Ainsi, nous pouvons facilement imaginer que la configuration et la densité du réseau pluviométrique vont jouer un rôle très important dans l'estimation de la pluie à plus grande échelle (prévision météorologique à l'échelle nationale, estimation de la pluie de bassin, etc.). Toutefois, il se peut que, selon les zones géographiques ou les besoins opérationnels, la densité et la configuration optimales de ce réseau varient. Par exemple, une zone montagneuse qui présente un fort gradient de pluie en fonction de l'altitude aura très probablement besoin d'une densité de pluviomètres plus élevée qu'une zone de plaine avec une pluie plus homogène.

Certaines études portant sur des bassins versants français ont eu pour but d'essayer de déterminer quel serait le meilleur échantillonnage spatial et temporel de la pluie afin de pouvoir capter au mieux les événements pluvieux. Berne *et al.* [2004] ont montré, à l'aide d'information radar et d'outils géostatistiques (variogrammes), que la résolution spatio-temporelle de la pluie sur des bassins versants urbains en région méditerranéenne (pouvant présenter des événements de pluie très intenses) dépendait surtout de la surface des bassins versants. Ils montrent que, pour des bassins urbains de 10 km², les résolutions temporelle et spatiale des pluies devraient être respectivement d'environ 5 min et 3 km. Pour des plus petits bassins de l'ordre de 1 km², des résolutions d'environ 3 min et 2 km seraient nécessaires. Ils proposent également une méthode permettant de calculer la résolution spatiale correspondant à un réseau pluviométrique existant. Ceci permet de déterminer si le réseau respecte les recommandations proposées et d'évaluer la taille minimale du bassin pour lequel les données du réseau peuvent être utilisées à des fins hydrologiques. Dans une autre étude, Berne *et al.* [2009] se sont intéressés à deux événements pluvieux intenses de nature différente ayant eu lieu dans cette même zone méditerranéenne en 2002. L'étude avait pour but de mettre en

place, à l'aide de méthodes géostatistiques, une méthode permettant d'obtenir la structure spatiale des pluies intenses.

Sous un climat davantage océanique, l'étude d'*Emmanuel et al.* [2012] visait également à déterminer la variabilité spatiale et temporelle de différents types d'événements pluvieux, à l'aide de variogrammes et d'images radar. Ils ont déterminé quatre types d'événements avec des résolutions spatiales et temporelles différentes. Selon le type d'événement, les distances de décorrélation varient entre 5 et 17 km et les temps de décorrélation sont estimés entre 5 et 15 min. En appliquant la méthode de *Berne et al.* [2004], afin d'avoir une représentation correcte du champ pluvieux, ils recommandent une distance inter-stations pluviométriques variant entre 2,5 km et 6,5 km selon le type d'événement, pour une résolution temporelle de 5 min.

Ainsi, ces études ont montré que la variabilité spatiale et temporelle des événements de pluie pouvait être très importante, notamment en région méditerranéenne où des événements très intenses peuvent avoir lieu. Pour pouvoir capter ces événements pluvieux, ceci implique des résolutions temporelles et spatiales plus ou moins fines, selon le type d'événement. Toutefois, comme nous l'avons déjà mentionné, pour des raisons économiques et pratiques, il paraît difficile de pouvoir installer des stations pluviométriques à des résolutions spatiales aussi fines. La densité spatiale des stations pluviométriques va alors avoir un impact probablement très important sur l'estimation des pluies de bassin versant, et donc potentiellement sur les applications hydrologiques qui utilisent ces pluies.

1.2.2 Variabilité spatio-temporelle des débits

Le débit d'un cours d'eau résulte de l'interaction entre divers processus : précipitation, évapotranspiration, infiltration, ruissellement hypodermique et écoulement souterrain. Ce débit est mesuré au niveau de l'exutoire du bassin versant. Un bassin versant peut être défini comme l'ensemble d'une région ayant un exutoire commun pour ses écoulements de surface. La surface de ce bassin versant est délimitée par tous les points hauts de la région qui forment une ligne de partage des eaux. La pluie tombant en dehors de cette surface ne contribue pas au débit de ce bassin.

La mesure des écoulements d'un bassin s'effectue généralement par la mesure de la hauteur d'eau à l'exutoire du bassin. Afin d'obtenir le débit de ce cours d'eau, une relation univoque hauteur-débit propre à cette station hydrométrique doit être établie. Il s'agit d'une courbe de tarage qui est simplement une table de correspondance hauteurs-débits. Il est important afin d'assurer la pérennité de la courbe de tarage, que la géométrie du lit au niveau de cette station reste stable dans le temps [*Cosandey et Robinson*, 2000].

Le débit est une grandeur non continue spatialement (excepté le long du cours d'eau) et également variable dans le temps et l'espace. A la différence des précipitations ou de la lithologie qui sont spatialement continus, l'interpolation de débit entre deux bassins est difficile car l'eau suit uniquement le réseau hydrographique (Figure 2). Deux bassins versants proches peuvent avoir des débits corrélés, mais cela va provenir essentiellement de la corrélation spatiale de la pluie sur de faibles distances ainsi que des caractéristiques physiographiques proches entre ces deux bassins versants. *Skoien et al.* [2006] soulignent la difficulté d'établir une interpolation de débit entre bassins versants, en raison du fait que les bassins versants sont organisés en sous-bassins. Par conséquent ils

estiment que la distance euclidienne séparant deux bassins n'est pas la façon naturelle de mesurer une distance entre bassins.

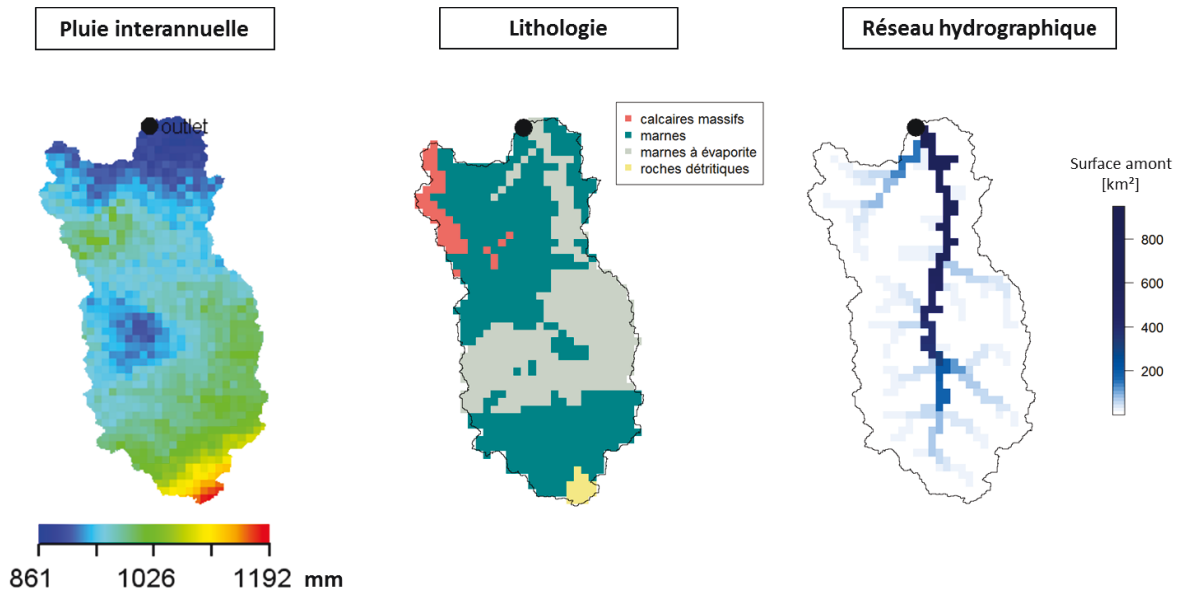


Figure 2. Pluie moyenne interannuelle, lithologie et réseau hydrographique du bassin du Madon à Pulligny (superficie de 948 km²).

L'échantillonnage temporel nécessaire aux données de débit va être dépendant des objectifs hydrologiques. Par exemple, le besoin de récolte des données pour la prévision des crues sera plutôt de l'ordre de l'horaire (voire un pas de temps encore plus fin), alors que la prévision des étiages se contentera de données journalières ou décadales.

L'objectif principal des réseaux hydrométriques est d'obtenir suffisamment d'information sur la disponibilité des ressources en eau, leur distribution géographique et leur variabilité dans le temps. En hydrologie, l'amplitude et la fréquence des crues ou des sécheresses sont particulièrement importantes [WMO, 2008]. Mais pour cela, les gestionnaires et scientifiques ont besoin de données suffisamment longues pour des études de tendance de long terme. Par ailleurs, les utilisateurs de ces données sont nombreux et variés (hydrologues, climatologues, agronomes, gestionnaires de ressources en eau et décideurs au sein des gouvernements et des secteurs privés), ce qui rend la gestion des réseaux de mesures encore plus difficiles.

Sene et Farquharson [1998] présentent un certain nombre d'arguments en faveur de la collecte continue des données hydrologiques. Ceux-ci sont basés sur le besoin d'avoir de longues chroniques afin d'obtenir une estimation raisonnable des variabilités des débits. Ils comparent les situations des données hydrologiques de l'Afrique et de l'Asie avec celles très contrastées de l'Europe. Les auteurs montrent que les faibles performances opérationnelles de plusieurs cas d'études en Afrique et Asie s'expliquent par l'insuffisance des données hydrologiques.

Par ailleurs, depuis plusieurs années, les impacts du changement climatique global sur les extrêmes hydrologiques (crues et sécheresses) en termes de fréquence et de sévérité ont été largement étudiés et documentés [IPCC, 2001]. *Mishra et Coulibaly* [2009] insistent sur l'importance des données hydrométéorologiques qui fournissent une information clef pour les chercheurs et utilisateurs, en contexte de changement climatique global. Elles permettent en particulier d'étudier

les effets de la variabilité du climat sur les bas et hauts débits, le régime des rivières, la distribution spatiotemporelle des précipitations sur le bassin, etc.

A l'heure où le changement climatique représente l'un des enjeux majeurs de ce 21^{ème} siècle, les scientifiques n'ont jamais eu autant besoin de ces données hydrométéorologiques pour mieux comprendre l'évolution des extrêmes hydrologiques et des pénuries en eau futures (dues à une constante croissance de la population mondiale et des demandes en eau). Il est de ce fait important d'assurer la pérennité des réseaux de mesure. En effet, l'un des enjeux majeurs des réseaux hydrométriques est de pouvoir enregistrer des chroniques de débit suffisamment longues pour les études de tendances de long terme. Toutefois, dans un contexte de changement climatique global, les chroniques passées ne sont plus suffisantes pour connaître le comportement hydrologique futur du bassin versant et pouvoir modéliser les conséquences du changement climatique. Seules les données de débit mesurées nous renseignent de manière fiable sur ces changements. Pour cela, nous avons de plus en plus besoin de stations hydrométriques pérennes pouvant enregistrer de longues chroniques de débit, et un maximum de variabilité hydroclimatique.

1.3. Coût des réseaux de mesure

L'existence de réseaux de mesure hydrométéorologiques suffisamment denses est essentielle pour limiter les incertitudes des calculs hydrologiques. Cependant, ces données ont un coût, qui est d'autant plus important qu'il n'existe pas encore de « robot métrologue » : les interventions humaines requises sont nombreuses et coûteuses (coûts d'installation additionnés aux coûts d'entretien et de maintenance).

1.3.1 Coût des réseaux hydrométéorologiques d'Electricité de France

A titre d'exemple, les coûts moyens des réseaux hydrométéorologiques de l'Electricité De France sont les suivants (Rémy Garçon, EDF, communication personnelle) :

- un poste pluviométrique reviendrait à EDF à 4700 € par an dont 1400 pour l'amortissement de l'investissement ;
- pour une station hydrométrique, ce coût s'élèverait à environ 15000 € par an. Le ratio investissement/exploitation serait voisin de celui des postes climatologiques (plus de 70% pour la gestion). Ce coût est toutefois influencé par une localisation majoritaire des stations en moyenne et haute montagne. De plus, des différences de coût sont à noter entre des profils stables et des stations dont le détarage du lit est fréquent.

1.3.2 Coût du réseau pluviométrique de Météo-France

Pour le réseau pluviométrique de Météo-France (Mathieu Creau, Météo-France, communication personnelle), un site automatique (composé bien souvent d'un pluviomètre et d'autres instruments de mesure) récoltant des données toutes les 6 minutes ou toutes les heures par réseau téléphonique filaire ou GSM, représente un coût initial de 4500 € en matériel, 6000 € en infrastructure (énergie, accès télécom, clôture, support, préparation du terrain...) et environ 1000 € en main d'œuvre, soit un total d'environ 11 500 €. En fonctionnement, un site de mesures coûte en moyenne sur tout le parc environ 3000 €/an (en matériel et frais de personnel) affecté à la maintenance, réparation, étalonnage du pluviomètre, etc. A noter toutefois, que ce coût de gestion englobe la maintenance de toutes les mesures effectuées sur ce site, c'est-à-dire généralement la pluie, le vent, l'humidité de

l'air, la température, etc. De plus, ces coûts de maintenance peuvent varier d'une zone à une autre. Par exemple, une station située en zone de relief aura un surcoût par rapport à l'investissement initial pouvant aller de 1000 à 7000 €, pour la mise en place par exemple d'un réchauffage du pluviomètre permettant la mesure des précipitations solides en temps réel.

Un poste manuel de réseau Météo-France, relevé quotidiennement par un bénévole, coûtera moins cher : 200 € d'investissement initial correspondant à l'achat du pluviomètre et environ 200 €/an pour l'indemnisation du bénévole ainsi que la visite annuelle de contrôle et d'accompagnement et l'achat de diverses fournitures comme le carnet d'observation par exemple.

1.3.3 Coût du réseau hydrométrique national

Le réseau hydrométrique national compte à ce jour 2800 stations de mesure actives. Le budget total annuel des services d'hydrométrie s'élève à environ 6 millions d'euros (Rachel Puechberty, Schapi, communication personnelle).

La création d'une station hydrométrique coûte initialement environ 25 000 € et son coût annuel de fonctionnement (maintenance, téléphone, ligne EDF...) revient à environ 1000 €. Ces coûts ne tiennent toutefois pas compte des frais de personnel.

Le coût des opérations de modernisation et de sécurisation des stations et des sites de mesure peut être évalué à environ 600 €/station/an. L'achat de matériel concernant le jaugeage de débit peut lui s'élever à 200 €/station/an. Toutefois, ces deux derniers coûts résultent d'une moyenne réalisée sur tout le parc national. Toutes les stations ne nécessitent pas obligatoirement ces maintenances chaque année.

1.4. L'idée de rationalisation des réseaux de mesure

Ces réseaux peuvent ainsi paraître onéreux pour leurs gestionnaires, d'où la réflexion sur leur rationalisation. Mais, cette rationalisation, qui se traduit souvent par une baisse de la densité du réseau, rend notre connaissance du cycle hydrologique plus incertaine, et peut notamment augmenter les incertitudes des calculs hydrologiques.

En France, la question de la conception optimale des réseaux hydrométéorologiques n'est pas nouvelle. Dans les années 1960, le Ministère de l'Agriculture en partenariat avec l'ORSTOM (Office de la Recherche Scientifique et Technique Outre-Mer, organisme aujourd'hui devenu l'IRD) a lancé des études pour concevoir de manière rationnelle un réseau hydrométrique, fournissant un maximum d'informations de qualité au moindre coût [Dubreuil, 1968]. Dans le cadre des attributions du Ministère de l'Agriculture en matière d'inventaire des ressources en eau superficielle, il s'agissait d'étendre harmonieusement le réseau existant et à l'échelle nationale, pour passer de 500 à 1000 stations. L'ORSTOM s'est également intéressé aux problèmes de rationalisation des réseaux. Roche [1967] explique dans son étude l'intérêt de mettre en place des réseaux de mesure hydrométéorologiques et présente une méthode de rationalisation de ces réseaux. Dans la continuité, Herbaud [1969] propose un modèle de rationalisation du réseau pluviométrique pour un délai et un coût global fixés, dans le but hydrologique d'estimer la hauteur moyenne annuelle des précipitations ponctuelles.

En 2006, le Ministère des transports, de l'équipement, du tourisme et de la mer et le Ministère de l'écologie et du développement durable ont mis en place une circulaire visant à organiser l'hydrométrie des services de l'Etat et plus précisément des DIREN et des Services de Prévision des Crues (SPC), dans le but d'unifier le réseau de mesure hydrométrique et d'améliorer l'efficacité du service rendu en optimisant l'utilisation des moyens humains et financiers existants. Cette circulaire est destinée aux préfets coordonnateurs des six grands bassins français et définit un schéma cible pour l'organisation des services déconcentrés et des structures nationales.

D'autre part, les SPC utilisent également des données pluviométriques en temps réel provenant de sources variées et financées par des organismes différents (le réseau de la Direction Générale de la Prévention des Risques (DGPR) du ministère en charge du développement durable et de trois réseaux Météo-France (SALAMANDRE, RADOME et PATAC)). Dans le but de pérenniser les postes indispensables à l'exercice de prévision des crues et de créer des postes dans les zones où l'information pluviométrique manquait, le SCHAPI a lancé une étude à l'échelle nationale visant à mutualiser ces réseaux pluviométriques, tout en tenant compte de l'amélioration constante de la couverture radar [Helloco, 2010]. Ceci devait alors conduire à rationaliser ces réseaux tout en satisfaisant les besoins hydrologiques pour la prévision des crues et à valoriser les données produites. Cette étude a permis de définir un réseau de « sauvegarde ». Ce réseau recommandé est composé de postes à pérenniser et d'installations nouvelles. L'étude conclut également que l'utilisation du radar en prévision des crues ne permet pas, pour le moment, de diminuer la densité spatiale des réseaux pluviométriques.

Cette circulaire et cette étude soulèvent la question de la cohérence des réseaux installés par différents gestionnaires pour leurs besoins propres, sans qu'il y ait vraiment une mise en cohérence entre eux. Il semble ainsi important, afin de réduire le coût de gestion et de maintien de ces réseaux de mesure, de proposer des solutions de mutualisation des divers réseaux.

Par ailleurs, dès les années 1960, l'Organisation Mondiale de la Météorologie (OMM) recommandait des densités minimales pour les réseaux hydrométriques en admettant qu'un réseau optimal devait comprendre autant de stations sur grands et petits bassins [WMO, 1965]. Actuellement, les densités minimales de stations de mesure dont celles de pluie et de débit, recommandées par l'OMM pour des zones climatologiques et géographiques différentes, sont présentées dans le Tableau 1 [WMO, 2008]. Pour l'OMM, un réseau minimal est celui qui permettra d'éviter de graves lacunes dans le développement et la gestion des ressources en eau en accord avec le niveau global de développement économique du pays et ses besoins environnementaux.

Toutefois, dans les décennies 1980 et 1990, pour des raisons principalement économiques, de considérables réductions de densité de réseaux hydrologiques ont été constatées dans de nombreux pays : Canada (21%), Finlande (7%), Nouvelle-Zélande (20%) [Pearson, 1998] et Etats-Unis (6%) ; et ces réductions se poursuivent [WMO, 2008].

Tableau 1. Densités minimales des stations de mesure recommandées par l'OMM (surface en km² par station) [WMO, 2008].

Physiographic unit	Precipitation		Evaporation	Streamflow	Sediments	Water quality
	Non-recording	Recording				
Coastal	900	9 000	50 000	2 750	18 300	55 000
Mountains	250	2 500	50 000	1 000	6 700	20 000
Interior plains	575	5 750	5 000	1 875	12 500	37 500
Hilly/undulating	575	5 750	50 000	1 875	12 500	47 500
Small islands	25	250	50 000	300	2 000	6 000
Urban areas	–	10–20	–	–	–	–
Polar/arid	10 000	100 000	100 000	20 000	200 000	200 000

1.5. La valeur économique des réseaux de mesure

1.5.1 La valeur économique des réseaux de mesure et des outils hydrométéorologiques en général

Un grand nombre d'études sont justifiées par une pression pour réduire les coûts de fonctionnement des services de l'Etat, dans la vague de privatisation et de corporatisation des services gouvernementaux « traditionnels » [Anaman *et al.*, 1998]. Des analyses coût-bénéfice sont nécessaires pour justifier les investissements.

Le bon sens exige que les coûts des réseaux soient comparés aux bénéfices associés aux calculs hydrologiques. Même si cette thèse n'a pas de vocation économique, il reste intéressant de passer en revue les travaux qui existent dans ce domaine :

- Au Royaume-Uni, en 2000, les bénéfices annuels associés au réseau hydrométrique national sont estimés à 21 millions £ (pouvant osciller entre 11 et 60 millions £). Le coût du réseau hydrométrique est quant à lui estimé à 9 millions £ [Walker, 2000], soit un ratio d'environ 1:2.
- Une revue sur le rapport d'activité du réseau hydrométrique dans la région de Colombie Britannique au Canada réalisée par le Ministère de la gestion de la ressource durable, estime un ratio coût/bénéfice de leur réseau composé de 461 stations, de 1:19. Cette revue précise également que des pertes considérables (de l'ordre de 82 millions de dollars) pourraient être évitées si la densité du réseau était au moins doublée [Azar *et al.*, 2003].

D'autres études ont porté sur l'estimation du rapport coût/bénéfice des services hydrométéorologiques en général, qui d'une certaine mesure dépend de la qualité et la quantité des données dont ils disposent. Freebairn et Zillman [2002] passent en revue les différentes méthodes proposées dans la littérature afin d'évaluer les bénéfices économiques associées aux services météorologiques et évaluent en particulier leur applicabilité pour la prise de décision concernant les dépenses associées à ces services. Les études suivantes donnent quelques valeurs de ratio coût/bénéfice répertoriées dans la littérature :

- *Mason* [1966] estime que le service météorologique anglais (« the Meteorological Office ») présente un rapport coût/bénéfice d'environ 1:20, avec une valeur économique du service national de prévision météorologique variant de 50 millions à 100 millions £ par an, pour un coût évalué à 4 millions £ par an.
- R. de Guzman, directeur de l'office de planification stratégique de l'OMM estime lors de sa participation à la conférence internationale « Secure and sustainable living : Social and Economic Benefits of Weather, Climate and Water Services » en 2007, « qu'un dollar dépensé dans la prévision d'une catastrophe naturelle peut empêcher sept dollars de pertes économiques liées à cette catastrophe ».
- Différentes études citées par *Frei* [2010] avancent des ratios coût-bénéfice de l'ordre de 1:4 à 1:6 [*Anaman et al.*, 1995; *Lazo et Chestnut*, 2002; *Leviäkangas et al.*, 2007] pour les services météorologiques occidentaux. Pour la Suisse, les bénéfices de la prévision climatique sont de l'ordre de centaines de millions de Francs Suisses [*Frei*, 2010] (à comparer avec le budget de MétéoSuisse qui était de 77 millions, et un total d'environ 100 millions de FS si on y ajoute les producteurs privés). Les Suisses estiment un ratio coût/bénéfice de 1:5.
- D'autres études ont évalué le consentement à payer pour des services tels que l'alerte cyclonique, avec une intéressante asymétrie du consentement à payer pour le service (env. 40 \$) et du consentement à recevoir une somme d'argent pour accepter de se passer du même service (env. 200 \$) [*Anaman et al.*, 1998].

De façon plus opérationnelle, durant sa thèse effectuée à EDF, Houdant [2004] a mené une étude sur la valorisation économique des prévisions opérationnelles (celles fournies à l'utilisateur). Il présente deux exemples de valorisation de prévisions : déterministes à court terme et probabilistes à long terme. La valorisation des prévisions déterministes à court terme concernait l'aménagement hydroélectrique de l'Ain équipé de six usines hydroélectriques. La gestion de cette chaîne est complexe pour plusieurs raisons : le risque de saturation de la chaîne de production, l'obligation de maintenir une cote de réservoir pour les activités récréatives et un débit minimal à l'aval du dernier aménagement. L'étude montre que les prévisions déterministes permettent à l'utilisateur ou l'exploitant de réaliser des gains substantiels (de l'ordre de 1 à 2%). Le deuxième exemple de valorisation de prévisions est un cas de prévisions probabilistes à long-terme sur le lac de Serre-Ponçon. Ces prévisions concernent l'apport de remplissage de réservoirs saisonniers. Les enjeux liés au barrage de Serre-Ponçon sont de plusieurs types : production électrique, enjeux touristiques (obligation de maintenir une cote admissible pour les activités touristiques, assurer un débit minimal durant l'été en amont). Les résultats ont montré qu'une meilleure gestion était obtenue avec les prévisions saisonnières qu'avec une gestion basée sur l'historique. Ceci représente un gain économique sur l'utilisation du barrage en tant que producteur hydroélectrique et pour l'activité touristique du lac de Serre-Ponçon (de 1 à 8% pour une année à fort enneigement, 5% pour une année « sans neige »).

Par ailleurs, au cours de sa thèse réalisée à Irstea, *Zalachori* [2013] a quant à elle essayé de valoriser les prévisions hydrologiques d'ensemble pour la gestion de la production d'énergie hydroélectrique d'une retenue d'eau. Elle a montré que la qualité des prévisions d'ensemble améliorerait les performances de la gestion et la rentabilité économique, en particulier si ces prévisions sont préalablement traitées à l'aide d'outils statistiques.

Palmer [2002] s'est également intéressé à la valeur économique des ensembles de prévision au Royaume-Uni. Les prévisions d'ensemble fournissent un outil quantitatif pour la météorologie et

l'évaluation du risque climatique. L'auteur souligne dans son article l'importance de l'interconnexion quantitative entre les développeurs de modèle (qui s'expriment en erreur quadratique du modèle) et les clients de ces prévisions (qui eux s'expriment davantage en euros ou dollars). Il propose un modèle de décision pour l'utilisateur permettant de relier analytiquement la valeur économique potentielle aux scores classiques de compétences météorologiques.

1.5.2 Optimisation du réseau hydrométrique

Dans la littérature, certaines études s'intéressent à la valeur hydrologique des données de débit, soit la détermination du gain hydrologique que la donnée de débit ajoutée (l'ajout de stations ou l'assurance de la pérennité d'une station) peut apporter.

Tasker [1986] présente une méthode permettant d'identifier un réseau hydrométrique efficace en respectant un budget de fonctionnement spécifié. Le calcul hydrologique visé est l'estimation de la crue cinquantennale au moyen d'une régression Generalized Least Squares (GLS). La fonction objectif permettant d'évaluer le réseau est l'erreur quadratique moyenne de l'échantillonnage moyen. Il ne mesure pas l'erreur due au modèle de régionalisation (erreur sur l'estimation des caractéristiques de débits par la vraie régression γ_r) mais l'erreur due à l'échantillonnage (erreur sur l'estimation de la caractéristique de débits γ_r par la régression de l'échantillon). L'exemple donné dans cette étude concerne un réseau de 31 stations situées dans le sud-ouest du Wisconsin ainsi que 8 stations fictives. L'auteur fait varier le nombre de stations (et donc leur coût total) et sous-échantillonne le réseau complet pour évaluer les erreurs dues à la régression effectuée dans un premier temps à l'aide des 31 stations puis à l'aide de ces 31 stations ainsi que des 8 fictives (stations situées à distance du cas d'étude). Les résultats indiquent que lorsque les 8 nouvelles stations sont ajoutées l'erreur obtenue sur l'estimation du débit cinquantennal est toujours inférieure à l'erreur obtenue avec la régression effectuée avec les 31 stations (Figure 3). Pour l'auteur, une stratégie future de jaugeage appropriée serait d'ajouter des nouvelles stations à des points distants. Toutefois, à budget équivalent, ajouter de nouvelles stations impliquerait de fermer un nombre équivalent d'anciennes stations.

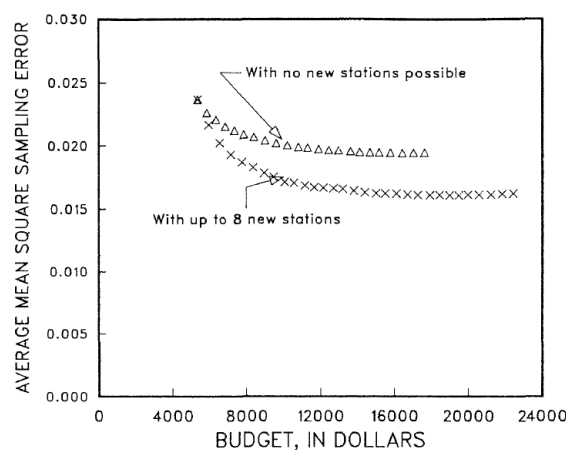


Figure 3. Erreur quadratique moyenne d'échantillonnage en fonction du budget alloué au réseau hydrométrique. Les triangles représentent le cas où les 31 stations sont utilisées pour établir la régression et les croix le cas où les 8 stations sont ajoutées au réseau pour la régression (d'après *Tasker* [1986]).

L'étude de *Cordery et Cloke* [1994] vise à estimer une valeur de la donnée de débit pour dimensionner un système de digues sur trois bassins australiens. Ils insistent sur l'importance d'avoir une longueur de chronique de débit suffisamment longue afin d'avoir une bonne représentation de la population de débits. Les données nécessaires à leur étude sont une longue chronique de débit ou de hauteur d'eau, une relation dommages-hauteur de crue et une relation coût-hauteur de protection. Le bénéfice de stratégie de réduction de crue est évalué par rapport à la valeur des pertes évitées grâce à la stratégie mise en place. Les pertes ici considérées sont la somme des dommages annuels apportés aux infrastructures et aux propriétés, de la perturbation des activités professionnelles et des routes de transport, du coût de nettoyage de débris et des inconvénients généraux à tous les utilisateurs de la zone inondée. Pour évaluer la valeur de la collecte de données, les auteurs ont développé une stratégie utilisant toutes les données collectées à une date et ont comparé celle-ci avec une stratégie développée en utilisant toutes les données qui seraient collectées à une date future. Par exemple, une stratégie serait d'utiliser 20 ans de données de 1936 à 1955, et de la comparer à une stratégie déterminée sur 30 ans de données de 1936 à 1965. Les auteurs regardent alors l'apport de ces 10 ans supplémentaires de données. L'étude a montré que les données recueillies pour la planification des projets ont un retour annuel sur investissement jusqu'à 80 fois les coûts d'exploitation. Les auteurs soulignent également le besoin considérable d'investissements publics dans l'installation et l'exploitation de stations de jaugeage de rivières pour collecter et archiver des données de haute qualité.

Mishra et Coulibaly [2010] ont quant à eux développé une méthode permettant d'évaluer les réseaux hydrométriques afin d'identifier les stations de débit essentielles et les zones de très faible densité de réseau, au sein de bassins majeurs répartis dans différentes provinces du Canada. Les résultats de l'étude montrent que la plupart des principaux bassins canadiens contiennent des réseaux hydrométriques déficients et que certains de ces bassins ne possèdent pas suffisamment de données historiques (moins de 20 ans de données de débit mesurées). Les auteurs concluent également que les zones de faible densité varient selon le bassin étudié, et dépendent du nombre de stations et de leur distribution spatiale.

1.6. Synthèse

Nous venons de voir que l'échantillonnage spatio-temporel des données de pluie et de débit peut avoir un impact sur les applications hydrométéorologiques (prévisions météorologiques et hydrologiques, étude des tendances à long-terme, études en lien avec la problématique du changement climatique, etc.). Toutefois, ces dernières années ont été marquées par une pression de la part des gestionnaires des réseaux hydrométéorologiques pour réduire leurs coûts de fonctionnement et ceci se traduit le plus souvent par une baisse des densités de ces réseaux de mesure.

La réduction des réseaux de mesure hydrométéorologiques est un moyen pour les gouvernements de faire des gains économiques à court terme, mais ceux-ci peuvent se traduire par une perte économique à long-terme très importante. Les ratios coût/bénéfice passés en revue dans cette étude témoignent de l'importance économique à long terme de ces réseaux. Les conséquences hydrologiques de long terme résultant d'une rationalisation extrême peuvent être nombreuses et variées : des mauvaises décisions concernant l'estimation des ressources en eau pour l'eau potable, la gestion hydroélectrique, les dimensionnements d'ouvrages entraînant dans un cas des pertes

matérielles et/ou humaines et dans l'autre une perte économique due à un surdimensionnement, les prévisions des extrêmes, les alertes aux populations etc.

Ainsi, cette rationalisation peut rendre notre connaissance du cycle hydrologique plus incertaine, et peut notamment augmenter les incertitudes des calculs hydrologiques. Toutefois, quantifier cette augmentation présente un certain nombre de difficultés, car elle dépend de nombreux facteurs : de l'objectif hydrologique, des outils utilisés et des caractéristiques des bassins versants étudiés. Dans la suite de ce mémoire, nous nous proposons d'étudier l'impact de la densité des réseaux hydrométriques (en diminuant la densité spatiale du réseau et en testant le contenu temporel des séries disponibles de débit), mais également celui de la densité des réseaux pluviométriques (en diminuant la densité spatiale du réseau) sur les performances de divers calculs hydrologiques.

Partie I

Impact de la densité du réseau hydrométrique sur les calculs hydrologiques

Chapitre 2.

Impact de la densité du réseau hydrométrique en hydrologie : état de l'art

2.1. Introduction

De nombreuses études portant sur la sensibilité des calculs hydrologiques à la densité du réseau hydrométrique ont été menées. On peut distinguer deux cas :

- les bassins versants étudiés ne sont pas jaugés (données de débit non disponibles) : dans ce cas, les calculs hydrologiques s'appuient notamment sur une information issue de bassins versants voisins, dont la proximité spatiale ou la similarité physique jouent un rôle prépondérant. En conséquence, les calculs hydrologiques seront sensibles à la **densité spatiale** des stations servant à l'estimation du débit au point non jaugé.
- les bassins sont jaugés : les études s'intéressent alors essentiellement à l'impact de la durée et de la qualité des chroniques sur les performances des modèles hydrologiques. On parlera ici de **densité temporelle** pour désigner la longueur et le contenu (en termes de fréquence d'observation ou de gamme de débit échantillonnée) des chroniques de débit.

Dans ce chapitre, nous dressons tout d'abord un état de l'art de l'impact de la densité spatiale du réseau hydrométrique en hydrologie, en soulignant l'importance d'un réseau de mesure relativement dense pour la modélisation des bassins versants non ou peu jaugés (bassins ne disposant pas de chroniques de données hydrométriques suffisamment longues pour caler les paramètres d'un modèle hydrologique ou d'estimer certaines caractéristiques hydrologiques).

Par la suite, nous analyserons les études menées sur l'impact de la durée des chroniques de débit, en abordant dans un premier temps leur impact sur l'identification des paramètres des modèles hydrologiques, puis celui de la longueur des chroniques de débit pour le calcul des débits caractéristiques (typiquement le module ou des quantiles de crue ou d'étiage).

2.2. Impact de la densité spatiale du réseau hydrométrique en hydrologie

2.2.1 Efficacité de différentes approches d'estimation de paramètres sur bassins versants non jaugés

Les modèles hydrologiques utilisés dans le cadre de cette thèse ont des paramètres qui ne peuvent pas être obtenus directement à partir de caractéristiques physiques du bassin versant : ils exigent un processus de calage basé sur des chroniques de débit observé pour l'identification de leurs paramètres. Cependant, malgré le besoin de données locales de calage, beaucoup de questions hydrologiques se posent en situation non jaugée, c'est-à-dire quand le bassin versant concerné ne possède pas de mesure de débit à son exutoire. Ces bassins exigent alors un traitement spécifique.

Depuis les années 1970, les hydrologues se sont penchés sur la façon dont ils pouvaient estimer les paramètres de leurs modèles sans avoir besoin de les caler [James, 1972; Magette et al., 1976]. Les approches correspondantes sont généralement appelées approches de régionalisation [Gottschalk et al., 1979]. Les avancées sur la régionalisation réalisées dans le cadre de l'initiative de recherche « Prediction in ungauged basins » (PUB) de l'AISH (2003-2012) ont été synthétisées par Hrachowitz et al. [2013]. Cette synthèse détaille la façon dont l'information peut être transférée d'un bassin jaugé vers un bassin non jaugé.

Parmi les principales approches de régionalisation, on peut distinguer deux grands groupes :

- 1) celles utilisant le principe de **similarité physique**, c'est-à-dire faisant l'hypothèse que les bassins avec des caractéristiques physiques similaires auront une réponse hydrologique similaire également. Les approches basées sur des régressions linéaires multiples, à partir desquelles les paramètres des modèles hydrologiques sont estimés indépendamment à partir de caractéristiques physiques de bassin, rentrent également dans cette catégorie.
- 2) celles basées sur la **proximité spatiale** entre le bassin d'intérêt et les bassins jaugés, basées sur l'hypothèse que les conditions physiques environnantes sont similaires et transférables entre bassins voisins.

Un certain nombre d'études se sont ainsi intéressées à la comparaison de ces méthodes. On se propose d'en citer ici quelques-unes :

- *Merz et Blöschl* [2004] ont montré que les méthodes basées sur la proximité spatiale fournissaient de meilleurs résultats que celles basées sur les caractéristiques physiques du bassin.
- *Parajka et al.* [2005] se sont intéressés au transfert des paramètres de modèles hydrologiques. Ils ont montré que deux méthodes étaient les plus performantes : l'approche par krigeage (basée sur la corrélation spatiale des paramètres) et l'approche basée sur la similarité physique.
- *Oudin et al.* [2008b] ont également comparé les résultats de trois méthodes de régionalisation sur deux modèles et 913 bassins français. L'étude a montré que dans le cas d'un réseau hydrométrique dense, comme c'est le cas en France, la proximité spatiale fournit la meilleure solution de régionalisation. L'approche par régression, qui établit des relations entre valeurs des paramètres et descripteurs du bassin, est la moins satisfaisante, avec des résultats très proches de ceux obtenus en utilisant un jeu de paramètres médian sur tout l'échantillon de bassins.
- *Samuel et al.* [2011] ont comparé quatre méthodes de régionalisation dans une province de l'Ontario couvrant 1 million de km² (bassins versants de surface comprise entre 100 et 100 000 km²). Ces dernières comprennent les approches classiques citées ci-dessus, ainsi qu'une approche couplant la proximité spatiale (pondération par l'inverse de la distance) et la similarité physique. L'étude révèle que les méthodes basées sur la proximité spatiale (krigeage, et pondération par l'inverse de la distance) sont plus performantes que celles basées sur la régression et la similarité physique. Toutefois, l'approche couplant proximité spatiale et similarité physique semble montrer une amélioration des performances.
- *Arsenault et Brissette* [2014] ont évalué l'incertitude de trois méthodes de régionalisation. Dans leur étude, la meilleure option de régionalisation est celle basée sur la similarité physique, suivie de près par celle basée sur la proximité spatiale. Comme dans les études précédentes, la méthode basée sur la régression linéaire multiple donne les moins bons résultats car les paramètres sont pour la plupart très faiblement corrélés aux caractéristiques des bassins.

Parajka et al. [2013] ont proposé une synthèse des études récentes, s'appuyant sur une analyse de 34 études publiées dans la littérature et concernant 3874 bassins. Cette analyse a pour but de comparer ces différentes méthodes de régionalisation dans divers cas d'étude. Les auteurs concluent que les méthodes de régionalisation montrent des performances similaires avec une dispersion

importante au sein de chaque groupe de régionalisation. Les résultats montrent tout de même que les méthodes de régression ont tendance à présenter les plus faibles performances lorsque celles-ci sont comparées aux autres méthodes dans une même région.

Nous venons ainsi de voir que les bassins versants donneurs sont indispensables pour la modélisation des chroniques de débit des bassins non jaugés. Mais les résultats de ces approches de régionalisation vont dépendre de la densité du réseau de bassins donneurs, et plus généralement du réseau hydrométrique existant.

2.2.2 Etudes sur le nombre de bassins voisins utilisés pour le transfert d'information à des bassins non jaugés

Une question essentielle pour les bassins non jaugés

Les paramètres des modèles hydrologiques peuvent varier d'un bassin à un autre, même lorsque ceux-ci sont jugés similaires en termes de caractéristiques physiographiques et climatiques, ou quand ils sont proches spatialement. Tenir compte de plusieurs bassins voisins permet de lisser ces différences imprévisibles et de réduire les erreurs d'estimation des paramètres du bassin non jaugé, afin d'obtenir des paramètres moyens pour la zone d'étude. Le nombre et les caractéristiques des bassins voisins pris en compte pour cette extrapolation ont un effet sur l'estimation des paramètres et aussi sur les performances obtenues en simulation de débit. Toutefois, peu d'études ont été menées à notre connaissance sur cette problématique.

Dans l'étude d'*Oudin et al.* [2008b] citée précédemment, les auteurs ont fait varier le nombre de bassins donneurs de 1 à 100. Ils ont montré que pour les deux méthodes de régionalisation (basées sur la proximité spatiale et la similarité physique), seulement cinq bassins donneurs sont nécessaires pour obtenir de bons résultats (Figure 4). Les auteurs interprètent cela de la façon suivante : prendre plusieurs bassins donneurs permet d'éviter de trop fortes erreurs dans la simulation de débit en lissant la réponse avec d'autres sources. Cependant, lorsque l'on augmente le nombre de bassins donneurs à plus de cinq, (bassins qui sont donc de plus en plus loin du bassin non jaugé), l'efficacité du modèle décroît presque linéairement. Notons que dans cette étude, aucune pondération de la distance n'était appliquée. Seuls les n donneurs les plus proches étaient retenus pour le transfert des paramètres.

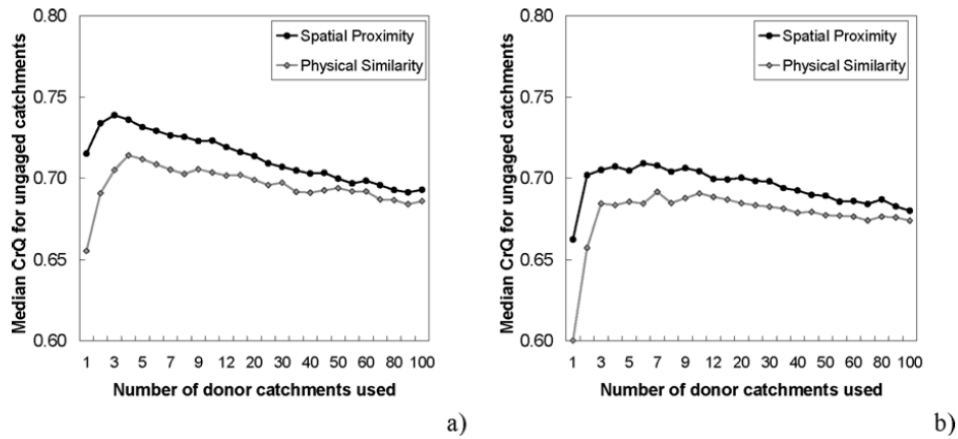


Figure 4. Impact du nombre de bassins voisins utilisés pour la simulation de débit d'un bassin non jaugé sur l'efficacité du modèle hydrologique. Méthodes de régionalisation basées sur la proximité spatiale et sur la similarité physique pour (a) GR4J et (b) TOPMO (d'après *Oudin et al. [2008b]*).

Arsenault et Brissette [2014], cités précédemment, ont fait également varier le nombre de bassins donneurs. Les auteurs ont montré que le nombre de bassins donneurs optimum varie en fonction de l'approche de régionalisation utilisée. Toutefois, ils soulignent l'importance d'utiliser plusieurs donneurs. Pour le modèle HSAMI, le nombre optimum de donneurs se trouverait entre 4 et 7 (Figure 5). Cependant, dépasser ce seuil entraîne une chute des performances de régionalisation à cause de l'ajout de donneurs trop éloignés (en termes de géographie ou de caractéristiques). D'autre part, l'étude montre qu'écarter les bassins dits « mauvais » en termes de performances en régionalisation améliore les performances moyennes sur les bassins non jaugés.

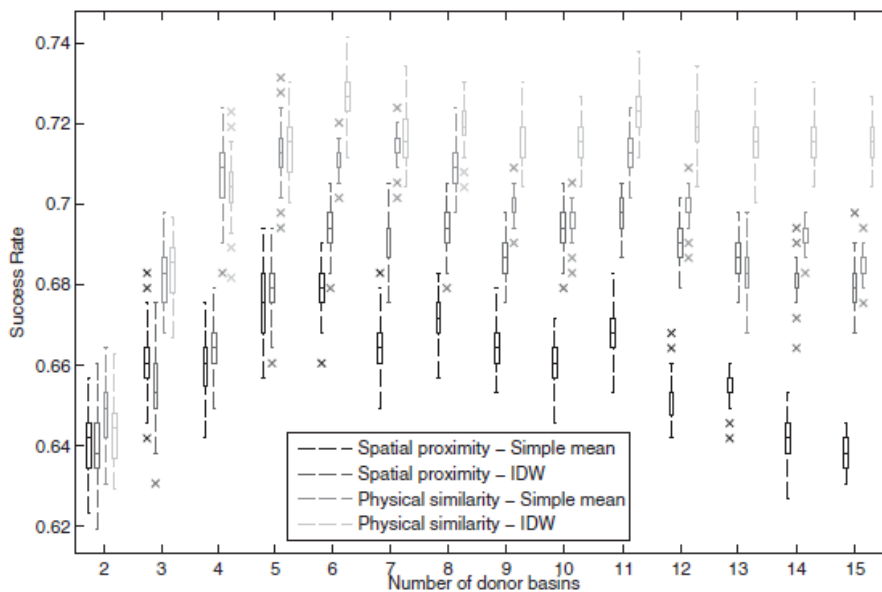


Figure 5. Distribution du taux de réussite des méthodes de régionalisation basées sur la similarité physique et la proximité spatiale en utilisant la simple moyenne des sorties du modèle ou la moyenne des sorties du modèle pondérée par l'inverse de la distance des donneurs (d'après *Arsenault et Brissette [2014]*).

Enfin, *Boldetti et al.* [2010] ont proposé une nouvelle méthodologie pour l'identification des bassins jaugés qui ne devraient pas être pris en compte en tant que « donneurs » dans le contexte d'un exercice de régionalisation. Dans cette étude, ils ont en partie testé l'impact de la densité des stations donneuses. Les résultats ont montré qu'avec un réseau moins dense, les améliorations suite au rejet de certains donneurs jugés indésirables, étaient moins importantes. Le pourcentage optimal de donneurs était réduit également et les performances chutaient plus rapidement quand trop de bassins étaient écartés.

Dans le domaine de la prévision hydrologique d'ensemble des bassins non jaugés, le nombre de bassins donneurs utilisés afin d'obtenir une prévision sur le bassin cible peut également avoir son importance. Les résultats de la thèse de *Randrianasolo* [2012] montrent que les scores de prévision d'ensemble sont améliorés lorsque le nombre de voisins augmente puis se stabilisent autour de 10 à 15 voisins. Au-delà de 15 à 20 bassins voisins, certains scores présentent une dégradation des performances.

Au Canada, un important plan de réduction des mesures hydrologiques a commencé au milieu des années 1990 en raison de pressions budgétaires au niveau fédéral. Dans ce contexte, l'étude de *Spence et al.* [2007] a eu pour objectif de quantifier l'effet de l'arrêt des mesures hydrométriques sur les capacités de prévision des débits, effectuée par méthode statistique, dans la région du Nord du Canada. La méthode utilisée trouve des voisins hydrologiques ou des groupes de bassins homogènes pour corrélérer les mesures de débit avec les caractéristiques physiographiques et climatiques de chaque bassin. Les résultats indiquent que lorsque 12 stations sur 34 sont fermées sur un domaine avoisinant les 1 500 000 km², les erreurs d'extrapolation augmentent de 16% pour tous les quantiles de débit. L'incertitude reste la même pour les quantiles moyens de débit mais double pour les quantiles de débits extrêmes.

Ainsi, la densité spatiale des stations hydrométriques semble avoir un impact sur les performances des modèles hydrologiques dans le cas de bassins non jaugés. Dans certains cas [*Oudin et al.*, 2008b], un nombre limité de bassins est suffisant pour obtenir de bons résultats, mais il serait dangereux d'étendre ce résultat de France au Canada, tant la densité du réseau varie entre ces deux pays (exemple : au Canada [*Arsenault et Brissette*, 2014], c'est la similarité physique qui est meilleure alors qu'en France c'est la proximité spatiale qui tendrait à être meilleure). On peut raisonnablement supposer que lorsque la densité du réseau diminue, les performances des méthodes de proximité géographique vont à un moment passer sous celles des méthodes de similarité physique.

Extension des enregistrements de débit basée sur les stations hydrométriques des bassins voisins (approches débit-débit)

Andréassian et al. [2012a] se sont intéressés à l'estimation de débits à l'aide d'une approche débit-débit. Leur étude a eu pour but de déterminer la manière de généraliser la procédure des bassins appariés, afin qu'elle puisse être appliquée à n'importe quelle paire de bassins voisins, en comparant en même temps un modèle pluie-débit (GR4J) et un modèle débit-débit. Le modèle débit-débit est évalué de deux façons : le cas où les paramètres du meilleur bassin sont utilisés et celui où la moyenne arithmétique des simulations issues de la meilleure combinaison de voisins variant

successivement de 1 à 7 bassins est utilisée. Les résultats révèlent qu'un modèle basé sur des voisins multiples permet un net gain de performances du modèle en validation (Figure 6). Toutefois, dans les deux cas, l'efficacité du modèle débit-débit diminue quand le réseau devient moins dense, mais au début avec une faible diminution (il y a peu de différences entre le réseau entier et celui réduit de 20%).

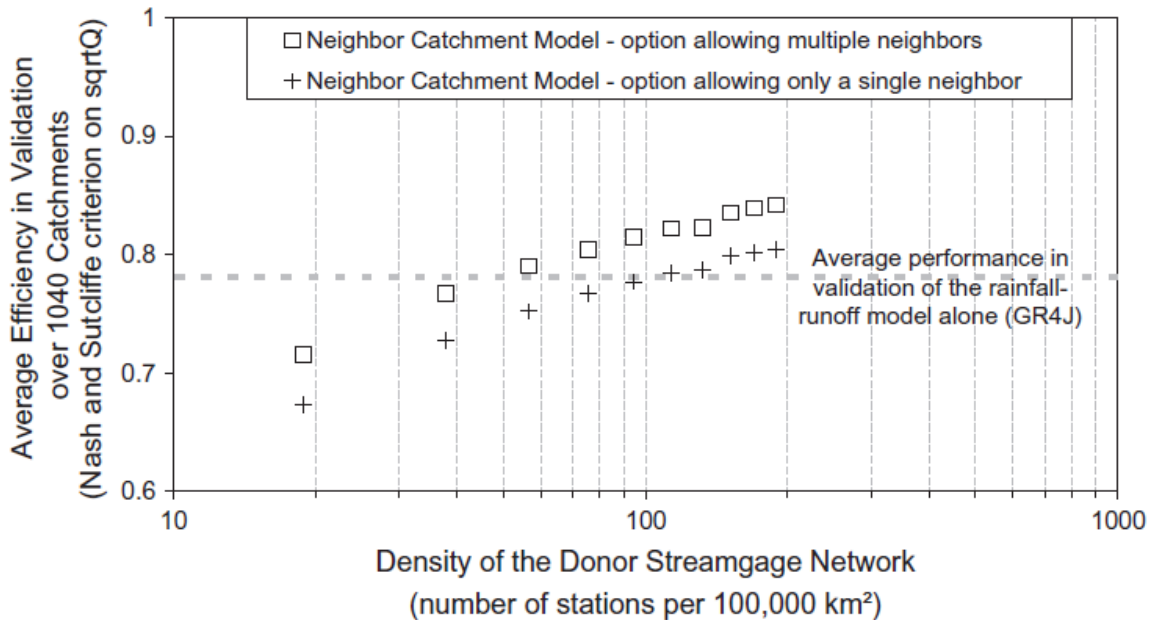


Figure 6. Impact de la réduction de la densité de voisins sur les performances du modèle de bassin voisin en validation (d'après Andréassian et al. [2012a]).

Dans un autre contexte, *Nicolle et al.* [2013] ont proposé une méthode permettant à des bassins ne disposant pas de longues chroniques de données hydrométriques, d'estimer des caractéristiques de débit d'étiage telles que le débit minimum mensuel de période de retour 5 ans (QMNA5), à partir de longues chroniques disponibles sur des bassins voisins. Cette méthode combine deux approches : une première méthode à base climatique et une seconde basée sur la proximité spatiale des bassins versants. Au cours de leur étude, les auteurs ont testé l'impact de la densité des bassins voisins sur l'efficacité de la méthode. Ils montrent que lorsque la distance des plus proches voisins augmente, les performances de la méthode basée uniquement sur la proximité spatiale diminuent. La combinaison avec la méthode à base climatique permet de limiter cette baisse de performances (Figure 7).

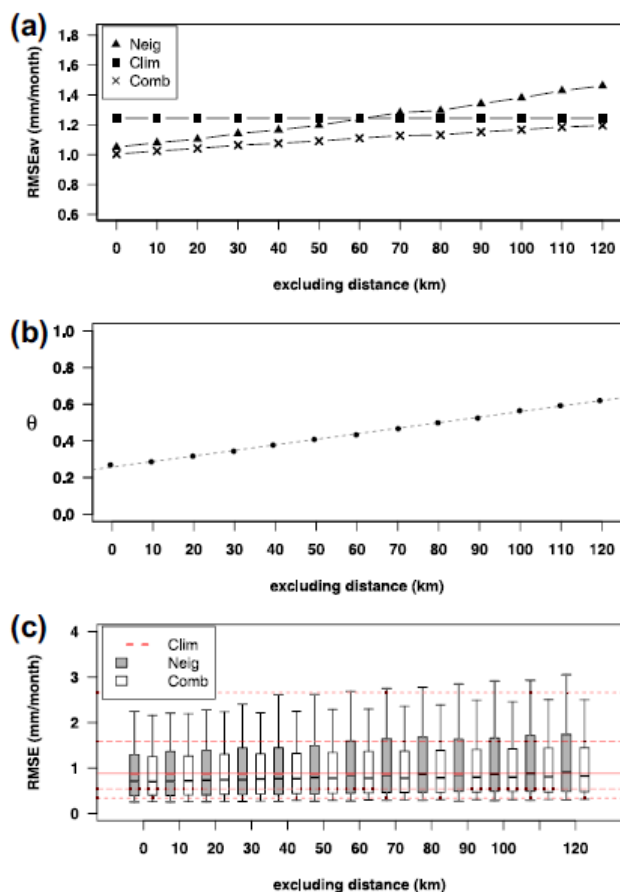


Figure 7. (a) Performances et (c) distributions des performances de l'ajustement climatique (Clim), l'ajustement à l'aide des voisins (Neig) et la combinaison de ces deux méthodes (Comb) en fonction de la distance d'exclusion des bassins donneurs ; (b) θ valeurs du paramètre de la méthode combinée en fonction de la distance d'exclusion des bassins donneurs (d'après *Nicolle et al. [2013]*).

Ainsi, la diminution de la densité du réseau hydrométrique amènerait une baisse des performances des calculs hydrologiques. Cependant, la question est de savoir en dessous de quel niveau de performance le modèle est jugé inutilisable, car présentant trop d'erreurs dans l'estimation du débit.

L'étude effectuée au Canada montre également que la fermeture de certaines stations hydrométriques clé peut dégrader très nettement les estimations de débit, ce qui peut représenter un enjeu majeur lorsque l'on s'intéresse à la détermination des quantiles extrêmes, par exemple pour des objectifs de dimensionnement d'ouvrages.

2.3. Impact de la longueur des chroniques de débit en hydrologie

2.3.1 Longueur des chroniques de débit pour le calage des paramètres

L'évaluation de la sensibilité des modèles hydrologiques à la quantité d'information disponible pour le calage des paramètres est un sujet qui a été très étudié en hydrologie. La réussite de toute procédure de calage est hautement dépendante de la nature (qualité et quantité) des données utilisées. Généralement, on cherche à avoir des données contenant des conditions climatiques et de débit suffisamment variées pour donner une bonne représentativité de leur variabilité naturelle. Il a

souvent été suggéré que les données de calage devraient être les plus représentatives possibles de la variété des phénomènes se produisant sur les bassins. Habituellement, de longues chroniques sont préférées aux courtes, car elles devraient mener à un meilleur calage, dans le sens où elles contiendraient davantage d'informations hydrologiques diverses pour pouvoir caler les paramètres (avec une probabilité plus élevée d'avoir à la fois des données représentatives de conditions humides et sèches).

Andréassian et al. [2012b] discutent de la qualité des calages des paramètres des modèles. L'un des problèmes majeurs lors du calage est que le jeu de paramètres calé sur une période, ne soit pas le jeu de paramètres optimum sur une autre période. L'une des principales raisons serait le manque d'exhaustivité des informations collectées durant la période de calage, qui ne couvrirait alors pas toutes les gammes possibles de conditions climatiques (dû à une période de calage souvent trop courte).

Certaines études ont ainsi analysé l'effet de la longueur des chroniques de débit (et donc indirectement de pluie) en calage sur les performances du modèle alors que d'autres ont plutôt testé la continuité de ces données, ou encore les caractéristiques climatologiques de ces données de calage (années humides/années sèches). Mais les conclusions sont très diverses.

Gan et al. [1997] ont conclu que la longueur des séries de données n'était pas cruciale dans la modélisation de débit, à condition qu'elle ne soit pas inférieure à une année hydrologique, et que les données utilisées contiennent "assez d'informations" pour caler les paramètres. *Sorooshian et al.* [1983] ont également conclu que les performances des modèles étaient peu sensibles à la variabilité et à la longueur des séries de données. Ils précisent aussi qu'il est raisonnable d'admettre qu'il ne faut pas utiliser moins d'une année de données. Cependant, ces conclusions sont remises en question par les études plus récentes.

Dans la prochaine section, un bilan de la littérature concernant la longueur des chroniques sera dressé dans le cas d'utilisation de longues séries de données continues puis dans celui de séries discontinues ou avec peu de mesures. Dans ce dernier cas, des exemples de méthodes mises en place pour tenir compte des quelques mesures disponibles à la station de jaugeage du bassin considéré afin d'améliorer le calage des paramètres seront présentées ainsi que leurs principaux résultats.

***Longues séries de données de débit continues*¹**

Les études suivantes évaluent la sensibilité des performances des modèles hydrologiques à la longueur des enregistrements utilisés pour le calage des paramètres du modèle :

- L'étude de *Brath et al.* [2004] avait pour but de fournir à l'utilisateur des indications pour sélectionner une base de données historiques appropriée pour le calage de modèles pluie-débit dans le cadre d'événements de crue. Les résultats ont révélé que de meilleurs résultats

¹ A noter : cet aspect, qui est déjà abondamment traité dans la littérature hydrologique, n'a pas été repris dans cette thèse.

pour le volume et le pic de crue, étaient obtenus avec le calage effectué sur les plus longues périodes.

- *Anctil et al.* [2004] ont évalué l'impact de la longueur des données sur les performances d'un modèle de réseaux de neurones, et d'un modèle conceptuel pluie-débit. Les deux types de modèle obtiennent les meilleures performances quand 3 et 5 ans de données sont utilisés (meilleures que celles obtenues avec 15 ans de données), bien que le modèle GR4J soit plus performant avec des années plus sèches. Le modèle basé sur les réseaux de neurones fournit de meilleurs résultats quand les données de la période d'apprentissage sont dominées par un temps humide. Toutefois, ce dernier modèle dépend également du nombre de nœuds du réseau de neurones.
- L'étude de *Juston et al.* [2009] visait à déterminer le nombre de données nécessaires pour le calage d'un modèle hydrologique. Leurs résultats indiquent que les calages utilisant une quantité limitée de données pourraient fournir des résultats similaires à ceux obtenus avec des séries complètes, à condition que les données limitées soient sélectionnées selon un raisonnement hydrologique (un événement échantillonné, particulièrement durant un événement dominant peut fournir une information plus valable pour le calage du modèle).
- *Liu et Han* [2010] fournissent une méthode basée sur des indices permettant de sélectionner des données adéquates de calage de modèles pluie-débit. L'étude révèle que l'information contenue dans les données de calage est plus importante que la longueur des données : 6 mois de données peuvent fournir plus d'information utile que des séries de données plus longues. En effet, certaines séries de 6 mois donnent de meilleures performances que d'autres de 12 ou 24 mois. Cependant, les performances des séries de 12 et 24 mois sont généralement supérieures à celles de 6 mois.

Le calage des paramètres des modèles hydrologiques peut également dépendre des caractéristiques hydrologiques des données sélectionnées pour ce calage. Certains auteurs se sont intéressés à la fois à la longueur des données nécessaires pour le calage des paramètres des modèles mais aussi aux caractéristiques hydrologiques des données :

- *Yapo et al.* [1996] ont évalué combien d'années étaient nécessaires pour obtenir un "bon" calage du modèle, tels que les résultats sont relativement insensibles à une période particulière. Ils se sont également intéressés aux caractéristiques que devraient avoir les données sélectionnées afin de maximiser les chances d'obtenir des estimations de paramètres fiables. Leurs résultats ont montré que huit années étaient nécessaires pour que les calages soient relativement indépendants des périodes sélectionnées. Ils précisent également qu'une réduction de l'incertitude des paramètres était obtenue quand des années plutôt humides étaient utilisées.
- *Gan et al.* [1997] se sont aussi intéressés à l'influence des caractéristiques des données. Dans cette étude, il apparaît que les années humides sont préférables pour le calage, car les années sèches ne contiennent pas assez de hauts débits pour activer suffisamment les paramètres responsables de la simulation des hauts débits. Les années humides seraient plus susceptibles de contenir à la fois des hauts et des bas débits et donc de fournir plus d'informations pour le calage des modèles.
- Enfin, *Wagener et al.* [2003] se sont également penchés sur l'identification des parties de l'hydrogramme jugées les plus informatives pour le calage des paramètres. Pour cela, ils ont

mis en place une approche d'analyse d'identification dynamique (DYNIA) des modèles conceptuels pluie-débit. La méthode est basée sur l'utilisation de l'échantillonnage de Monte Carlo de l'espace des paramètres ainsi que sur l'utilisation de la distribution de la fonction objectif pour caractériser la qualité des paramètres. Utilisant une fenêtre glissante pour calculer la fonction objectif le long de l'hydrogramme enregistré, chaque pas de temps peut être alors caractérisé par son utilité dans l'identification des valeurs des paramètres. L'identification temporelle de chaque paramètre sera par conséquent liée à la séquence de débit. Les résultats de cette étude indiquent que certaines parties de l'hydrogramme contiennent bien plus d'informations pour l'identification de certains paramètres.

- Dans un autre registre, *Berthet et al.* [2010] ont quant à eux essayé de déterminer une limite inférieure de la longueur des chroniques de données utilisées, qui assure une valeur de RMSE proche de son espérance statistique. Les auteurs ont montré que la chronique de données devait être plus longue que plusieurs décennies pour s'assurer que les RMSE calculées soient non influencées par le contenu des données.

Le Tableau 2 donne un récapitulatif des conclusions des études citées précédemment concernant le nombre optimal d'années nécessaire pour obtenir un « bon » calage des paramètres de modèle hydrologique ainsi que les caractéristiques hydrologiques adéquates de ces données de calage.

En conclusion, dans le cas de bassins versants jaugés, on ne trouve pas de consensus clair sur la quantité d'années nécessaires au calage d'un modèle. Il semble toutefois que le nombre d'années disponibles pour le calage des modèles ait moins d'impact que l'information hydrologique contenue dans les séries de données. *Beven et Westerberg* [2011] soumettent la possibilité que certaines données puissent être « désinformatives » pour le calage et la validation des modèles hydrologiques et discutent de la façon dont on pourrait identifier ces périodes « désinformatives ».

D'autre part, le choix d'années plutôt humides pour le calage des paramètres semblerait donner de meilleures performances, car ces dernières contiendraient à la fois des hauts et des bas débits, et fourniraient ainsi davantage d'informations pour le calage des modèles. La méthode DYNIA a été proposée afin d'identifier quelles périodes pourraient être avantageuses pour le calage de chaque paramètre du modèle.

Tableau 2. Tableau récapitulatif des conclusions de différentes études sur le nombre optimal d'années nécessaire pour obtenir un « bon » calage des paramètres de modèle hydrologique, et sur les caractéristiques hydrologiques adéquates de ces données de calage.

Auteurs	Bassin d'étude	Variable hydrologique étudiée	Modèle hydrologique	Nombre d'années optimal	Caractéristiques hydrologiques des données de calage
Sorooshian et al. [1983]	Bassin de la rivière Leaf (Etats-Unis)	Modélisation et prévision de débit	Modèle conceptuel pluie-débit de prévision de débit NWSRFS-SMA	> 1 an hydrologique	Les données devraient être choisies de façon à obtenir le plus d'informations précieuses.
Yapo et al. [1996]	Bassin de la rivière Leaf (Etats-Unis)	Modélisation et prévision de débit	Modèle conceptuel pluie-débit de prévision de débit NWSRFS-SMA	8 ans	Années plutôt humides
Gan et al. [1997]	3 bassins secs (2 en Afrique et 1 aux Etats-Unis)	Modélisation de débit journalier	4 ou 5 modèles conceptuels pluie-débit	> 1 an hydrologique	Années humides
Wagner et al. [2003]	Bassin du Lower Meway à Teston (Angleterre)	Modélisation de débit journalier	Modèle conceptuel	-	Certaines parties de l'hydrogramme contiennent plus d'informations pour l'identification de certains paramètres.
Brath et al. [2004]	Le Reno (Italie)	Evènements de crue	Modèle continu distribué spatialement	> 3 mois	-
Anctil et al. [2004]	Bassin du Serein (France)	Modélisation de débit	Modèle conceptuel pluie-débit GR4J et modèle de réseau de neurones (ANN)	3 et 5 ans	Années plus sèches pour GR4J et années humides pour réseau de neurones
Juston et al. [2009]	Bassin suédois	Modélisation de débit	Modèle conceptuel et global HBV, modifié et adapté pour le site d'étude	-	Sélection des données selon un raisonnement hydrologique
Liu et Han [2010]	Bassin du Brue (Royaume-Uni)	Modélisation de débit	Modèle pluie-débit PDM	Entre 12 et 24 mois	Certaines séries de six mois peuvent être meilleures que des plus longues.
Berthet et al. [2010]	178 bassins français	Prévision de débit	Modèle de prévision GRP	Plus longue que plusieurs décennies	-

Séries de données de débit partielles : Cas des bassins peu ou non jaugés

Des conditions climatiques et de débit variées permettraient de mieux représenter la variabilité naturelle du bassin versant. Toutefois, dans certains cas, aucune mesure de débit n'est disponible sur le bassin et dans d'autres, seulement quelques jaugeages ponctuels ont été effectués. Cependant, le risque majeur de prendre en compte ces quelques mesures de débit est de rendre le modèle peu robuste, avec des paramètres ne représentant pas le comportement général du bassin. C'est pourquoi quelques études ont essayé de définir un cadre spécifique pour la prise en compte des données mesurées sur le bassin pour l'évaluation des paramètres dans le cas de bassins versants non jaugés :

- L'étude de *Rojas-Serna et al.* [2006] a eu pour but de déterminer une stratégie pour caler les paramètres d'un bassin versant partiellement non jaugé, tout en tenant compte des quelques années de données disponibles sur ce même bassin. De bonnes estimations de paramètres peuvent être obtenues en combinant l'information de quelques jours d'observations et une estimation régionale des paramètres.
- *Droque et Plasse* [2014] ont également appliqué sur un échantillon de 149 bassins, une méthode issue des travaux de thèse de *Rojas-Serna* [2005] combinant les quelques données disponibles ponctuellement et l'information régionale (cette méthode est présentée plus en détail par la suite dans le chapitre 6 de cette thèse). Les auteurs concluent aussi que l'information hydrométrique ponctuelle observée sur le bassin cible permet une amélioration du processus de régionalisation du débit journalier.
- *Seibert et Beven* [2009] se sont également intéressés à la simulation des débits de bassins non jaugés ayant quelques mesures de débit disponibles. Leurs résultats ont indiqué que peu de mesures pouvaient contenir beaucoup d'informations. Toutefois, les auteurs soulignent que les résultats peuvent différer significativement entre les bassins et qu'ils dépendent des jours choisis pour le choix des mesures.
- L'étude de *Seibert et McDonnell* [2015] a eu pour but d'explorer la valeur d'une campagne de jaugeage simple pour effectuer des prévisions sur des bassins non jaugés. Ils montrent qu'un événement observé ou seulement 10 observations effectuées durant une période de haut débit fournissent presque autant d'informations que trois mois de débit continu utilisé pour caler le modèle hydrologique.
- *Catalogne et al.* [2014] se sont intéressés à l'estimation d'un quantile de débit (QMNA5, débit de référence d'étiage) à partir de quelques données de jaugeage mesurées sur le site cible peu jaugé. La méthode proposée consiste à rechercher le bassin voisin jaugé le plus similaire, en termes de corrélation entre les quelques mesures de débit disponibles sur le site cible et les débits observés à la même date sur le site voisin. L'hypothèse établie est que la relation obtenue précédemment peut être utilisée afin d'avoir une approximation de la relation entre le QMNA5 au site cible et celui au site jaugé connu. L'étude montre un gain de performances quand le nombre de jaugeages augmente jusqu'à 20 puis se stabilise. De plus, pour un nombre de jaugeages identiques, la réalisation de plusieurs jaugeages au cours d'une même saison d'étiage fournit un gain de performance moins important que s'ils étaient répartis sur plusieurs années. Les auteurs préconisent alors un nombre de jaugeage annuel égal à trois.

D'autres auteurs se sont intéressés à la prise en compte de données de débit discontinues pour le calage des paramètres :

- *Perrin et al.* [2007] ont évalué la sensibilité des modèles pluie-débit continus GR4J et TOPMO à la quantité d'information utilisée pour leur calage, en sélectionnant aléatoirement des données ponctuelles dans l'hydrogramme observé (période de calage non continue). En validation, les performances des modèles semblent s'améliorer quand le nombre de données sélectionnées pour le calage augmente. Le modèle identifierait donc mieux le comportement hydrologique du bassin quand le contenu de l'information de la période de calage augmente. Toutefois, à partir de 350 données sélectionnées aléatoirement (non équivalent à une année civile ou hydrologique), les performances atteignent un plateau signifiant que l'apport d'information par des mesures supplémentaires serait redondant.
- *Kim et Kaluarachchi* [2009] ont eux aussi étudié comment les limites de la continuité et de la durée des données impactent le calage des modèles hydrologiques. Leurs résultats révèlent que les données échantillonnées aléatoirement nécessitent une plus courte période de calage que les données continues afin d'obtenir de bonnes performances du modèle. Toutefois, le nombre de mois à haut débit dans les données de calage peut affecter l'efficacité du modèle, suggérant qu'un taux de mois à haut débit de qualité suffisante est nécessaire.
- L'étude de *Tan et al.* [2008] avait pour objectif d'évaluer le calage de modèles pluie-débit sous un nombre limité de données. Les résultats diffèrent selon les objectifs hydrologiques. Dans le cas d'une simulation globale de l'hydrogramme de crue, ainsi que de l'amplitude et du « timing » du pic de crue, le calage basé sur des événements donne généralement de meilleures performances que le calage effectué sur une chronique continue. Alors que pour simuler le volume de débit, le calage continu est nettement supérieur à un calage basé sur des événements isolés.

Ainsi, ces études ont révélé que tenir compte de quelques mesures ponctuelles de débit dans le cadre de bassins peu jaugés permettrait d'améliorer le calage des paramètres des modèles et donc les performances de ces derniers. L'utilisation de données discontinues pour le calage des modèles permettrait aussi, dans certains cas, d'améliorer les performances des prévisions des crues (pour le pic de crue et son « timing »). Toutefois, les résultats dépendent énormément de l'objectif hydrologique fixé et des bassins considérés. *Perrin et al.* [2007] soulignent également que plus le nombre de données discontinues utilisées est élevé, meilleures seront les simulations. Toutefois, le nombre de données utiles peut atteindre un pallier, à partir duquel l'ajout d'informations hydrologiques devient redondant et n'apporte pas d'amélioration.

2.3.2 Longueur des chroniques de débit pour le calcul des débits caractéristiques²

Un autre aspect de l'impact de la densité temporelle sur les calculs hydrologiques semble important d'être abordé. Il s'agit de l'influence du nombre d'observations de débit sur le calcul de certains débits caractéristiques tels que le débit moyen à long terme, appelé aussi module, ainsi que certaines caractéristiques de débits extrêmes.

² Cet aspect étant déjà bien abordé au sein de la thèse de *Catalogne* [2012], il ne sera pas développé dans la suite de la thèse.

On trouve dans la littérature des recommandations concernant la longueur des chroniques de débit utilisée pour une estimation suffisamment fiable de certains débits caractéristiques. Toutefois, il n'existe pas de consensus sur la quantité des observations nécessaires. Par exemple, la réglementation Suisse recommande une période de 10 ans afin de calculer le Q347 (débit dépassé 347 jours par an, soit 95% du temps) [Aschwanden et Kan, 1999]. Tallaksen et van Lanen [2004], quant à eux conseillent de travailler avec des chroniques plus longues, composées d'au minimum 15 à 20 ans d'observation. Enfin, Dyck [1976] considère qu'utiliser des chroniques de plus de 3 ans serait déjà suffisant en contexte climatique humide. Mais cette quantité d'observations devrait dépendre de la variable hydrologique que l'on cherche à estimer. Par exemple, pour estimer les crues de période de retour T de manière fiable, on considère que la chronique doit se composer d'au moins T/3 années [Catalogne et Sauquet, 2011].

En France, Catalogne [2012] a, durant sa thèse, montré sur un échantillon de 274 bassins français, que 18 années d'observation de débit étaient nécessaires pour le calcul du module interannuel, 24 et 26 ans pour les quantiles Q90 et Q95 (débits dépassés dans 90% et 95% des cas, extraits de la courbe des débits classés interannuelle), 15 ans pour l'indice de saisonnalité (date moyenne des jours où le débit est inférieur au Q95) et moins de 5 années pour le « Base Flow Index » (proportion de l'écoulement de base par rapport à l'écoulement total). Toutefois, ces longueurs de chroniques optimales présentent de fortes disparités selon la variable et le bassin versant considérés.

Dans leur étude, Nicolle et al. [2013] montrent que l'erreur d'estimation du QMNA5 sur des sites peu jaugés dépendaient du nombre d'années d'observation de débit (moins il y avait d'observations, plus l'erreur était importante).

2.4. Synthèse

La densité spatiale du réseau hydrométrique semble influencer les performances des modèles hydrologiques pluie-débit et débit-débit. En effet, la baisse de la densité du réseau hydrométrique entrainerait une chute des performances des modèles hydrologiques dans le cas de bassins peu ou non jaugés. Or, les mesures effectuées sur les bassins versants donneurs sont les seules données disponibles pouvant aider à paramétrer les modèles hydrologiques utilisés pour simuler le débit d'un bassin non jaugé. Il semble alors important de pouvoir quantifier au mieux les effets d'une rationalisation des stations hydrométriques dans ce cas.

Une autre question majeure en modélisation hydrologique est la longueur et la continuité des chroniques de débit utilisées pour le calage des paramètres des modèles. La règle générale est que les données de débit devraient regrouper suffisamment de conditions climatiques et hydrologiques diverses pour permettre l'identification des paramètres en fonction des objectifs fixés. Les études que nous avons passées en revue indiquent que la longueur des chroniques de débit est moins importante que l'information hydrologique contenue dans ces chroniques. Le choix d'années plutôt humides pour caler les paramètres, semblerait également fournir de meilleures performances que les années sèches. De plus, utiliser un certain nombre de mesures de débit dans le cadre d'estimation de débit d'un bassin non jaugé en combinaison avec une connaissance préalable des paramètres régionalisés, permettrait d'améliorer les performances des modèles pluie-débit.

Chapitre 3.

Matériel et méthodes

3.1. Introduction

Dans l'équipe hydrologique d'Irstea Antony, il est commun de travailler sur de larges échantillons de bassins versants afin d'assurer la généralité et la robustesse des résultats et de rendre les résultats davantage exploitables en opérationnel grâce à la diversité de l'échantillon. De plus, pour les problématiques de modélisation sur des bassins versants non jaugés et d'évaluation d'impact des densités de réseau de mesure, il semble indispensable de travailler sur de grandes bases de données.

Dans ce chapitre, nous présenterons tout d'abord les données hydro-climatiques et l'échantillon de 609 bassins versants utilisés au cours de la première partie de cette thèse portant sur l'évaluation de l'impact de la densité du réseau hydrométrique sur les calculs hydrologiques.

Ce chapitre vise également à donner des éléments de méthodologie utiles pour la suite des chapitres de cette partie 1. Le modèle hydrologique utilisé au cours de cette thèse est présenté. Il s'agit du modèle pluie-débit journalier GR4J, largement utilisé dans les études menées à Irstea. Par la suite, nous présentons la méthode de régionalisation retenue pour la modélisation de chroniques de débit journalier des bassins versants non jaugés. Améliorer les méthodes de régionalisation n'est pas un objectif de la thèse, c'est pourquoi nous utiliserons une méthode de régionalisation éprouvée basée sur la proximité spatiale.

Afin de pouvoir appliquer cette méthode de régionalisation, nous avons besoin d'estimer une distance entre bassins versants, afin d'identifier quels bassins donneurs sont les plus proches de notre bassin cible non jaugé. Pour cela, différentes distances hydrologiques seront testées et comparées afin de déterminer quelle distance serait la plus appropriée pour notre étude d'évaluation de l'impact de la densité du réseau hydrométrique en hydrologie.

Enfin, l'une des propriétés importantes attendues d'une méthode de régionalisation est sa robustesse vis-à-vis de la disponibilité des données hydrométriques. Nous avons proposé et comparé deux méthodes d'évaluation de la robustesse de la méthode de régionalisation basée sur la proximité spatiale. La méthode retenue sera par la suite utilisée pour le reste des tests de cette première partie.

3.2. Présentation de la base de données de la partie I

3.2.1 Les données hydro-climatiques

Les données hydro-climatiques utilisées dans la première partie sont issues d'une base de données mise en place en 2011 au sein de l'équipe hydrologie d'Irstea Antony. Cette base de données regroupe des données hydro-climatiques au pas de temps journalier sur 3700 bassins versants français.

Les données pluviométriques sont issues de la réanalyse SAFRAN de Météo-France à la résolution 8 km [Quintana-Seguí et al., 2008; Vidal et al., 2010]. Elles sont disponibles au pas de temps journalier sur la période 1958-2010 et sans lacune. On trouve également dans cette base de données des valeurs journalières d'évapotranspiration potentielle calculée à partir de la formule d'Oudin et al. [2005] basée sur la température de l'air et la radiation extraterrestre. Les données météorologiques nécessaires pour le calcul de l'ETP d'Oudin et al. [2005] sont également issues de la réanalyse SAFRAN.

Les données de débit sont des données journalières extraites de la banque HYDRO. La longueur des chroniques de débit est variable selon les bassins versants.

3.2.2 La sélection d'un échantillon de bassins versants

Critères de sélection

A partir de cette base de données, une sélection de bassins versants a été effectuée suivant divers critères.

Tout d'abord, les bassins devaient être non influencés. Ceux-ci ont donc été sélectionnés en prenant en compte une estimation des capacités de stockage à l'amont de la station, fournies dans notre base de données. Si cette capacité de stockage était supérieure à 10% du débit moyen annuel, alors le bassin versant était écarté de notre base de données.

De plus, afin de pouvoir calculer les débits caractéristiques, tels que le module, les débits d'étiage et les débits de crue, de longues chroniques de débits sont nécessaires (cf. partie 2.3.2.). Nous avons donc gardé les bassins sur lesquels au moins 25 années de données de débits étaient disponibles (chaque année devait comporter moins de 10% de lacune).

Enfin, les bassins versants de notre base de données devaient avoir des données avec moins de 10% de lacunes pour la période 1997-2006. Ces années correspondent aux chroniques de pluie ponctuelles disponibles (c'est-à-dire aux données originales des postes pluviométriques), utilisés pour la deuxième partie de cette thèse concernant l'impact de la densité du réseau pluviométrique sur les calculs hydrologiques. Dans un premier temps, afin de faciliter la liaison entre les deux parties ainsi que pour l'interprétation des résultats, nous souhaitons travailler sur les mêmes périodes en gardant pour les deux parties cette période commune pour le calage des paramètres. Nous verrons par la suite que finalement les périodes de calage vont différer d'une partie à l'autre.

Echantillon de bassins sélectionnés

Ainsi, après cette sélection, notre base de données se compose de 609 bassins versants français (Figure 8). On observe sur la Figure 8 que les bassins versants sélectionnés couvrent l'ensemble du territoire français et représentent des conditions climatiques variées (continentales, océaniques, méditerranéennes). Une liste des 609 bassins versants est présentée dans l'annexe A.

La Figure 9 et le Tableau 3 présentent les principales caractéristiques hydro-climatiques et morphologiques de l'échantillon de bassins versants, en termes de précipitation moyenne annuelle, d'évapotranspiration potentielle moyenne annuelle (ETP), de débit moyen annuel, de surface du bassin versant, d'altitude moyenne du bassin, de fraction solide moyenne des précipitations et de longueur de chronique de débit disponible.

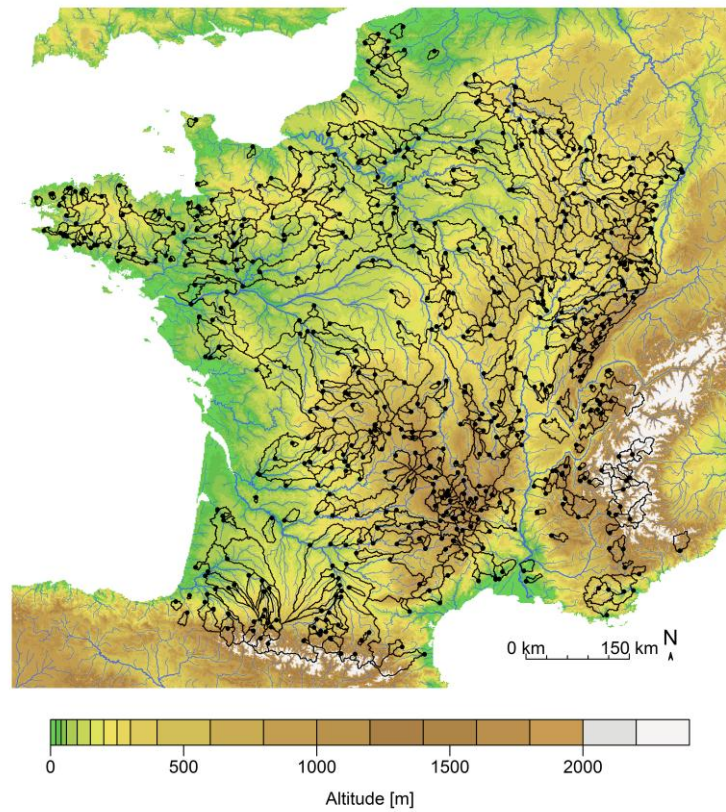


Figure 8. Echantillon des 609 bassins versants sélectionnés.

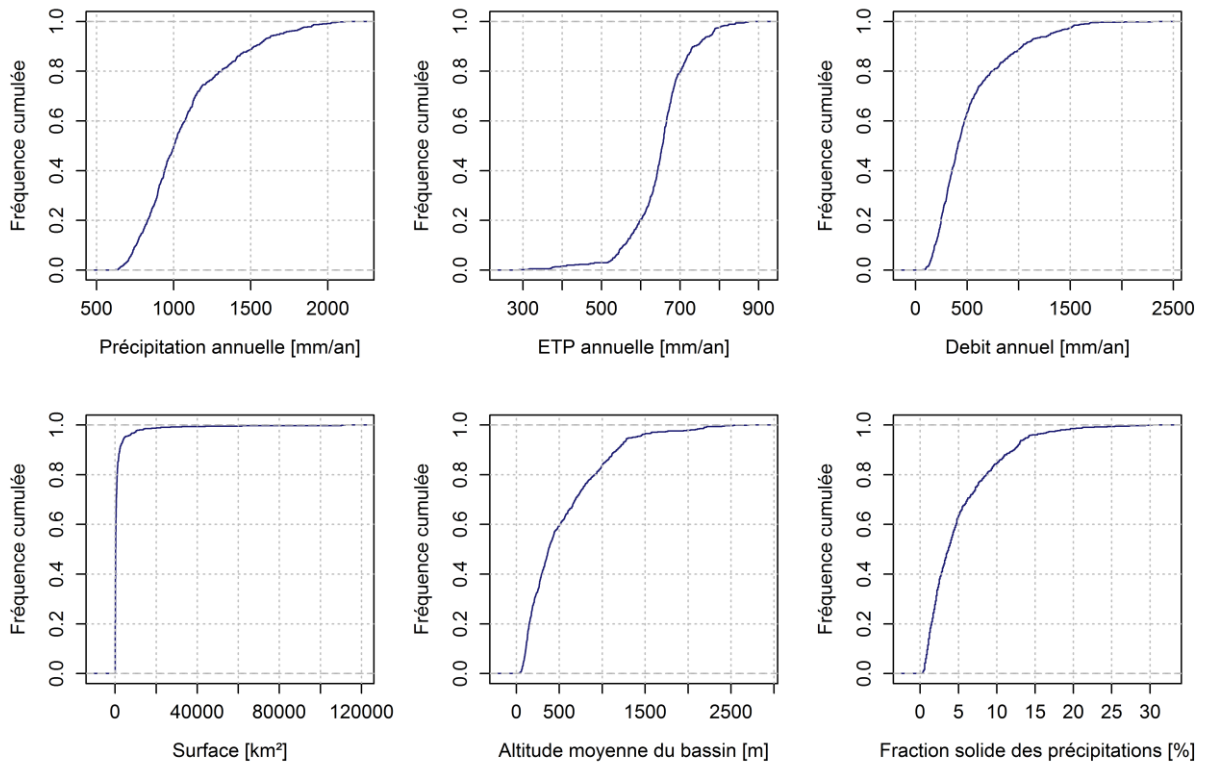


Figure 9. Principales caractéristiques de l'échantillon de bassins versants français sélectionnés dans le cadre de cette thèse.

Tableau 3. Principales caractéristiques de l'échantillon de bassins versants français sélectionnés dans le cadre de cette thèse.

Quantiles	0.05	0.25	0.5	0.75	0.95
Surface (km²)	34	109	270	833	4514
Altitude moyenne du bassin (m)	87	180	375	781	1350
Fraction solide des précipitations (%)	1	2	4	7	14
Précipitation moyenne annuelle (mm/an)	714	863	1003	1209	1688
ETP moyenne annuelle (mm/an)	533	616	655	687	782
Débit moyen annuel (mm/an)	159	272	411	643	1308
Longueur des chroniques de débit disponibles (années)	27	36	39	43	51

Ces différentes caractéristiques hydro-climatiques et morphologiques présentent une grande diversité entre les bassins. En effet, nous retrouvons dans cette base de données des bassins en majorité de basse et moyenne altitudes et d'autres situés en haute montagne (dans les Alpes et les Pyrénées). Pour les bassins de montagne, les précipitations solides auront un rôle très important sur l'écoulement du bassin versant et les températures joueront sur la fonte nivale des neiges (le modèle GR4J sera bien entendu utilisé avec une routine d'accumulation et de fonte de la neige).

La surface des bassins versants est également très diversifiée puisque la surface des plus petits bassins est de l'ordre de quelques dizaines de km² alors que les plus grands ont une surface supérieure à 10 000 km² (avec deux grands bassins dont la surface est supérieure à 100 000km²).

La Figure 10 présente les distributions des débits caractéristiques observés de crue et d'étiage. Les débits de crue décennale et centennale ont été obtenus à l'aide de la loi de Gumbel [Lang et Lavabre, 2007] alors que les débits caractéristiques d'étiage correspondent aux quantiles 0.05 et 0.10 de la distribution des débits journaliers, ainsi qu'au débit minimal mensuel de période de retour 5 ans pour le QMNA5. Nous observons également sur ces distributions, une grande diversité de ces débits caractéristiques au sein de l'échantillon de bassins versants.

Une analyse adimensionnelle ($Q/P = f(P/E)$) de l'ensemble de l'échantillon est également effectuée dans le chapitre 4.

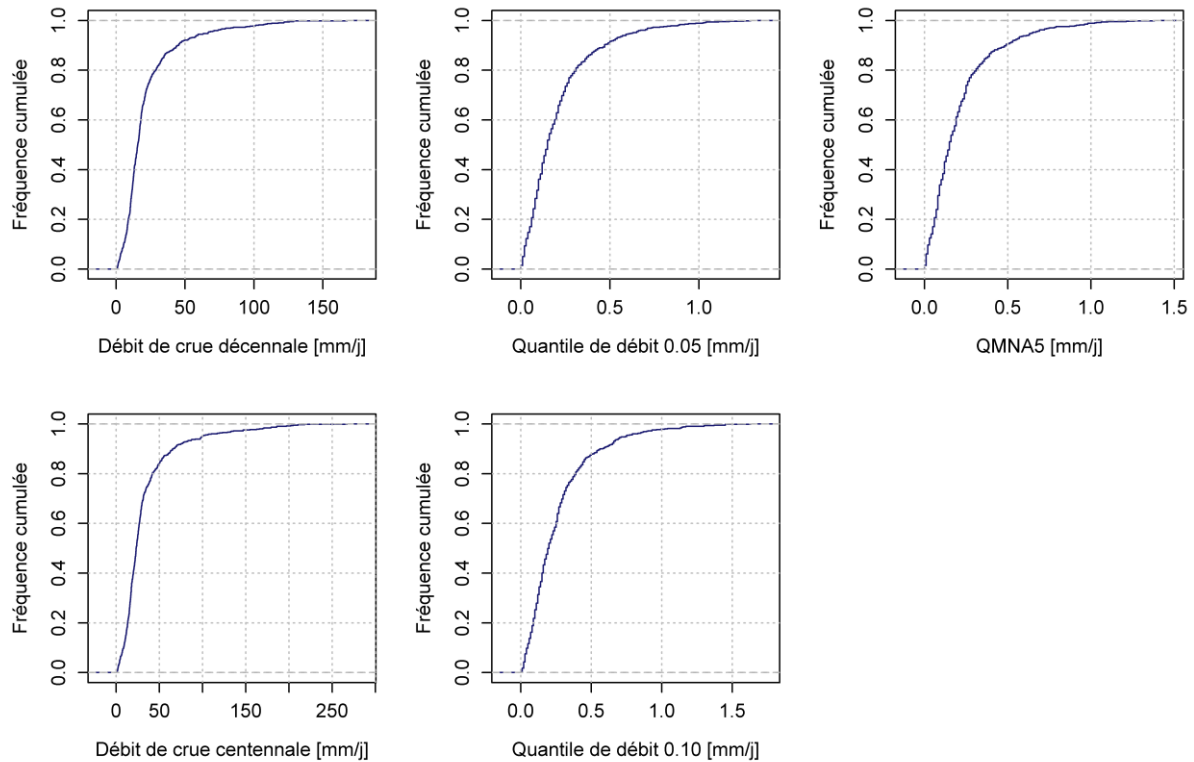


Figure 10. Distributions des débits caractéristiques observés de crues (décennale et centennale) et d'étéage (quantiles 0.05 et 0.10 de la distribution des débits journaliers et QMNA5, débit minimal mensuel de période de retour 5 ans).

3.3. Le modèle hydrologique pluie-débit GR4J

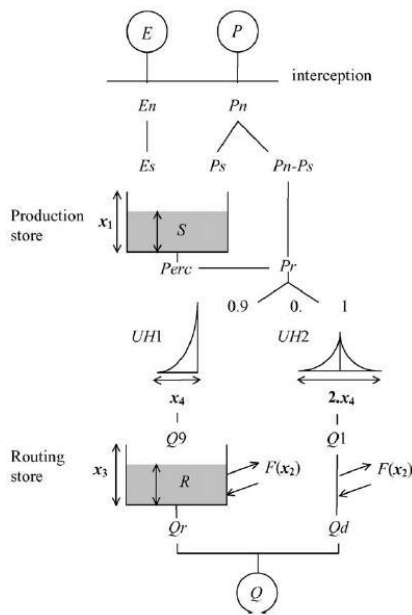
3.3.1 Description du modèle

Le modèle hydrologique GR4J [Perrin *et al.*, 2003], modèle journalier continu avec quatre paramètres libres, a été utilisé au cours de cette thèse, mais l'approche peut être appliquée à n'importe quel autre modèle.

Le modèle GR4J est un modèle à réservoir et global, l'apparentant à des modèles conceptuels malgré son mode de développement empirique. Il possède deux réservoirs : un réservoir de production, qui calcule une pluie efficace, et un réservoir de routage combiné à un hydrogramme unitaire pour le transfert d'eau au sein du bassin. Les entrées du modèle sont des données de pluie et d'évapotranspiration potentielle, mesurées et calculées à l'échelle du bassin versant.

La Figure 11 présente la structure du modèle GR4J ainsi que ses quatre paramètres libres.

Un avantage de ce modèle pour des objectifs de régionalisation réside dans son faible nombre de paramètres libres, limitant les problèmes d'identifiabilité.



Paramètres du modèle :

X1 : capacité maximale du réservoir de production [mm]

X2 : coefficient d'échange souterrain [-]

X3 : capacité à un jour du réservoir de routage [mm]

X4 : temps de base de l'hydrogramme unitaire [j]

Figure 11. Structure et paramètres du modèle pluie-débit GR4J

3.3.2 Utilisation du module neige CemaNeige

Nous avons constaté dans le chapitre précédent que notre base de données était constituée d'une part importante de bassins influencés par la neige. Le modèle GR4J a alors été couplé au module neige CemaNeige afin de tenir compte des précipitations solides dans la modélisation des écoulements sur ces bassins.

Le module CemaNeige a été développé durant la thèse de Valéry [2010]. Ce module a été testé sur un échantillon de 380 bassins versants répartis dans quatre pays (Canada, France, Suède et Suisse). Il comprend deux paramètres libres, le facteur de fonte (K_f) et le coefficient de pondération de l'état thermique du manteau (C_{TG}). Le module CemaNeige considère une distribution spatiale en cinq bandes d'altitude de surfaces égales.

3.3.3 Calage des paramètres du modèle GR4J et du module CemaNeige

Calage des paramètres et évaluation des performances du modèle GR4J

Les paramètres du modèle GR4J ont besoin d'être calés (dans le cas de bassins jaugés) ou transférés à partir d'un calage réalisé sur des bassins voisins jaugés (dans le cas de bassins non jaugés). Le calage des paramètres s'effectue à partir des données de débit observées à l'exutoire du bassin versant. Dans notre étude, le modèle est calé sur chaque bassin versant sur les données disponibles entre 1997 et 2006 (avec dix ans de mise en route pour initialiser les états du modèle) utilisant un algorithme combinant exploration exhaustive de l'espace des paramètres (phase de « préfiltrage ») et recherche locale de l'optimum, particulièrement efficace pour ce modèle [Edijatno et al., 1999; Mathevet, 2005]. Cette méthode de calage permet de trouver les paramètres du modèle qui maximisent une fonction objectif donnée.

La fonction objectif utilisée pour l'optimisation du modèle GR4J est le critère de Nash et Sutcliffe [1970] (NS) calculé sur la racine carrée des débits, comme recommandé par Oudin et al. [2006b] :

$$NS = 1 - \frac{\sum_{i=1}^n \left(\sqrt{Q_{sim}^i} - \sqrt{Q_{obs}^i} \right)^2}{\sum_{i=1}^n \left(\sqrt{Q_{obs}^i} - \sqrt{\overline{Q_{obs}}} \right)^2} \quad \text{Eq. 1}$$

avec i le pas de temps utilisé pour la simulation (ici journalier), Q_{obs}^i le débit observé au pas de temps i , Q_{sim}^i le débit simulé au pas de temps i et $\overline{Q_{obs}}$ le débit observé moyen.

Pour évaluer les performances moyennes sur l'ensemble de l'échantillon de bassins et pour éviter le biais introduit par de fortes valeurs négatives du NS , la version bornée du critère de NS proposé par *Mathevet et al.* [2006] ($C2M$) a été utilisée :

$$C2M = \frac{NS}{2 - NS} \quad \text{Eq. 2}$$

Notons que le critère garde la même référence zéro que le critère de Nash-Sutcliffe ($C2M=0$ quand $NS=0$), possède le même optimum (1 signifie une parfaite simulation pour les deux critères), mais fournit des valeurs positives plus petites que celles du critère NS (par exemple, $C2M=0.67$ quand $NS=0.80$).

Estimation des paramètres du module neige CemaNeige

Pour des bassins jaugés, les deux paramètres de CemaNeige sont en général optimisés. Cependant, *Valéry* [2010] précise dans sa thèse qu'en fixant les paramètres libres à leur valeur médiane, la dégradation des performances est limitée, ce qui représente un atout majeur pour une application sur des bassins non jaugés.

Pour l'estimation des paramètres du module neige pour les bassins non jaugés, nous n'avons ainsi pas utilisé la méthode de régionalisation basée sur la proximité spatiale, comme pour les autres paramètres du modèle (voir ci-dessous). Nous avons utilisé comme paramètres de CemaNeige, ceux résultant d'une autre méthode de régionalisation, précédemment mise en place au sein de l'équipe Hydrologie d'Irstea [*Brigode et al.*, 2014].

Cette étude de *Brigode et al.* [2014] a été menée sur 432 bassins versants influencés par la neige (Figure 12) et a conduit à une régionalisation des paramètres de CemaNeige par massifs montagneux :

- Pour les bassins jugés peu influencés par la neige (dits « de plaine »), les paramètres utilisés sont les paramètres par défaut établis par *Valéry* [2010] (valeurs médianes de la distribution des valeurs calées sur les 380 bassins versants de l'étude ($C_{TG} = 0.25$ [-] et $K_f = 3.74$ mm/j/°C)).
- Pour les bassins jugés influencés par la neige, des valeurs médianes des paramètres calés sur les 432 bassins ont été obtenues par massif montagneux et sont répertoriées dans le Tableau 4.

La Figure 12 indique la localisation des 432 bassins montagneux de l'étude ainsi que la distribution des deux paramètres calés sur ces bassins et classés par massif montagneux.

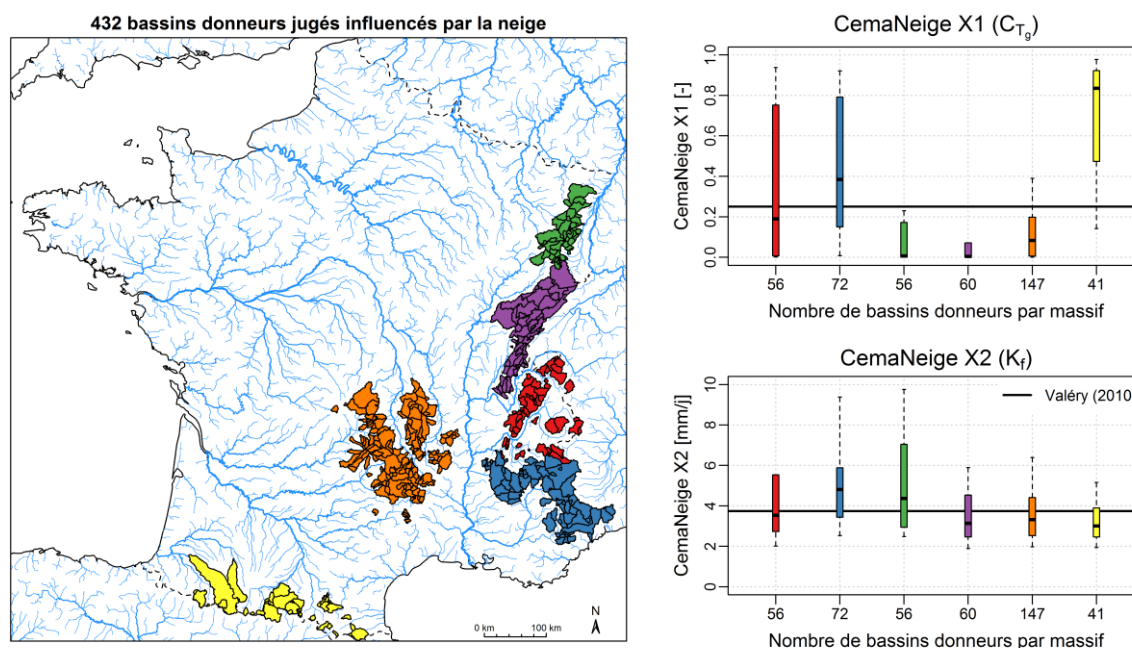


Figure 12. Localisation et massif montagneux (Vosges, Jura, Massif Central, Alpes du Nord, Alpes du Sud ou Pyrénées) des 432 bassins donneurs considérés comme influencés par la neige (à gauche) et distributions des valeurs de paramètres de CemaNeige par massif montagneux (à droite). Les boîtes à moustaches indiquent les quantiles 0.10, 0.25, 0.50, 0.75 et 0.90 des valeurs de paramètres. Figure extraite de *Brigode et al.* [2014].

Tableau 4. Valeurs médianes des paramètres de CemaNeige de référence [Valéry, 2010] et établies pour chaque massif montagneux français sur une base de 432 bassins versants jaugés. Tableau extrait de *Brigode et al.* [2014].

	C_{TG} [-]	K_f [mm/j/°C]
Par défaut [Valéry, 2010]	0.250	3.74
Vosges	0.006	4.37
Jura	0.003	3.14
Massif Central	0.830	3.32
Alpes du Nord	0.189	3.53
Alpes du Sud	0.384	4.81
Pyrénées	0.835	3.00

3.4. La méthode de régionalisation basée sur la proximité spatiale

Les bassins versants non jaugés ne disposent pas de mesures de débit à leur exutoire pour permettre le calage des paramètres du modèle hydrologique. Ils nécessitent alors un traitement spécifique pour leur modélisation de chroniques de débit. Dans le chapitre 2, nous avons énuméré et comparé plusieurs méthodes de régionalisation. Puisque notre but ici n'est pas de développer une nouvelle méthode de régionalisation mais seulement de les évaluer, nous utilisons une méthode existante jugée performante par *Oudin et al.* [2008b] : l'approche basée sur la proximité spatiale. Pour rappel, *Oudin et al.* [2008b] ont montré qu'en France cette méthode, utilisée avec des modèles pluie-débit,

fournissait la meilleure solution de régionalisation (en comparaison avec la régionalisation utilisant la régression et celle basée sur des similarités physiques). Cette conclusion est à relier avec la densité importante de stations disponible en métropole.

La régionalisation basée sur la proximité spatiale consiste à transférer les paramètres calés sur les bassins voisins pour calculer le débit sur le bassin non jaugé. Pour la suite de la thèse, nous qualifierons de «donneurs» les bassins jaugés voisins du bassin non jaugé (qui devient donc «receveur») et qui lui transmettent leurs paramètres.

Ainsi, pour chaque bassin considéré comme non jaugé, 608 jeux de paramètres sont disponibles comme donneurs potentiels. Les performances du modèle en calage seront utilisées plus tard comme une référence pour évaluer la perte d'efficacité due à la régionalisation.

Pour un bassin versant non jaugé, l'utilisation du modèle GR4J a consisté à caler les paramètres du modèle pour les n (ici $n = 10$)³ bassins versants voisins sélectionnés sur une période de 10 ans (1997-2006). Puis, n simulations du débit journalier du bassin non jaugé ont été effectuées (avec chacun des n jeux de paramètres transférés depuis les voisins calés), en utilisant les données de pluie et d'ETP du bassin non jaugé en entrée du modèle. On a ainsi obtenu n chroniques de débit journalier pour le bassin cible. La chronique finale retenue résulte d'une pondération par l'inverse de la distance au carré entre bassins voisins et bassin non jaugé, comme présenté dans l'équation 3. *Oudin et al.* [2008b] et *McIntyre et al.* [2005] ont montré que cette méthode était plus cohérente que celle consistant à moyenniser les paramètres des bassins voisins, puisqu'elle utilisait des jeux de paramètres non modifiés. Ceci permettait ainsi d'utiliser une information totale contenue dans les paramètres des voisins calés localement.

$$Q_{\text{bassin non jaugé}}(j) = \frac{\sum_{i=1}^n \left(Q_{\theta i}(j) \times \frac{1}{d_i^2} \right)}{\sum_{i=1}^n \frac{1}{d_i^2}} \quad \text{Eq. 3}$$

avec $Q_{\theta i}$ débit du bassin non jaugé obtenu avec le jeu de paramètre ϑ du bassin voisin i , d_i : distance hydrologique (la même distance que celle utilisée pour sélectionner les bassins les plus proches (voir la partie 3.5.)) entre le bassin non jaugé et le bassin voisin i , et j : jour de la simulation.

Notons que l'utilisation de débits spécifiques pour le calage des paramètres des bassins donneurs, nous permet de nous affranchir de la question concernant la différence de taille entre bassins donneurs et bassins receveurs. D'autre part, une relation empirique entre le paramètre X4 du modèle GR4J, représentant le temps de base de l'hydrogramme unitaire, et la surface (S) des bassins versants donneur et receveur, a tout de même été établie⁴ et appliquée lors du transfert de paramètres entre bassin voisin et bassin cible non jaugé :

$$X4_{\text{non jaugé}} = X4_{\text{voisin}} \times \left(\frac{S_{\text{non jaugé}}}{S_{\text{voisin}}} \right)^{0,1} \quad \text{Eq. 4}$$

³ Des tests préliminaires (non présentés ici) ont montré que considérer les 10 premiers voisins était suffisant pour obtenir de bonnes performances. Par la suite, nous aurons donc $n = 10$.

⁴ Relation établie d'après des tests préliminaires (non présentés ici).

3.5. Etude de l'impact de la distance hydrologique utilisée pour la modélisation de bassins versants non jaugés

3.5.1 Pourquoi est-il important d'évaluer la sensibilité des simulations régionalisées au choix de la distance hydrologique ?

Calculer la distance entre deux points est facile (du moins dans l'espace euclidien). Mais comment peut-on calculer une distance entre deux ensembles de points, ou plus précisément entre deux bassins versants ? Comment peut-on dire lequel est le plus proche et lequel est le plus éloigné d'un autre ? Différentes méthodes peuvent être suggérées : la méthode la plus simple serait de calculer la distance entre deux points de référence, par exemple les exutoires de deux bassins (d_{exutoire}) ou leurs centroïdes ($d_{\text{centroïde}}$). Naturellement, ces solutions simples ont des limites, comme identifier comme infiniment proches deux bassins avec le même exutoire dans le premier cas (confluence), ou avec le même centroïde dans le second cas (Figure 13). D'autres solutions pourraient être proposées, et nous les discuterons plus tard dans cette partie.

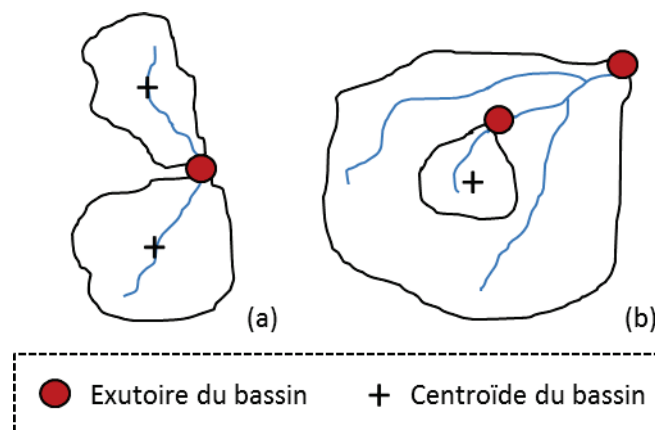


Figure 13. Schémas de deux cas où de simples distances basées sur des points de références peuvent donner des résultats incohérents : (a) deux bassins différents avec des exutoires extrêmement proches situés à leur confluence, (b) deux bassins emboîtés avec des centroïdes extrêmement proches.

Dans la littérature hydrologique, la question des distances des bassins versants a surtout été traitée par des auteurs intéressés par la covariance du débit (voir en particulier *Gottschalk* [1993a, 1993b], *Sauquet et al.* [2000], *Skoien et al.* [2006] et *Laaha et al.* [2014]). *Skoien et al.* [2006] mentionnent que « les distances euclidiennes entre bassins ne sont pas la façon naturelle de mesurer la distance spatiale des bassins ». Une récente étude de *Gottschalk et al.* [2011] a comparé la corrélation spatiale des données de débit de bassin obtenues avec deux distances bien connues : la distance entre les centroïdes de deux bassins et la distance de Ghosh [*Ghosh*, 1951], qui consiste à combiner toutes les distances possibles entre deux bassins, précisément la méthode utilisée dans le récent travail de « top-kriging » [*Laaha et al.*, 2014; *Skoien et al.*, 2006]. Par la suite, nous suivrons *Gottschalk et al.* [2011] pour nommer cette distance la *distance de Ghosh*.

Cette partie a pour but d'évaluer l'impact du choix d'une mesure de distance entre deux bassins sur l'efficacité d'une méthode de régionalisation, basée ici sur la proximité spatiale. Nous présenterons

les différents paramètres évalués pour calculer la distance entre les bassins versants. Enfin, une comparaison en termes d'efficacité de régionalisation est réalisée et discutée.

3.5.2 Les différentes mesures de distance étudiées

Les différentes mesures de distance testées ici sont les suivantes :

- Parce que la méthode la plus simple consiste à calculer la distance entre deux points de référence, nous testons la distance entre deux exutoires de bassins (d_{exutoire}) et la distance entre leurs centroïdes ($d_{\text{centroïde}}$). La distance combinant ces deux mesures de référence est également testée :

$$d(\alpha) = \alpha \times d_{\text{exutoire}} + (1 - \alpha) \times d_{\text{centroïde}} \quad \text{Eq. 5}$$

avec α variant entre 0 et 1 avec un incrément de 0.1.

Notons que quand α est égal à 0, la distance combinée est égale à la distance séparant les centroïdes ($d(0)=d_{\text{centroïde}}$), alors que quand α est égal à 1, la distance combinée est égale à la distance séparant les exutoires ($d(1)=d_{\text{exutoire}}$).

- La distance de Ghosh [Ghosh, 1951; Gottschalk et al., 2011] qui consiste à moyenner toutes les distances possibles entre deux jeux de points distribués au sein de deux bassins versants.
- Deux distances proposées par Oudin (communication personnelle), où les bassins sont identifiés comme des rectangles d'emprise minimum. Pour le premier (d.Oudin1), la distance minimale entre les sommets des rectangles est mesurée. Pour le second (d.Oudin2), cette même distance est calculée mais en tenant compte de l'orientation du bassin, définie par l'angle de la droite centroïde-exutoire.

Le Tableau 5 résume les différentes distances testées ici.

Tableau 5. Liste des distances mesurées entre les bassins versants testées ici.

Distance	Nom
d.exutoire	Distance entre les exutoires des bassins
d.centroïde	Distance entre les centroïdes des bassins
$\alpha \times d_{\text{exu}} + (1-\alpha) \times d_{\text{cent}}$	Distances combinées (exutoires + centroïdes) avec des valeurs de α variant de 0.1 à 0.9 par pas de 0.1
d.Ghosh	Distance de Ghosh
d.Oudin1 et d.Oudin2	Distances des rectangles d'emprise minimum proposées par Oudin

La Figure 14 illustre l'impact de la distance sélectionnée pour calculer la distance entre un bassin non jaugé et ses voisins. Elle montre comment les dix plus proches bassins diffèrent pour chaque distance.

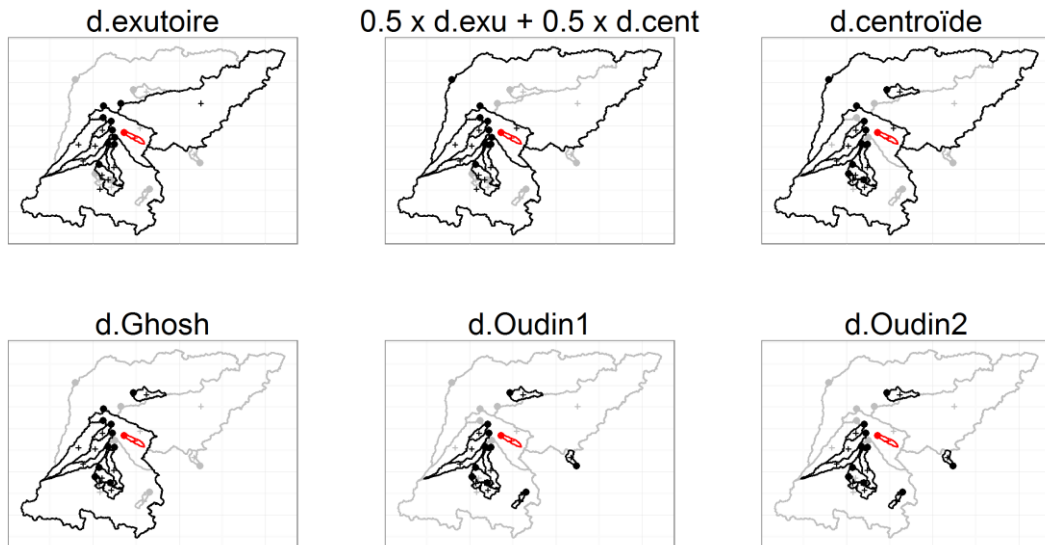


Figure 14. Illustration d'un exemple de sélection de bassins versants pour différentes distances considérées pour mesurer la distance entre un bassin non jaugé et ses 10 plus proches voisins. En rouge, le bassin non jaugé (receveur), en noir les bassins voisins (donneurs) et en gris le jeu complet de bassins voisins.

La Figure 15 présente la distribution du nombre de voisins parmi les dix sélectionnés, qui diffèrent selon que l'on choisit la distance séparant les exutoires ou celle séparant les centroïdes, comme distance pour mesurer l'éloignement entre un bassin non jaugé et ses dix plus proches voisins. Nous pouvons observer que pour 50% de l'échantillon des 609 bassins versants, au moins deux voisins sélectionnés diffèrent entre les deux options. Seuls 10% des bassins n'ont aucun voisin qui diffère selon le choix entre l'une et l'autre de ces deux distances.

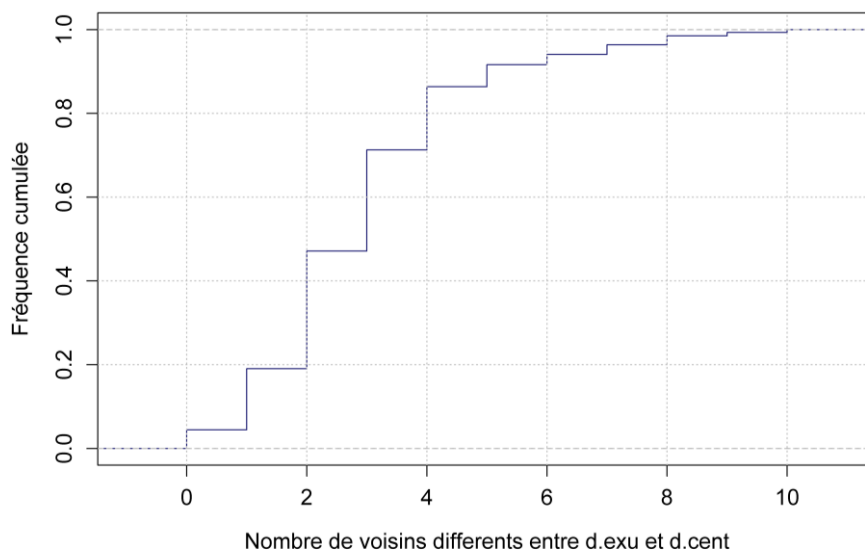


Figure 15. Distribution sur l'ensemble de l'échantillon de bassins versants du nombre de voisins sélectionnés qui diffèrent sur dix lorsque l'on choisit la distance séparant les exutoires ou la distance séparant les centroïdes, comme distance pour mesurer la distance entre un bassin non jaugé et ses dix plus proches voisins.

3.5.3 Comparaison des différentes distances hydrologiques dans une perspective d'une application sur un bassin versant non jaugé

Nous avons testé les différentes distances sur l'ensemble de l'échantillon de 609 bassins versants. Chaque bassin a été successivement considéré comme non jaugé. L'approche de régionalisation a été appliquée pour chaque cas, et les performances du modèle ont été évaluées en utilisant le critère C2M.

La Figure 16 présente la distribution des résultats. Comme attendu, les performances du modèle régionalisé, même avec l'échantillon complet de donneurs, sont plus basses que le modèle calé (voir *Oudin et al.* [2008b]). On observe également que l'efficacité de la méthode de régionalisation varie peu en fonction de la distance utilisée pour sélectionner les bassins donneurs. La distance optimale est obtenue avec la combinaison des distances d_{exutoire} et $d_{\text{centroïde}}$, et pour des petites valeurs de α (0.1 ou 0.2). La valeur 0.2 semble toutefois un peu meilleure, avec un quantile de 0.1 plus élevé. La combinaison de ces deux distances est bénéfique, à condition que le poids de la distance séparant les exutoires reste faible. Mais dans l'ensemble, de faibles différences en termes d'efficacité entre les différentes distances sont observées. Ces résultats peuvent être expliqués par le fait que pour un grand nombre de bassins traités comme non jaugés, les donneurs sélectionnés ne varient pas significativement entre les différents choix de distance.

Les deux dernières distances (d.Oudin1 et d.Oudin2) fournissent des résultats presque similaires et sont proches des résultats de la distance optimale, avec toutefois des résultats pour le quantile 0.1 qui semblent plus faibles.

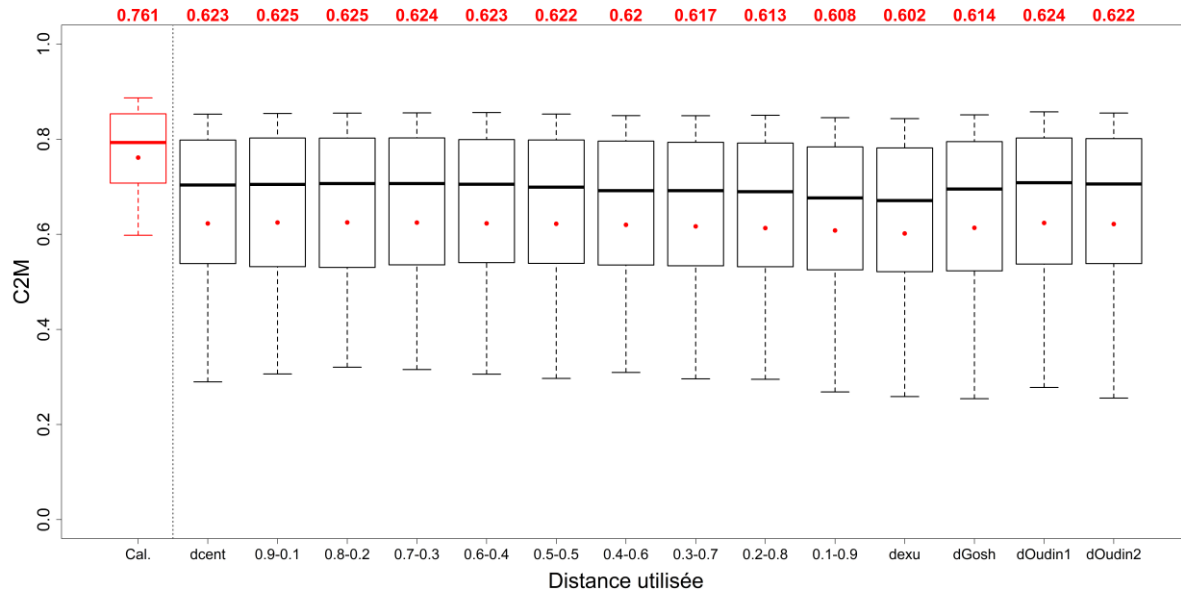


Figure 16. Distributions des performances du modèle GR4J calé et régionalisé sur les 609 bassins, montrant l'impact du choix de la distance en régionalisation. Les résultats de calage (à gauche) sont utilisés comme référence. Les boîtes à moustaches montrent les quantiles 10, 25, 50, 75 et 90 de la distribution, du bas vers le haut. Les performances moyennes sont indiquées par un point et la valeur correspondante est montrée en haut du graphique.

Comme les différences d'efficacité restent très limitées lorsque l'on regarde l'ensemble de l'échantillon de bassins versants, nous nous intéressons maintenant aux bassins qui ont au moins six donneurs qui diffèrent entre le choix de la distance séparant les exutoires et le choix de la distance séparant les centroïdes (Figure 17). Nous pouvons remarquer que les différences de performances sont plus marquées pour ce sous-échantillon et nous observons une différence de 7 points en moyenne de C2M entre la distance optimale (d.Oudin2) et la moins bonne (d.Ghosh). Pour la distance combinant les distances d_{exutoire} et $d_{\text{centroïde}}$, la valeur α optimale est toujours la valeur 0.2 en médiane et en quantiles 0.75 et 0.9. Toutefois, ces résultats ne représentent qu'un faible nombre de bassins versants de notre étude (seulement 51 sur 609) et nous garderons comme valeur optimale celle correspondant au mieux à l'ensemble des 609 bassins versants.

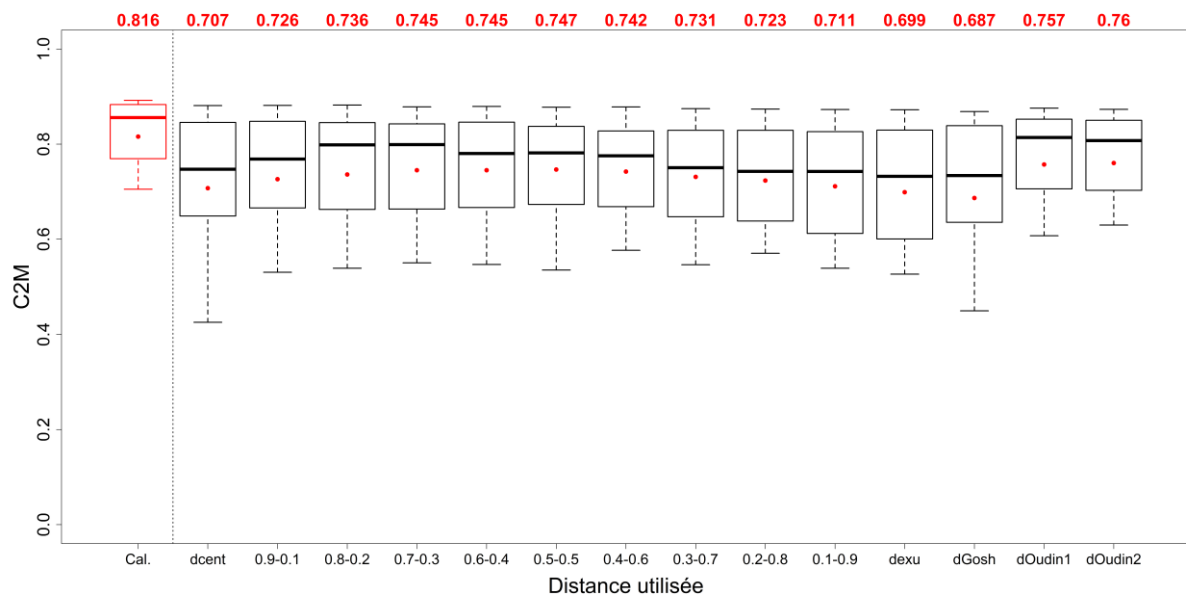


Figure 17. Distributions des performances du modèle GR4J calé et régionalisé, montrant l'impact du choix de la distance en régionalisation, pour les 51 bassins qui ont au moins 6 donneurs différents entre l'utilisation de la distance séparant les exutoires et celle séparant les centroïdes. Les résultats de calage (à gauche) sont utilisés comme référence. Les boîtes à moustaches montrent les quantiles 10, 25, 50, 75 et 90 de la distribution, du bas vers le haut. Les performances moyennes sont indiquées par un point et la valeur correspondante est montrée en haut du graphique.

3.5.4 Conclusion

Bien que les performances du modèle régionalisé semblent rester assez peu sensibles dans notre cas au choix de la distance utilisée, nous devons tout de même décider d'une distance pour effectuer nos tests de régionalisation. Les résultats ont montré que les performances étaient optimales lors de l'utilisation d'une distance combinant la distance séparant les exutoires de deux bassins et celle séparant leurs centroïdes, mais avec des faibles valeurs de α , soit pour un poids limité pour la distance séparant les exutoires de bassins versants. Pour la suite, nous garderons donc un α égal à 0.2 et nous obtenons alors la distance hydrologique suivante :

$$d = 0.2 \times d_{\text{exutoire}} + 0.8 \times d_{\text{centroïde}} \quad \text{Eq. 6}$$

Pour illustrer l'intérêt de ce choix, les Figure 18 et Figure 19 présentent la localisation de deux bassins versants pour lesquels la distance hydrologique retenue (Eq. 6) présente un avantage par rapport à l'utilisation de la distance séparant simplement les centroïdes, ainsi que la configuration de leurs réseaux de bassins donneurs. On distingue le cas où la distance hydrologique retenue (Eq. 6) est utilisée, du cas où il s'agit de celle séparant les centroïdes.

Nous constatons que considérer la distance séparant les exutoires des bassins dans la distance hydrologique présente l'avantage pour ces deux cas, d'écarter des bassins donneurs très grands par rapport au bassin cible (Figure 18) ou au contraire de sélectionner des bassins donneurs de tailles comparables au bassin cible et dont les exutoires sont très proches (Figure 19).

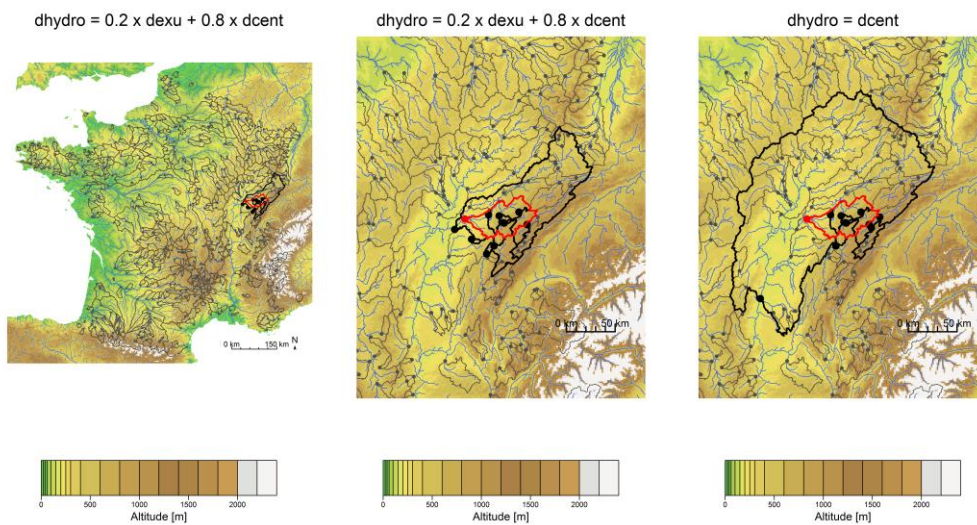


Figure 18. Localisation du bassin versant de la Loue à Parcey (en rouge) ainsi que la configuration de ses réseaux de bassins donneurs (en noir) lorsque la distance hydrologique retenue (figures de gauche et du milieu) et la distance séparant les centroïdes (figure de droite) sont utilisées.

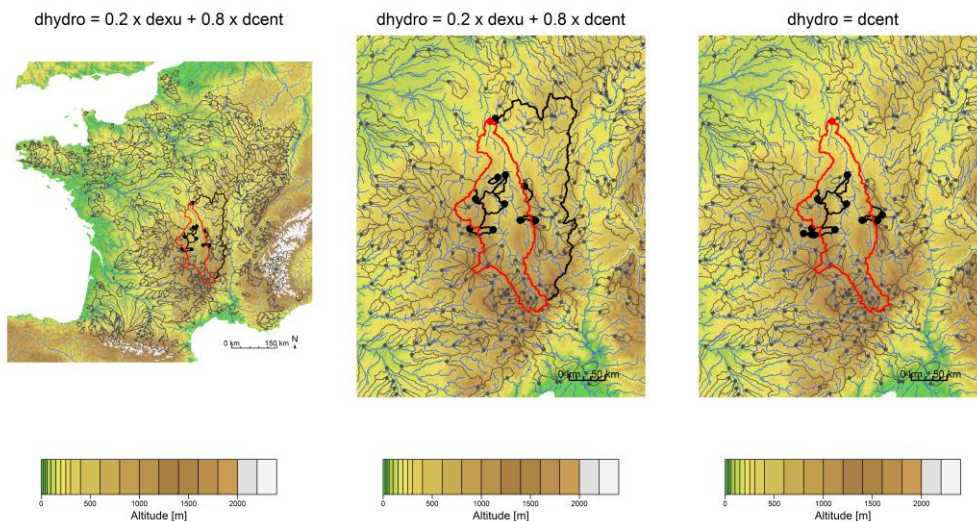


Figure 19. Localisation du bassin versant de l'Allier à Cuffy (en rouge) ainsi que la configuration de ses réseaux de bassins donneurs (en noir) lorsque la distance hydrologique retenue (figures de gauche et du milieu) et la distance séparant les centroïdes (figure de droite) sont utilisées.

3.6. Comparaison de deux méthodes d'évaluation de la robustesse de méthodes de régionalisation

3.6.1 Pourquoi est-il important d'évaluer la sensibilité d'une méthode de régionalisation à la densité du réseau hydrométrique ?

L'une des propriétés importantes attendues d'une méthode de régionalisation est sa robustesse. Deux méthodes de régionalisation peuvent obtenir des performances très similaires dans un environnement riche en données mais ces dernières peuvent différer sous des conditions de disponibilité de données plus limitées : évaluer la sensibilité d'une méthode de régionalisation à la disponibilité d'information (typiquement la densité du réseau avoisinant dans le cas d'approches de proximité spatiale) est un bon moyen d'éviter les déceptions quand deux méthodes académiques sont comparées à un jeu de données réel [Andréassian et al., 2010b]. En effet, les réseaux opérationnels ne sont jamais aussi denses que les hydrologues le souhaiteraient.

Etonnamment, notre recherche bibliographique indique que l'impact de la densité des données hydrométriques sur l'efficacité de la régionalisation n'est pas un sujet de préoccupation dans la littérature. Pour notre part, nous tenons à défendre le point de vue que ce test de sensibilité n'est pas une simple question de « confort hydrologique », mais plutôt une nécessité scientifique, une sorte de « crash test » essentiel pour garantir la crédibilité avant l'utilisation pratique [Andréassian et al., 2009].

Cette partie compare deux méthodes pour évaluer l'impact de la densité du réseau hydrométrique sur les performances de régionalisation. Nous commencerons par présenter les deux méthodes alternatives : la méthode de réduction aléatoire du réseau et la méthode du désert hydrométrique, puis nous détaillerons les résultats de leur comparaison.

3.6.2 Deux méthodes alternatives pour l'analyse de sensibilité d'une méthode de régionalisation basée sur la proximité spatiale

Analyse de sensibilité basée sur une réduction aléatoire du réseau hydrométrique (HALéa)

La méthode de réduction aléatoire du réseau hydrométrique (HALéa) consiste à conserver aléatoirement 90%, 80%, 70%,..., 10% du réseau hydrologique disponible. Le nombre de bassins donneurs reste le même (10), mais ils sont situés en moyenne de plus en plus loin du bassin receveur (même si l'échantillonnage peut permettre de garder des bassins proches).

La procédure de réduction procède de la manière suivante : pour chaque bassin receveur, les bassins voisins sont rangés du plus proche au plus éloigné. Pour chaque bassin donneur, un nombre x entre 0 et 1 est tiré aléatoirement, et un taux d'acceptation x_{seuil} est considéré : si x est plus grand que le seuil, le bassin voisin n'est pas utilisé pour la régionalisation. L'analyse de sensibilité commence avec le réseau complet ($x_{seuil} = 1$) et évolue progressivement vers un réseau réduit correspondant à 90% ($x_{seuil} = 0.9$), 80%,..., et 10% ($x_{seuil} = 0.1$) du réseau initial. La sélection donne toujours les 10 bassins versants les plus proches parmi les autres donneurs, mais pour le réseau réduit, ces donneurs sont situés en moyenne plus loin que pour le réseau complet. La sélection aléatoire est réalisée une seule fois par bassin. Puisque le nombre de bassins est grand, nous pensons que cela ne nous empêche pas d'obtenir des résultats robustes.

Cette méthode a pour but de réduire aléatoirement le réseau de bassins donateurs. La Figure 20 illustre un exemple de l'impact de la réduction du réseau sur la sélection des bassins voisins. Cette figure montre que même si les bassins versants voisins sont situés en moyenne plus loin à mesure que la réduction devient plus forte, des voisins proches peuvent toujours rester dans la sélection, même à des taux élevés de réduction.

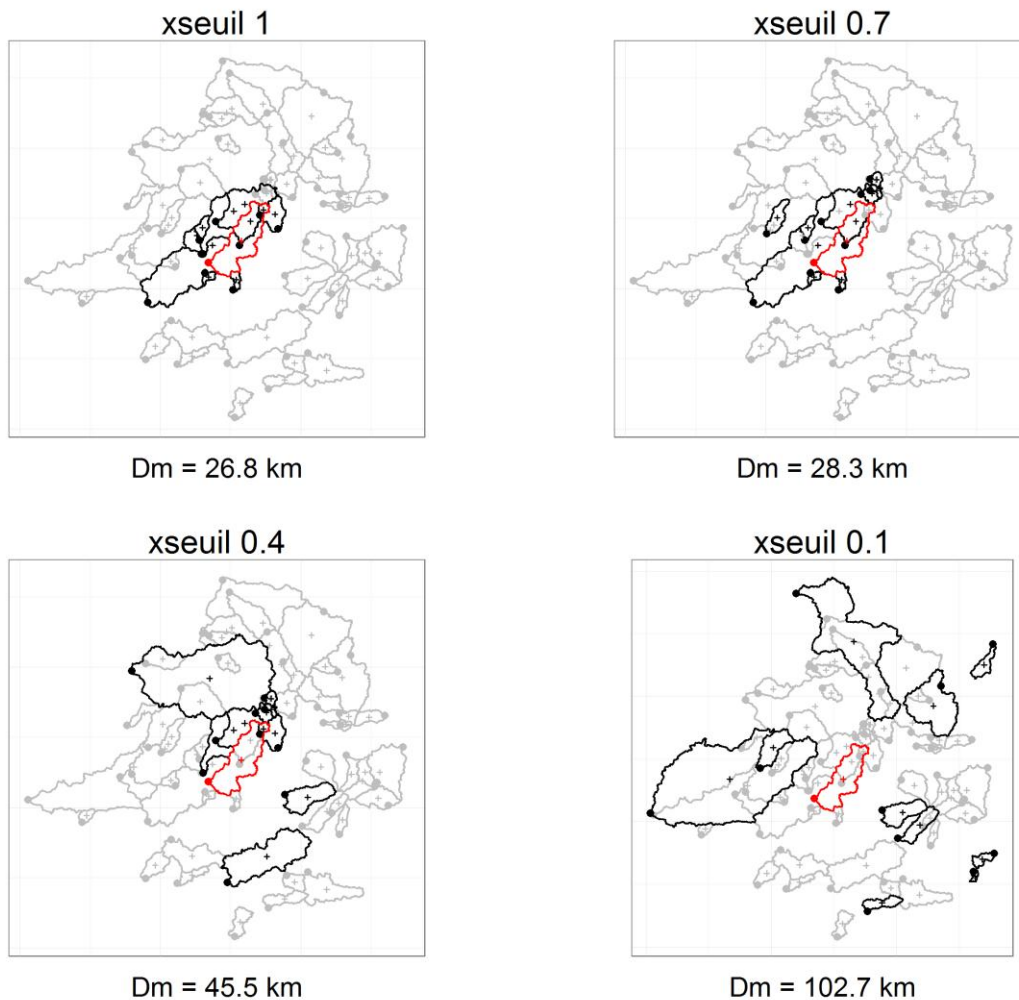


Figure 20. Illustration d'un exemple de sélection de bassins voisins avec le réseau complet (xseuil = 1) et lorsque la procédure de sélection aléatoire du réseau hydrométrique est appliquée (HALéa) pour différents niveaux de réduction (xseuil = 0.7, 0.4 et 0.1). En rouge, le bassin non jaugé (receveur), en noir les bassins voisins (donneurs) et en gris le jeu complet de bassins voisins. Dm est la distance moyenne séparant les 10 bassins donateurs du bassin non jaugé.

Analyse de sensibilité basée sur la méthode du désert hydrométrique (HDes)

La méthode du désert hydrométrique développée par *Boldetti* [2012] consiste à exclure progressivement les bassins donateurs les plus proches : les paramètres sont transférés de voisins qui sont de plus en plus éloignés du bassin versant cible non-jaugé, en fixant une limite inférieure en dessous de laquelle les voisins sont écartés (Figure 21). Ici, nous proposons un test permettant

d'étudier la sensibilité de la méthode de régionalisation quand l'information doit être transférée sur des distances de plus en plus éloignées.

Nous avons donc testé les seuils de distance suivants pour la méthode HDes : 0 (sans limite de distance), 10, 20, 30, ..., 100, 150 et 200 km. Pour mesurer l'impact potentiel de ces seuils d'éloignement, on peut les comparer à la distribution des distances des dix bassins versants les plus proches pour les 609 bassins versants de l'échantillon. La Figure 22 montre que les quantiles 10% et 90% de cette distribution sont de l'ordre de 20 et 50 km, et la valeur minimum de l'ordre de 15 km. La gamme d'éloignement choisie pour le plus proche voisin viendra donc clairement impacter le choix des bassins donneurs.

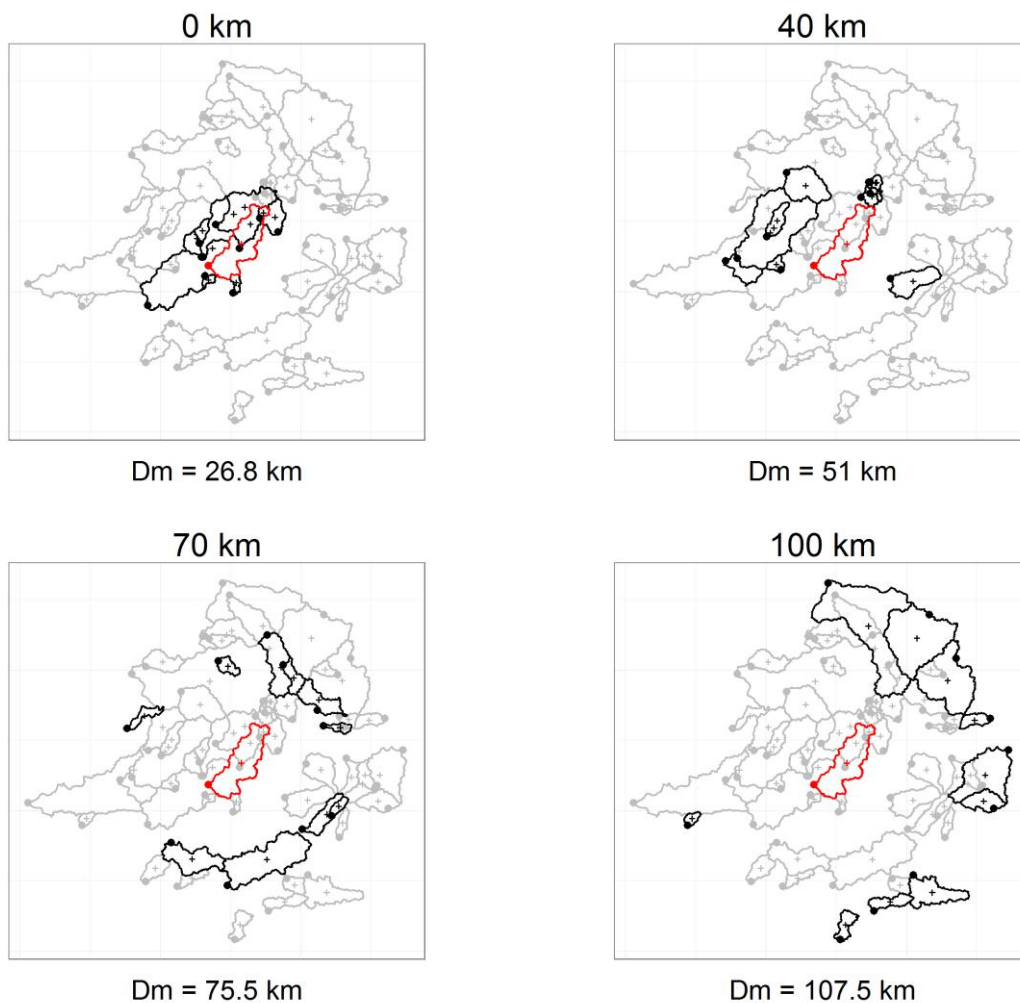


Figure 21. Illustration d'un exemple de sélection de bassins voisins avec le réseau complet (distance = 0 km) et lorsque la méthode du désert hydrométrique est appliquée (HDes) pour différentes limites de distance (40, 70 et 100 km). En rouge, le bassin non jaugé (receveur), en noir les bassins voisins (donneurs) et en gris le jeu complet de bassins voisins. Dm est la distance moyenne séparant les 10 bassins donneurs du bassin non jaugé.

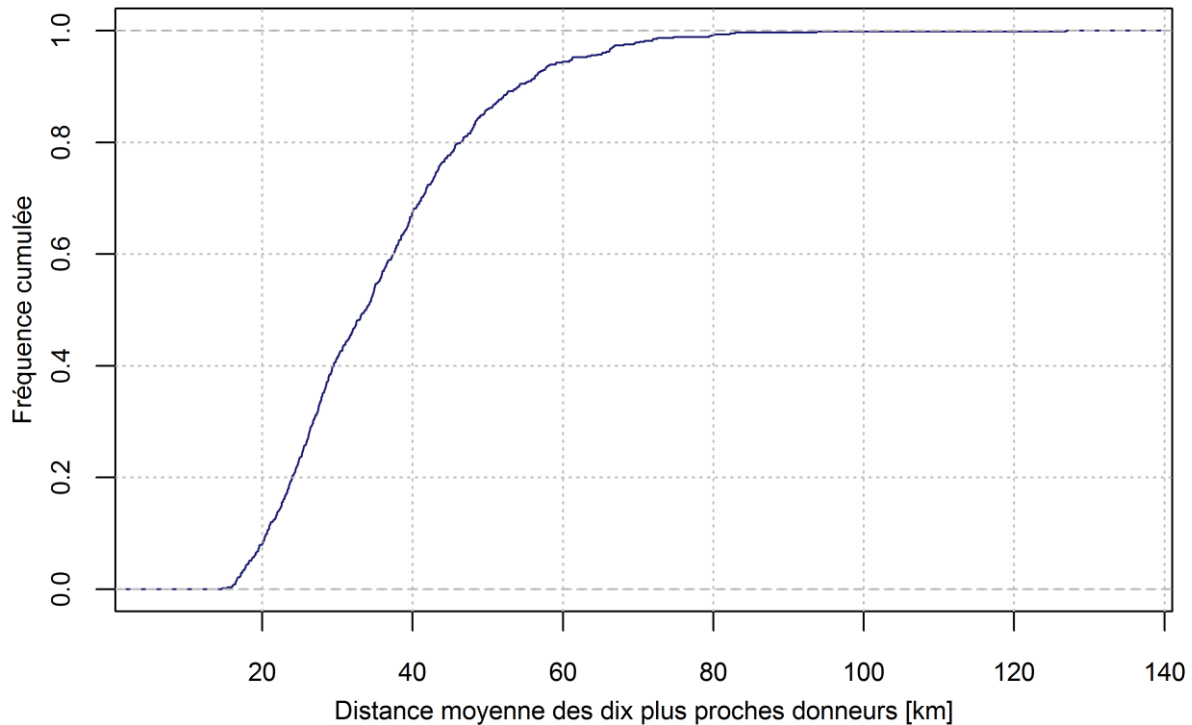


Figure 22. Distribution sur l'ensemble de l'échantillon des 609 bassins de la distance moyenne des dix plus proches donneurs [km].

3.6.3 Comparaison des deux méthodes d'analyse de sensibilité

Les deux approches de sensibilité HALéa et HDes ont été appliquées à la régionalisation du modèle GR4J sur notre base de données de 609 bassins. Chaque bassin a été successivement considéré comme non jaugé. L'approche de régionalisation a été appliquée pour chaque cas, et les performances du modèle ont été évaluées en utilisant le critère C2M. La distance entre bassins est celle retenue dans la partie 3.5.4.

La Figure 23 présente les distributions des performances sur l'ensemble de l'échantillon de bassins versants. L'effet le plus évident de l'application des deux méthodes d'analyse de sensibilité est une tendance de décroissance nette dans l'efficacité de la régionalisation. Le second résultat important est que le désert hydrométrique (HDes) semble fournir une décroissance plus abrupte que la méthode de réduction aléatoire (HALéa). La raison semble résider dans le fait que la méthode HALéa permet des situations où le premier donneur est assez proche du receveur (et potentiellement plus similaire et meilleur donneur), comme nous pouvons le constater sur la Figure 24. Parce qu'il oblige l'exclusion des plus proches donneurs, le désert hydrométrique semble plus exigeant.

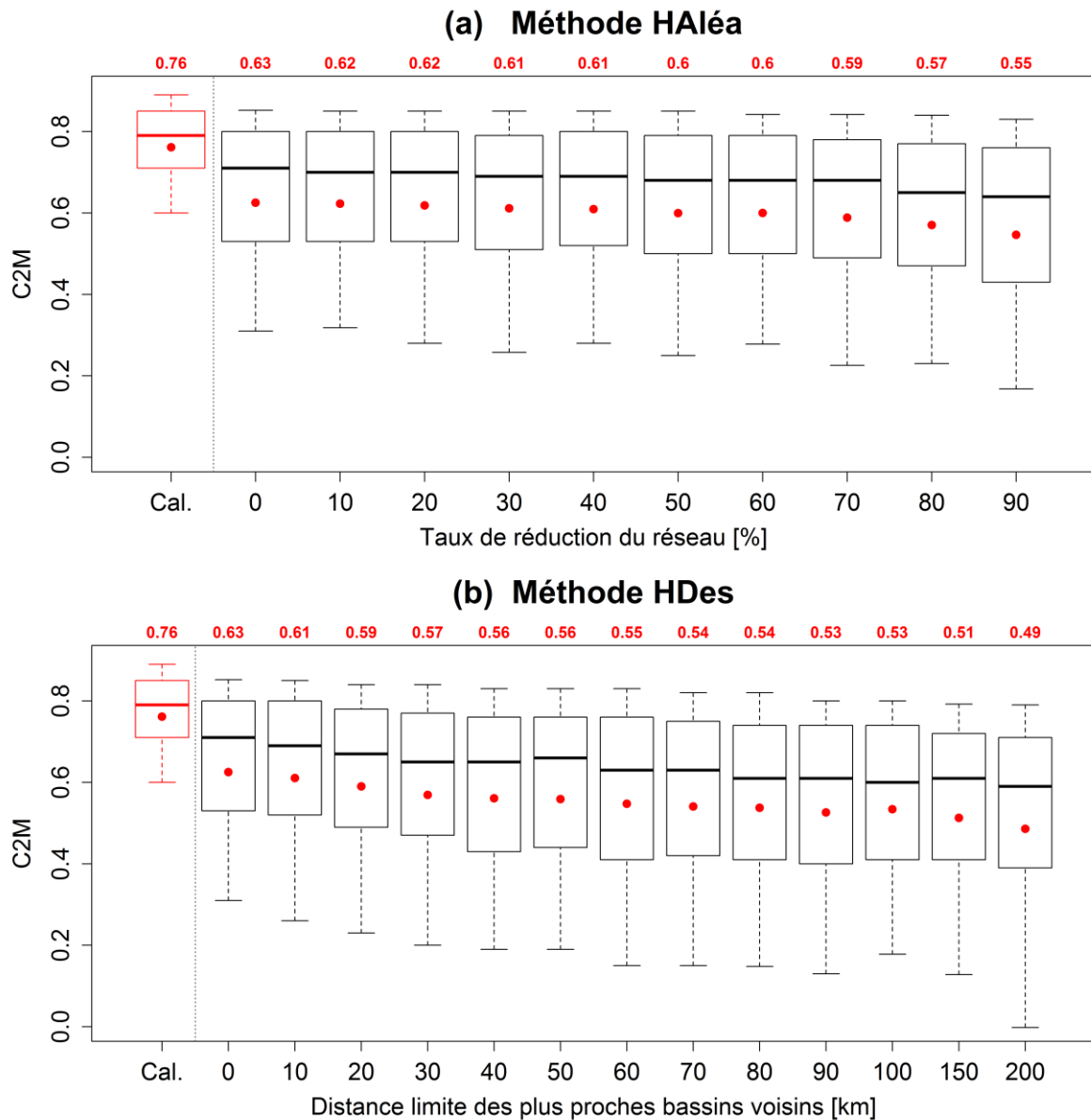


Figure 23. Distributions des performances du modèle GR4J calé et régionalisé sur les 609 bassins, montrant l'impact des deux méthodes d'évaluation de robustesse (a : la méthode de réduction aléatoire du réseau – HALéa ; b : la méthode du désert hydrométrique – HDes). Les résultats de calage (à gauche) sont utilisés comme référence. Les boîtes à moustaches montrent les quantiles 10, 25, 50, 75 et 90 de la distribution, du bas vers le haut. Les performances moyennes sont indiquées par un point et la valeur correspondante est montrée en haut du graphique.

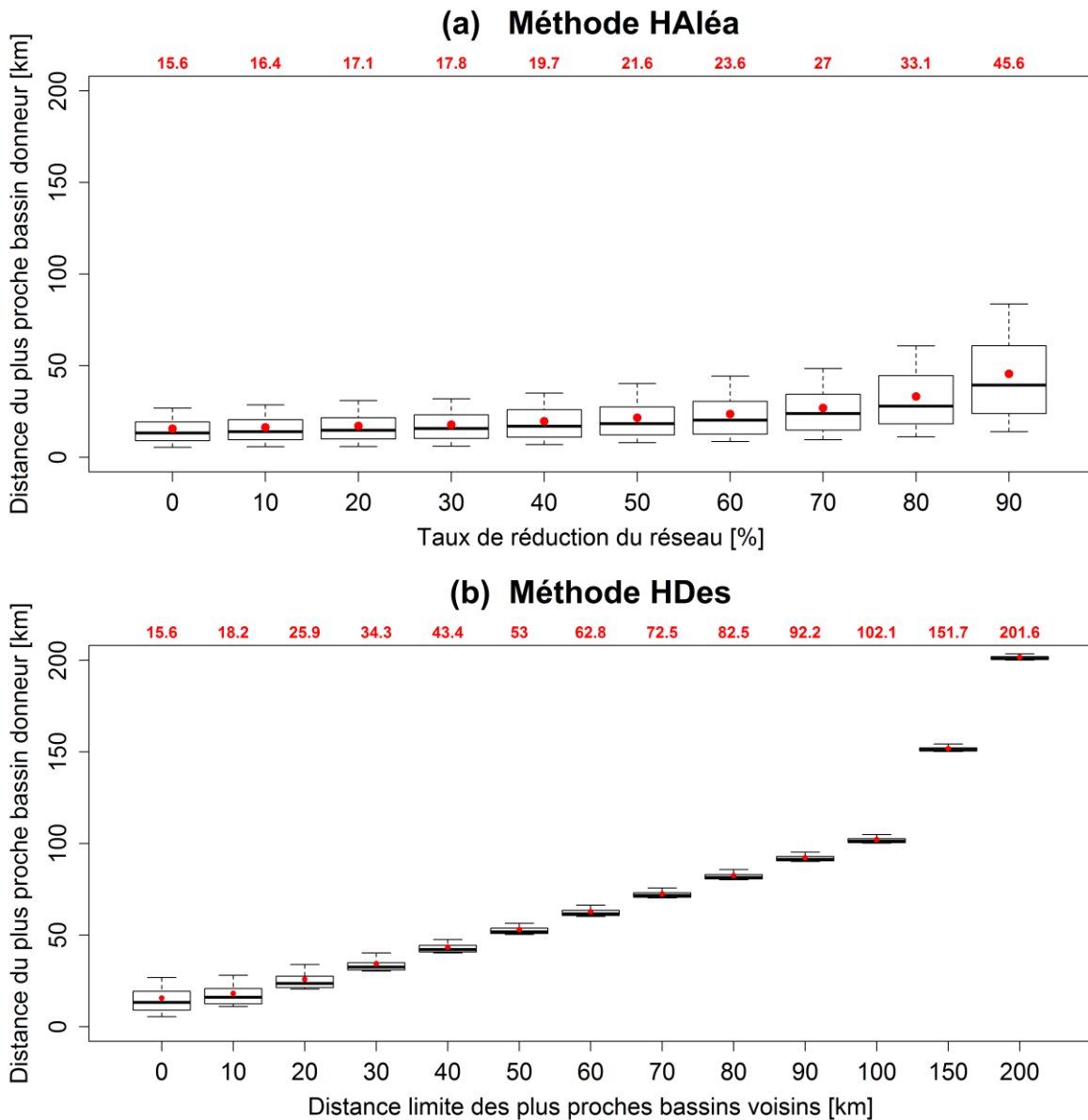


Figure 24. Distributions de la distance du plus proche bassin donneur pour les deux méthodes utilisées pour évaluer la robustesse des méthodes de régionalisation (a : la méthode de réduction aléatoire du réseau – HALéa ; b : la méthode du désert hydrométrique – HDes). Les boîtes à moustaches montrent les quantiles 10, 25, 50, 75 et 90 de la distribution, du bas vers le haut. Les distances moyennes sont indiquées par un point et la valeur correspondante est montrée en haut du graphique.

Toutefois, la comparaison entre les deux méthodes semble difficile en raison d'axes des abscisses différents. Afin de rendre les méthodes HALéa et HDes comparables malgré des unités différentes (taux de réduction pour HALéa et distance limite des plus proches voisins pour HDes), il nous fallait trouver une échelle commune pour les deux méthodes. L'évolution de la distance moyenne des dix bassins donneurs peut être obtenue pour les deux méthodes et ainsi devenir leur dénominateur commun. La Figure 25 présente la distribution de ces distances utilisées pour chaque réduction de densité pour la méthode HALéa (a) et la méthode HDes (b). La distance moyenne des dix bassins

donneurs évolue en médiane d'environ 35 km à 120 km pour la méthode HALéa, alors que pour la méthode HDes les distances médianes évoluent d'environ 35 km à 200 km.

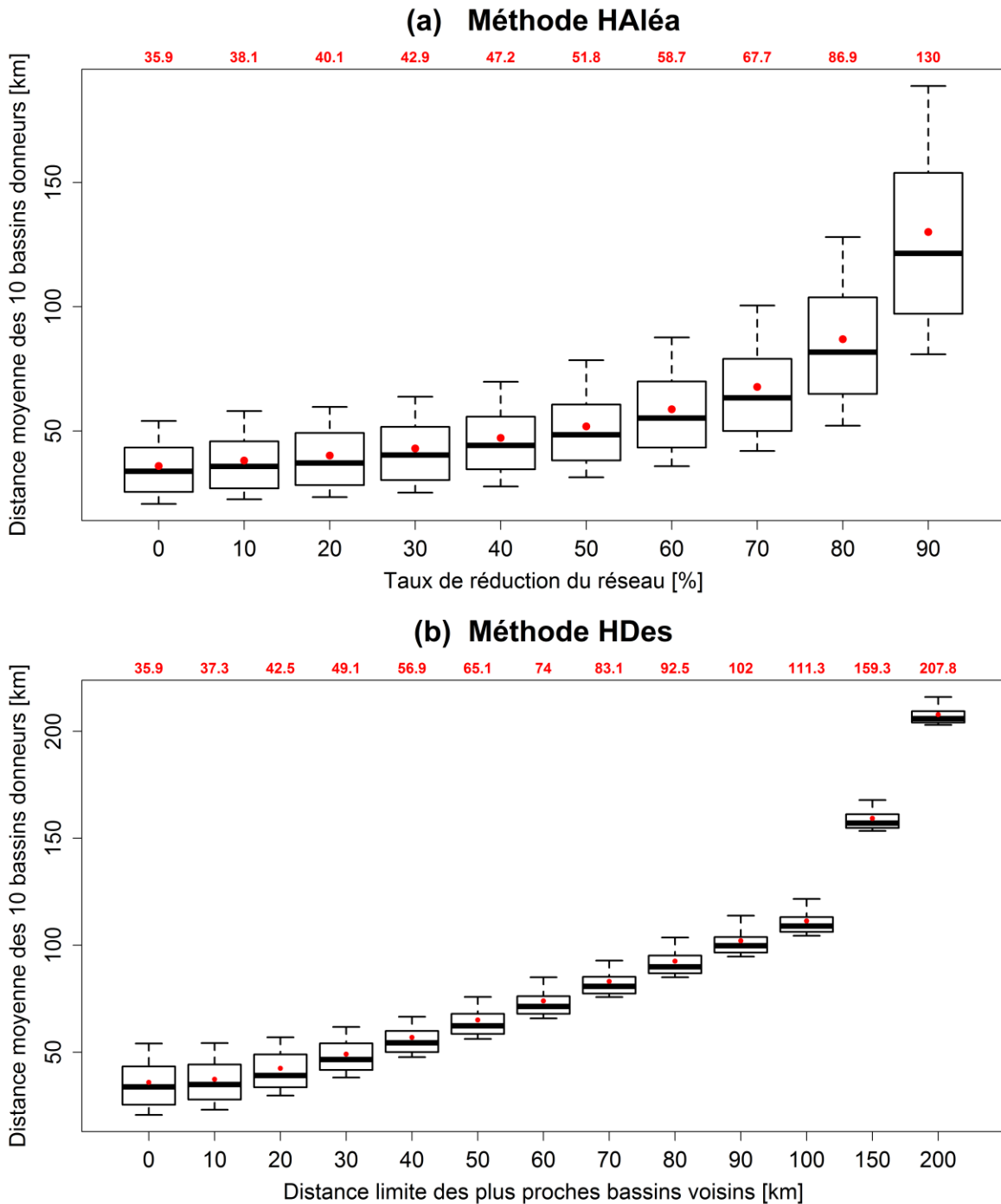


Figure 25. Distributions de la distance moyenne des 10 bassins donateurs pour les deux méthodes utilisées pour évaluer la robustesse des méthodes de régionalisation (a : la méthode de réduction aléatoire du réseau – HALéa ; b : la méthode du désert hydrométrique – HDes). Les boîtes à moustaches montrent les quantiles 10, 25, 50, 75 et 90 de la distribution, du bas vers le haut. Les distances moyennes sont indiquées par un point et la valeur correspondante est montrée en haut du graphique.

Nous allons donc prendre comme nouvel axe d'abscisses commun aux deux méthodes, la distance moyenne des 10 bassins donateurs. Celle-ci sera associée aux valeurs du taux de réduction du réseau pour HALéa, et à la distance limite des bassins donateurs pour HDes. Ces résultats sont présentés sur la Figure 26. Ainsi, nous pouvons confirmer ce qui a été remarqué plus haut à l'aide de ces graphiques puisque nous observons une perte de performances, un peu plus abrupte pour la méthode HDes que pour la méthode HALéa, notamment pour les plus faibles distances moyennes des dix donateurs.

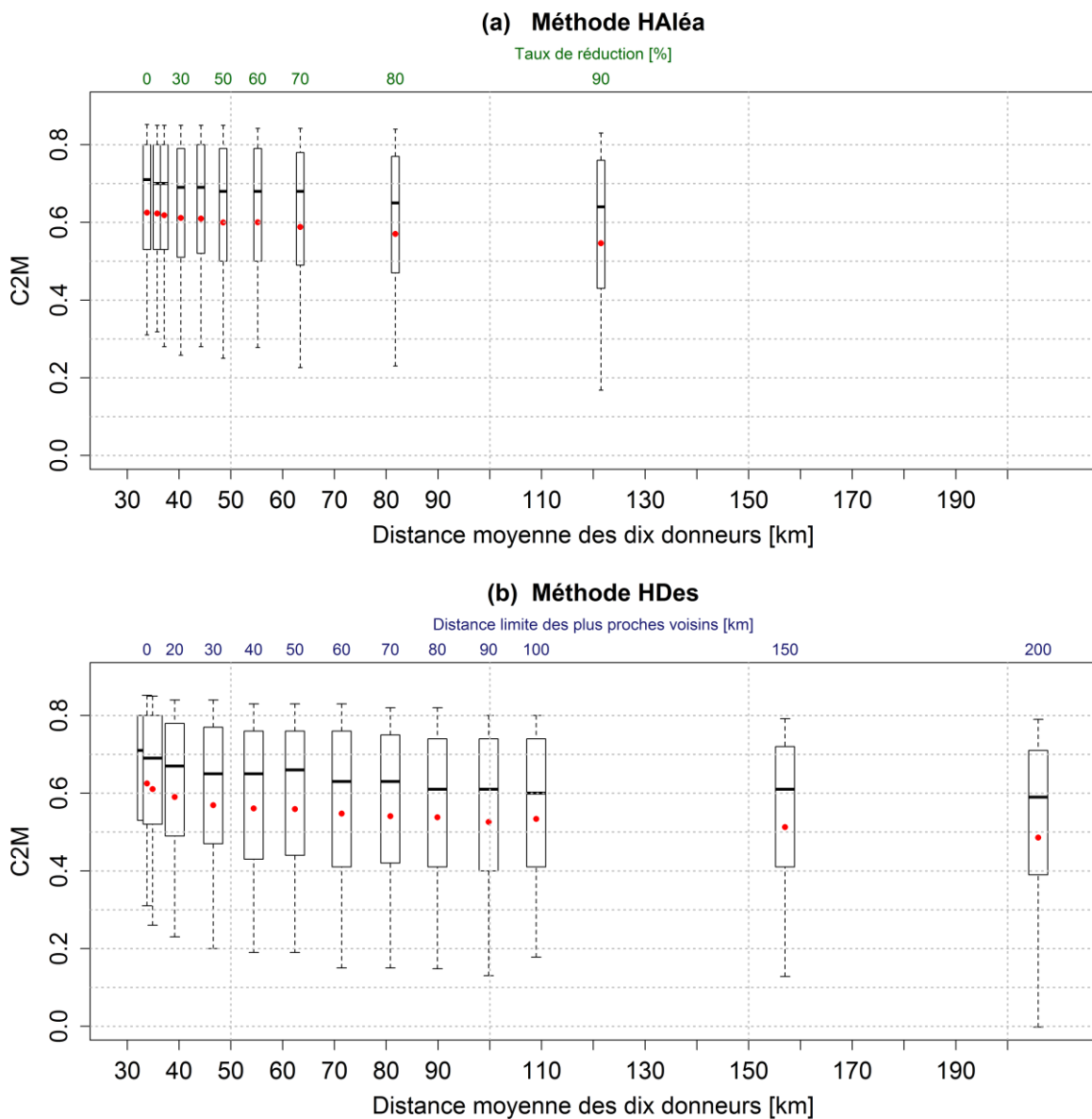


Figure 26. Distributions des performances du modèle GR4J régionalisé sur les 609 bassins en fonction de la distance moyenne des dix donateurs pour les deux méthodes d'évaluation de robustesse (a : la méthode de réduction aléatoire du réseau – HALéa ; b : la méthode du désert hydrométrique – HDes). Les boîtes à moustaches montrent les quantiles 10, 25, 50, 75 et 90 de la distribution, du bas vers le haut. Les valeurs de taux de réduction pour la méthode HALéa et des distances limites des plus proches voisins pour la méthode HDes correspondantes sont indiquées en haut du graphique.

Pour mieux évaluer le niveau d'exigence des méthodes relativement à la dispersion du réseau de bassins traité ici, nous avons calculé la distribution de toutes les distances possibles entre les 609 bassins versants, c'est-à-dire $(609 \times 608)/2$ soit environ 185000 combinaisons possibles de bassins (Figure 27). Ainsi, nous pouvons constater que les distances de bassins comprises entre 0 et 200 km (distance limite maximale pour HDes) se situent en deçà du quantile 20% des distances possibles entre les 609 bassins versants de notre échantillon. On constate donc que les tests d'éloignement des donneurs restent à une échelle que l'on peut qualifier de régionale relativement à l'ensemble de l'échantillon. Pour une application de la méthode à d'autres zones d'études, le choix de la gamme des distances d'éloignement pourrait nécessiter des ajustements relativement aux caractéristiques du réseau et de la zone étudiée, en prenant en compte cette distribution des distances entre bassins et la distribution des distances des donneurs (cf. Figure 22).

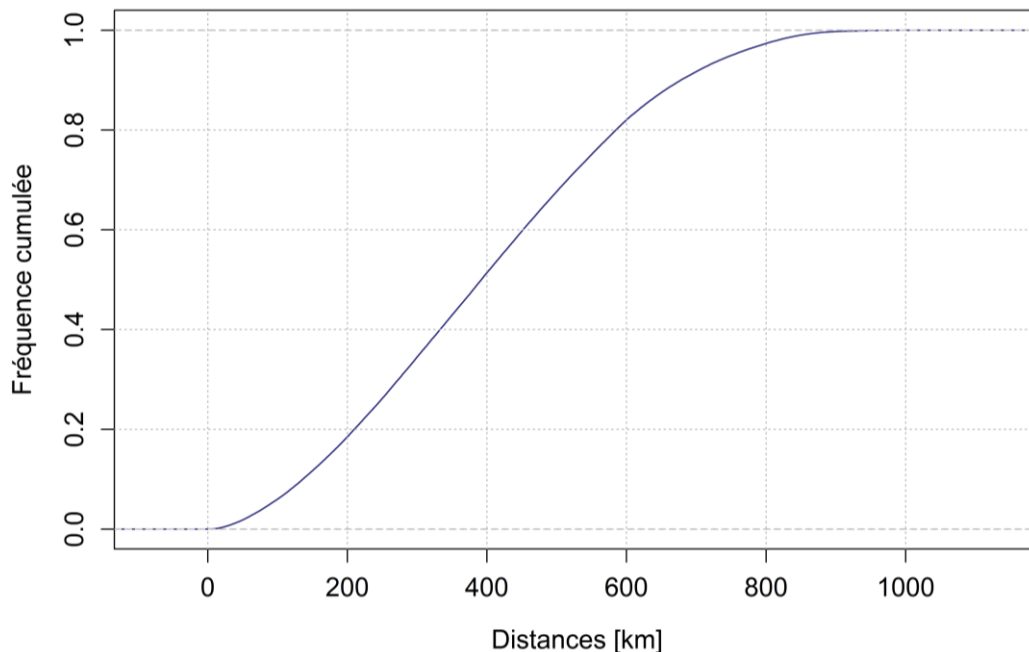


Figure 27. Distribution de toutes les distances possibles entre les 609 bassins versants.

3.6.4 Conclusion

Bien que les deux méthodes permettent d'analyser et quantifier la sensibilité des méthodes de régionalisation à la décroissance de la disponibilité des données hydrométriques, nous avons pu observer que la méthode HDes est plus exigeante, puisqu'elle produit la décroissance de l'efficacité du modèle la plus rapide. Dans une perspective de « crash test » hydrologique, comme discuté par *Andréassian et al.* [2009], nous recommanderons d'utiliser la méthode HDes qui, nous le pensons, fournit un point de vue plus réaliste (même s'il est plus pessimiste) de l'évolution de la robustesse d'une régionalisation basée sur la proximité spatiale. Une autre raison est que, pour une application opérationnelle sur un bassin versant non-jaugé donné, il est toujours plus facile de calculer la distance entre ce bassin et son voisin le plus proche que d'évaluer la densité régionale des bassins versants. Une fois cette distance connue, la Figure 23-b peut être utilisée pour donner une indication de la valeur attendue de l'efficacité de GR4J régionalisé, avec un intervalle d'incertitude sur la valeur attendue.

3.7. Synthèse

Afin de répondre aux questions concernant l'impact de la densité du réseau hydrométrique en hydrologie, nous avons constitué un large échantillon de 609 bassins versants pour lesquels des données pluviométriques, hydrométriques et d'évapotranspiration potentielle sont disponibles au pas de temps journalier sur une période minimale commune de 10 ans. Les bassins versants disposent également de chroniques de débit avec au minimum 25 ans de données disponibles afin de pouvoir calculer les débits caractéristiques, tels que le module, les débits d'étiage et les débits de crue. L'analyse des caractéristiques morphologiques et hydro-climatiques nous indiquent que ces bassins présentent des caractéristiques variées et couvrent l'ensemble du territoire français.

Dans ce chapitre, nous avons également présenté les différents éléments de méthodologie utiles pour les tests sur l'impact de la densité du réseau hydrométrique sur les calculs hydrologiques. En résumé, les éléments retenus sont les suivants :

- Le modèle hydrologique considéré est le modèle pluie-débit global GR4J [Perrin *et al.*, 2003], couplé au module neige CemaNeige [Valéry, 2010]. Ces outils de modélisation ont été largement utilisés lors de recherches antérieures menées dans l'équipe Hydrologie d'Irstea Antony.
- En raison de son efficacité sur les bassins versants français due à une forte densité de stations en métropole, la méthode de régionalisation sélectionnée pour répondre aux objectifs de recherche fixés dans cette partie, est celle basée sur la proximité spatiale. Elle consiste à transférer les paramètres du modèle des bassins voisins vers le bassin versant non jaugé. Les différentes chroniques de débit ainsi obtenues sont par la suite moyennées afin d'obtenir une chronique de débit unique sur le bassin cible.
- Le choix de la distance séparant deux bassins versants est important dans la sélection des bassins voisins utiles à la régionalisation. Nous avons retenu ici la distance hydrologique combinant la distance séparant les exutoires de deux bassins et celle séparant leurs centroïdes, en donnant davantage de poids à cette dernière. La distance hydrologique s'écrit alors de la façon suivante :

$$d = 0.2 \times d_{\text{exutoire}} + 0.8 \times d_{\text{centroïde}} \quad \text{Eq. 7}$$

- Nous étudierons la robustesse de la méthode de régionalisation vis-à-vis de la disponibilité spatiale des mesures de débit, à l'aide de la méthode du désert hydrométrique, qui consiste à sélectionner des bassins voisins de plus en plus éloignés du bassin cible non jaugé. Cette méthode est beaucoup plus exigeante et permet une application plus opérationnelle.

Chapitre 4.

Une régionalisation de la formule de bilan en eau de Turc-Mezentsev

Présentation et résumé de l'article en français

Ce chapitre est basé sur un article publié dans la revue *Water Resources Research* :

Lebecherel, L., V. Andréassian & C. Perrin. 2013. On regionalizing the Turc-Mezentsev water balance formula. Water Resources Research, 49, doi:10.1002/2013WR013575

Celui-ci porte sur l'évaluation des possibilités de régionalisation d'un calcul hydrologique bien particulier, celui de l'estimation du module à l'aide de la formule de Turc-Mezentsev. Mais au-delà, l'article évalue également la robustesse des options de régionalisation testées, vis-à-vis de la densité spatiale du réseau hydrométrique. Il s'agit d'un travail préliminaire aux travaux de cette thèse à proprement parler, mais qui a permis de bien poser les concepts, notamment les modalités d'évaluation de la robustesse dans le cas de bassins versants non-jaugés

La formule de bilan hydrique de Turc-Mezentsev est une formule mathématique simple et efficace reliant le débit moyen à long terme aux précipitations et évapotranspiration potentielle moyennes à long terme :

$$\frac{Q}{P} = 1 - \frac{1}{\left[1 + \left(\frac{P}{E_0}\right)^n\right]^{\frac{1}{n}}}$$

Habituellement, elle est appliquée sous sa forme originale, c'est-à-dire sans paramétrisation, et même si dans sa version originale, elle dispose d'un paramètre libre (n), ce dernier est souvent fixé ($n=2$ et $n=2.5$ sont les valeurs les plus courantes).

Dans l'article, nous avons cherché à évaluer les possibilités de régionalisation de cette formule, afin d'améliorer son efficacité sur des bassins non jaugés. Quatre options de régionalisation ont été comparées, basées respectivement sur : (i) l'erreur relative, (ii) l'erreur absolue, (iii) le facteur correcteur de l'indice d'aridité et (iv) le facteur de forme (c'est-à-dire le paramètre n). Les quatre alternatives ont été évaluées en termes d'efficacité et de robustesse sur un large échantillon de bassins versants français (celui présenté dans le chapitre précédent).

Globalement, cette étude montre que c'est en introduisant un facteur corrigeant l'indice d'aridité que l'on obtenait la meilleure régionalisation. Si on entre un peu plus dans les détails, on se rend compte que si la solution du facteur de forme n n'est pas retenue, c'est essentiellement à cause des bassins versants non conservatifs (c'est-à-dire des bassins qui ne « perdent » pas seulement de l'eau par évaporation et écoulement de rivière, mais qui « fuient » vers un aquifère régional). Enfin, l'étude montre, quelle que soit l'option de régionalisation retenue, l'importance de la densité du réseau de bassins jaugés pour maximiser les performances de la régionalisation.

On regionalizing the Turc-Mezentsev water balance formula

L. Lebecherel¹, V. Andréassian¹ and C. Perrin¹

Water Resources Research, 49(11), 7508-7517, doi: 10.1002/2013wr013575.

¹ Irstea, Hydrosystems and Bioprocesses Research Unit, Antony, France

Abstract

The Turc-Mezentsev water balance formula is a simple and efficient mathematical formulation relating long-term average streamflow to long-term average precipitation and potential evaporation. Usually, it is applied in its original form, i.e., without parameterization. This paper aims to assess the possibilities of regionalizing this formula for improved efficiency on ungaged catchments. Four regionalization options are compared based on: (i) the relative error, (ii) the arithmetic difference, (iii) the correcting factor of the aridity index and (iv) the shape factor. The four alternatives are assessed in terms of efficiency and robustness on a large set of 609 French catchments. We show that introducing a correcting factor of the aridity index provides the best regionalization solution.

4.1. Introduction

4.1.1 The need for a reliable assessment of long-term average streamflow

Knowledge of long-term average streamflow is at the basis of any water resource assessment. Actual measurements provide the best information, but given that the number of existing streamgages is limited and that the climate may not be as stationary as we once thought, precipitation-flow type methods can also be useful: based on an estimation of climatic forcings and catchment behavior, these methods allow simulating long-term average streamflow.

A variety of models and formulas exist to represent the precipitation-flow transformation. This paper will focus on one that is both simple and efficient: the Turc-Mezentsev formula. Our aim is to improve its predictions in ungaged situations, through regionalization of either its parameter or its residuals.

4.1.2 Scope of the paper

The paper starts with a short review of the origins of the Turc-Mezentsev formula and its applications in hydrology (section 4.2). Then the methodology is detailed: section 4.3.1 presents the data set of 609 French catchments on which our regionalization exercise will be conducted. We discuss the different alternatives that can be proposed for regionalizing the formula in section 4.3.2, and section 4.3.3 presents the quantitative criteria selected to assess the success of the regionalization procedure. A specific experimental setting ensuring the robustness of the regionalization procedure is also discussed in section 4.3.4. Section 4.4 presents the results, compares the efficiency and the robustness of the alternatives, and last, the paper concludes with perspectives for using this formula.

4.2. Origins and applications of the Turc-Mezentsev formula

4.2.1 Origins of the Turc-Mezentsev formula

In the 1950s, a formula linking long-term average evaporation to long-term average precipitation and potential evaporation was quasi-simultaneously proposed in France and in the Soviet Union. Because Turc and Mezentsev published in French and Russian, respectively, their work has received limited attention. A short account of their published work is provided below.

❖ Lucien Turc's work in France [1954]

Lucien Turc, a French soil scientist, worked at the Versailles soil science laboratory of the French National Agronomical Research Institute. For agronomic purposes, he was interested in developing a formula to estimate actual evaporation from precipitation and temperature data. Since lysimeter data were so scarce at that time, he came up with the idea of deriving such a formula from the catchment water balance. With the help of the hydrologist Maurice Pardé, he assembled a set of data on 254 catchments throughout the world. He derived long-term average actual evaporation (E) from estimates of long-term average precipitation (P) and long-term average discharge (Q) by writing (all variables in mm/yr):

$$E = P - Q \quad \text{Eq. 8}$$

Then Turc looked for a mathematical function linking two nondimensional terms: the aridity index (the ratio of precipitation P to potential evaporation E_0) and the evaporation rate (the ratio of E to E_0). He mentioned that the formulation should respect the two following constraints:

- $(P - Q)/E_0 \sim P/E_0$ when P/E_0 is small (i.e., $Q \sim 0$ when $P \ll E_0$)
- $(P - Q)/E_0 \sim 1$ when P/E_0 is large (i.e., $Q \sim P - E_0$ when $P \gg E_0$)

Turc (1954, p. 504) writes that the simplest function respecting these two conditions would be $y = x/(1 + x')$ with $y = (P - Q)/E_0$ and $x = P/E_0$ and the most general would be $y = x/[(1 + x^n)^{1/n}]$, in which n is an exponent to estimate. Turc graphically looked for the most convenient value for n and concluded that the best fit was "with $n=3$, or maybe $n=2$ " (Turc, 1954, p.563). Since the choice of $n=2$ allowed the simplest computations, he finally retained this value and the corresponding formulation:

$$\frac{E}{P} = \frac{1}{\left[1 + \left(\frac{P}{E_0}\right)^2\right]^{\frac{1}{2}}} \quad \text{Eq. 9}$$

❖ Varfolomeï Mezentsev's work in the Soviet Union [1955]

Mezentsev [1955] worked at the University of Omsk in Siberia. He started his analysis from a formula proposed by *Bagrov* [1953]:

$$\frac{dE}{dP} = 1 - \left(\frac{E}{E_0}\right)^n \quad \text{Eq. 10}$$

The Bagrov formula can be interpreted as follows: when E/E_0 is small, i.e., when water is the limiting factor, an increase in precipitation P is transformed into an increase in actual evaporation E . Conversely, when E/E_0 approaches 1, i.e., water does not limit evaporation, none of the additional P is transformed into E because no more energy is available for evaporation. This formula presents the appealing property of integrating into the *Schreiber* [1904] water balance formula for $n=1$:

$$E = P \cdot \left[1 - \exp\left(-\frac{E_0}{P}\right)\right] \quad \text{Eq. 11}$$

and into the *Ol'Dekop* [1911] water balance formula for $n=2$:

$$E = E_0 \cdot \tanh\left(\frac{P}{E_0}\right) \quad \text{Eq. 12}$$

but it has no analytical solution for other values of n .

Mezentsev [1955] modified Bagrov's formula by rewriting it as follows:

$$\frac{dE}{dP} = \left[1 - \left(\frac{E}{E_0}\right)^n\right]^{1+\frac{1}{n}} \quad \text{Eq. 13}$$

with the same interpretation as equation 10.

Equation 13 can be integrated analytically and yields equation 14 :

$$\frac{E}{P} = \frac{1}{\left[1 + \left(\frac{P}{E_0}\right)^n\right]^{\frac{1}{n}}} \quad \text{Eq. 14}$$

which is identical to the general formulation proposed by Turc. Based on a set of 35 catchments on the Siberian plateau, Mezentsev suggested using the value of 2.3 for parameter n , which is close to the value recommended by Turc.

This general formulation will hereafter be referred to as the Turc-Mezentsev formula.

❖ Related formulas

The Turc-Mezentsev formula is mathematically different but very close in its result to a formula proposed by *Fu* [1981] (see *Yang et al.* [2008] for a comparison of the formulas). It is also a different expression from the well-known *Budyko* [1974] formula, which actually corresponds to the geometric average of the above-mentioned water balance formulas proposed by *Schreiber* [1904] and *Ol'Dekop* [1911]. Note that *Donohue et al.* [2011] have shown that Budyko formula can be approximated by the Turc-Mezentsev formula with $n=1.9$.

4.2.2 Physical and mathematical interpretation of the Turc-Mezentsev formula

In proposing his formula, Turc sought two first-order mathematical properties:

$$\frac{E}{P} \rightarrow 1 \text{ when } \frac{P}{E_0} \rightarrow 0$$

$$\frac{E}{P} \rightarrow 0 \text{ when } \frac{P}{E_0} \rightarrow \infty$$

Similarly, Mezenstev sought:

$$\frac{E}{E_0} \rightarrow 0 \text{ and } \frac{Q}{P} \rightarrow 0 \text{ when } \frac{P}{E_0} \rightarrow 0$$

$$\frac{E}{E_0} \rightarrow 1 \text{ and } \frac{Q}{P} \rightarrow 1 \text{ when } \frac{P}{E_0} \rightarrow \infty$$

Yang *et al.* [2008] underlined that the Turc-Mezentsev formula has additional second-order mathematical properties:

$$\frac{\partial E}{\partial P} \rightarrow 0 \text{ when } \frac{P}{E_0} \rightarrow \infty$$

$$\frac{\partial E}{\partial E_0} \rightarrow 0 \text{ when } \frac{E_0}{P} \rightarrow \infty$$

$$\frac{\partial E}{\partial P} \rightarrow 1 \text{ when } P \rightarrow 0$$

$$\frac{\partial E}{\partial E_0} \rightarrow 1 \text{ when } E_0 \rightarrow 0$$

4.2.3 Transposing the Turc-Mezentsev formula for streamflow

The Turc-Mezentsev formula has mostly been used to compute actual evaporation, but it is based on streamflow data and can thus be transposed to streamflow by writing $E = P - Q$. It becomes:

$$\frac{Q}{P} = 1 - \frac{1}{\left[1 + \left(\frac{P}{E_0}\right)^n\right]^{\frac{1}{n}}} \quad \text{Eq. 15}$$

Table 1 details the mathematical properties of this streamflow formulation, with the corresponding hydrological interpretation (the corresponding partial differential equations are presented in Annex 1).

Table 1. Hydrological interpretation of the mathematical properties of the Turc-Mezentsev formula (applied to streamflow). Notations: E_0 , long-term average potential evaporation; P , long-term average precipitation; Q , long-term average discharge (all variables in mm/yr).

	Mathematical property	Hydrological interpretation
1	$Q < P$	A catchment cannot produce more water than it receives from precipitation
2	$Q/P \rightarrow 1$ when $P \gg E_0$	The water yield of very wet catchments tends towards 1
3	$P - Q < E_0$	A catchment cannot lose more water than it receives energy to evaporate water
4	$P - Q \rightarrow E_0$ when $P \gg E_0$	The evaporation losses of very wet catchments tend towards potential evaporation
5	$\frac{\partial Q}{\partial P} \rightarrow 1$ when $\frac{P}{E_0} \rightarrow \infty$	On very wet catchments, all additional precipitation is transformed into streamflow
6	$\frac{\partial Q}{\partial E_0} \rightarrow -1$ when $E_0 \rightarrow 0$	On low-energy catchments, any increase of E_0 reflects an equivalent decrease of Q .
7	$\frac{\partial Q}{\partial E_0} \rightarrow 0$ when $\frac{E_0}{P} \rightarrow \infty$	On high-energy catchments, streamflow is not affected by an augmentation of E_0
8	$\frac{\partial Q}{\partial P} \rightarrow 0$ when $P \rightarrow 0$	On very dry catchments, none of the additional precipitation is transformed into streamflow

The curve representing equation 15 also has noticeable graphical properties. On the nondimensional graph relating the runoff coefficient (Q/P) to the aridity index (P/E_0) [Andréassian et Perrin, 2012], the family of Turc-Mezentsev curves lies within the limits (water, energy and measurement limits) defining the white domain shown in Figure 28.

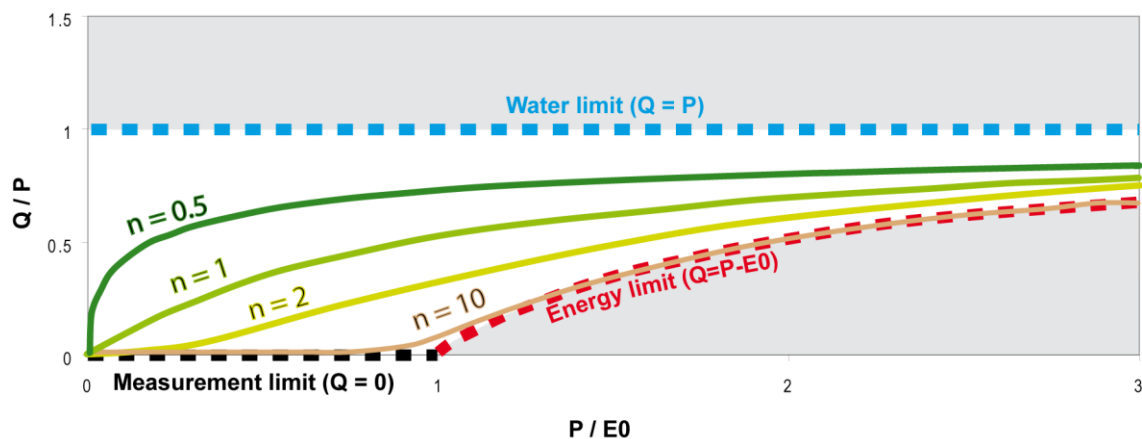


Figure 28. Domain occupied by the family of Turc-Mezentsev curves: for large n values, curves will tend towards the lower limit (zero flow limit in black and energy limit in red); for small n values, curves will tend towards the water limit (in blue). Shaded areas represent the part of the nondimensional space that is physically impossible according to the Turc-Mezentsev representation.

4.2.4 Recent studies based on the Turc-Mezentsev formula

Considerable interest has arisen over the past few years in the use of the Turc-Mezentsev formula. Table 2 lists a few key studies:

Table 2. Key studies based on the Turc-Mezentsev formula

Reference	Aim	Use of the Turc-Mezentsev formula
<i>Dooge</i> [1992]	Derive a general expression of the sensitivity of runoff to climate change.	Used a fixed $n=2$.
<i>Choudhury</i> [1999]	Derive long-term average actual evaporation.	On 10 experimental sites, n is set at 2.6 based on micro-meteorological measurements. On 10 continental-scale catchments, n is set at 1.8 based on results of a biophysical model.
<i>Sankarasubramanian et al.</i> [2001]	Parameterize the 'abc' model and compute the climate elasticity of streamflow in the US.	Used a fixed $n=2$.
<i>Arora</i> [2002]	Compare several long-term formulas to assess the impact of aridity changes on annual runoff.	Used a fixed $n=2$.
<i>Mouelhi et al.</i> [2006]	Base their annual rainfall-runoff model on the Turc-Mezentsev formula.	Used a fixed $n=2$.
<i>Le Moine et al.</i> [2007]	Examine the possible solutions to account for unknown intercatchment groundwater flows.	n is considered as a free parameter, a second parameter α is added to the Turc-Mezentsev formula.
<i>Oudin et al.</i> [2008a]	Investigate the relation between land cover and mean annual streamflow.	Used a fixed $n=2$. Compared several long-term formulas.
<i>Potter et Zhang</i> [2009]	Compare several long-term formulas to determine whether they could also be used to represent the interannual variability of water balance in Australia.	Used a fixed $n=2$.
<i>Roderick et Farquhar</i> [2011]	Relate variations in streamflow to variations in climatic conditions and catchment properties.	Found an average $n=1.7$ for the Murray-Darling basin.
<i>Donohue et al.</i> [2011]	Propose plots of runoff sensitivity to aridity change as a function of n .	n is considered as a free parameter for the Murray-Darling basin.
<i>Donohue et al.</i> [2012]	Test a new approach combining ecohydrological and Turc-Mezentsev models to estimate n .	n is estimated using a combination of two pre-existing models

Several of the studies mentioned in Table 2 attempted to adjust the Turc-Mezentsev formula. Two approaches were tested.

❖ Inferring parameter n from catchment water balance

Choudhury [1999] concluded that “factors determining the value of n are not fully understood and no physical explanation is apparent for the large discrepancies” between catchments.

On 1040 French catchments, *Le Moine et al.* [2007] found that water balance elements converge toward a Turc-Mezentsev formula corresponding to $n=2.5$. Note that these authors added a second parameter α to the formula, as a multiplicative parameter to E_0 .

Donohue et al. [2011] proposed maps of the n value for different subcatchments in the Murray-Darling basin.

Donohue et al. [2012] presented a combination of the ecohydrological model of *Porporato et al.* [2004] and the Turc-Mezentsev model. This new model estimates the parameter n from effective rooting depth, and differentiates between trees and grass.

❖ Keeping n constant but introducing a correction parameter that acts as a multiplier of E_0/P in the formula

Oudin et al. [2008a] attempted to modify several water balance formulas, including the Turc-Mezentsev formula with $n=2$, and to relate this adjustment factor to the type of land cover (they used four classes: forest, arable land, shrubland and grassland). They found that adding a land-cover-specific correcting factor slightly improved the efficiency of the formula, although not all categories were equally informative.

4.3. Catchments and method

4.3.1 Data set

To test the different regionalization options, we used a data set consisting of 609 French catchments, where daily streamflow, rainfall and potential evaporation time series were available over at least 30 years. These catchments are spread over France and encompass various climatic conditions (oceanic, Mediterranean, continental).

The potential evaporation is calculated for all catchments with *Oudin's* formula [*Oudin et al.*, 2005] and precipitation data comes from a countrywide interpolation of raingage data (SAFRAN analysis) [*Quintana-Seguí et al.*, 2008; *Vidal et al.*, 2010].

Figure 29 shows the position of the 609 catchments on the nondimensional graph presented in Figure 28, and Figure 30 shows their location. Eighty percent of this data set (490 catchments) is conservative in the sense used by *Le Moine et al.* [2007]: these catchments fall strictly within the water and energy limits presented in Figure 28. A total of 119 catchments are located on the part of the nondimensional space that is physically impossible according to the Turc-Mezentsev representation. These catchments either “leak” (i.e., those catchments located below the energy limit) or “gain” water (those located above the water limit). They will be referred to as the nonconservative part of the data set.

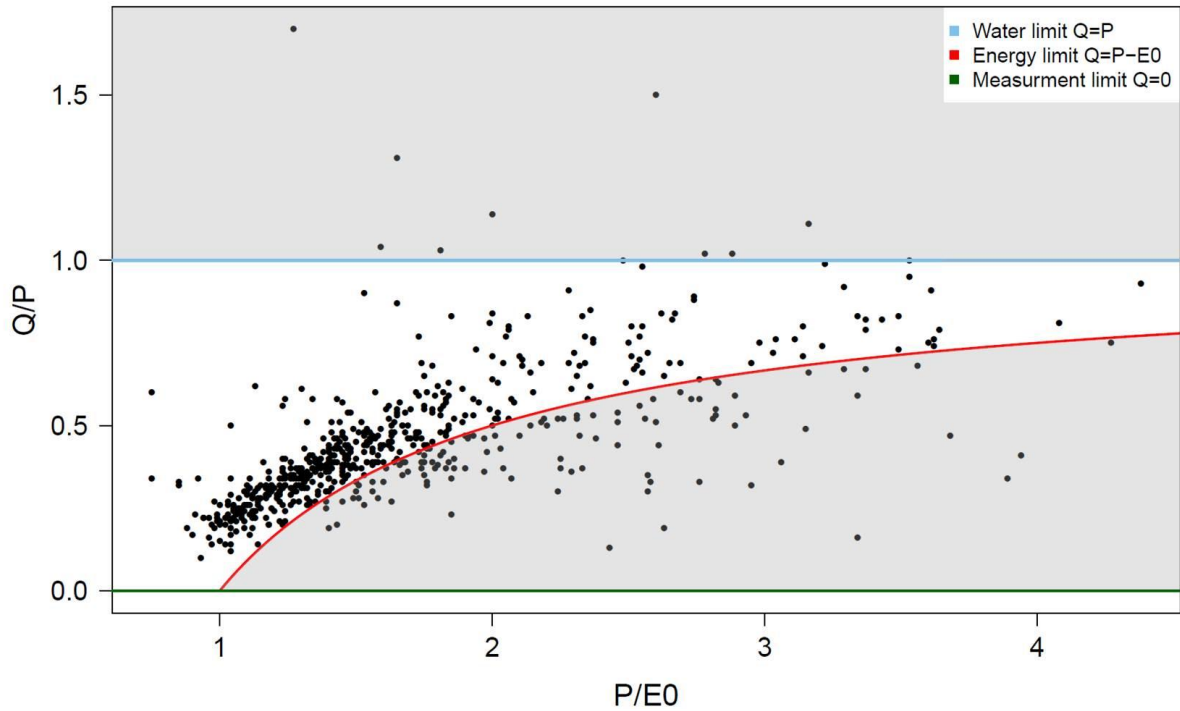


Figure 29. Position of catchments studied on the nondimensional graph relating the runoff coefficient (Q/P) and aridity index ($P/E0$). Shaded areas represent the part of the nondimensional space that is physically impossible according to the Turc-Mezentsev representation.

To meet the objectives of this study, we did not wish to do any data set cleansing [Boldetti et al., 2010] and to remove the catchments presenting water balance problems. We first worked on the conservative part of the data set, i.e., on those catchments that fall strictly within the water and energy limits presented in Figure 28. Then we checked whether the conclusions hold on the catchments that fall outside these limits. We did this in order to provide a more realistic (perhaps less optimistic) view of the possibilities of the regionalization scheme. Figure 30 shows that these catchments mostly correspond to mountainous areas. However, note that although nonconservative catchments are mostly found in mountainous areas, not all mountain catchments are nonconservative.

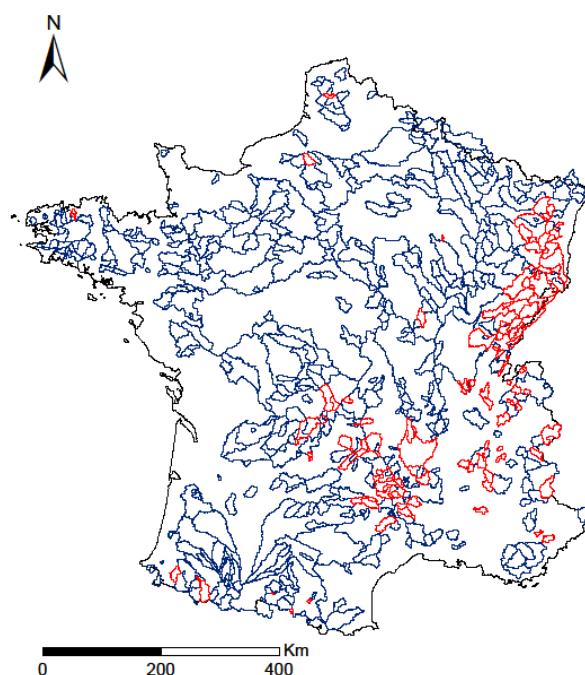


Figure 30. Location of the 609 catchments used in this study. In blue, the conservative basins and in red, the nonconservative catchments (which either leak to or gain water from adjacent catchments, see Figure 29).

Table 3 gives the main characteristics of the two data sets in terms of catchment area, mean annual streamflow, precipitation and potential evaporation.

Table 3. Main characteristics of the two parts of the data set (490 conservative catchments vs 119 nonconservative catchments)

	Quantiles				
	0.00	0.25	0.50	0.75	1.00
Conservative part of the data set					
Area (km ²)	5	118	304	908	111975
Mean annual precipitation, P (mm/year)	620	832	944	1121	2112
Mean annual runoff, Q (mm/year)	67	249	383	579	2020
Mean annual potential evaporation, E ₀ (mm/year)	312	632	661	693	875
Nonconservative part of the data set					
Area (km ²)	10	80	183	435	5169
Mean annual precipitation, P (mm/year)	848	1084	1256	1533	2066
Mean annual runoff, Q (mm/year)	180	388	523	813	2296
Mean annual potential evaporation, E ₀ (mm/year)	289	570	611	646	775

4.3.2 Testing geographical (spatial proximity-based) regionalization

Regionalization based on spatial proximity consists in transferring information from a gaged catchment to a neighboring ungaged catchment [Oudin *et al.*, 2008b]. Hereafter, the ungaged catchment (where no streamflow data are available) is called the target or receiver, while the gaged catchment is referred to as the donor. The idea behind this transfer is that the catchments close to each other share certain characteristics that cause them to deviate in a similar way from the general

pattern, represented by the Turc-Mezentsev formula. Thus, their residual should be similar and consequently transferable between neighbors.

This paper focuses exclusively on regionalizing the residual between the observed and long-term streamflow computed by the Turc-Mezentsev formula. But since there is no single view on what kind of residual should be used, the efficiency of four different regionalization solutions will be compared (Table 4).

Table 4. Definition of the variables to be regionalized in the Turc-Mezentsev formula. For the first three solutions, we use a fixed value for n (n=2); solution 4 uses parameter n as regionalized variable.

1. Relative error	$R_1 = Q_{obs}/Q_{cal}$
2. Arithmetic difference	$R_2 = Q_{obs} - Q_{cal}$
3. Correcting factor of the aridity index	$Q_{cal} = P \left(1 - \frac{1}{\left(1 + \left(\frac{P}{R_3 E_0} \right)^2 \right)^{1/2}} \right)$ <p>i.e.:</p> $R_3 = \frac{P/E_0}{\sqrt{\left(\frac{P}{P - Q_{obs}} \right)^2 - 1}}$
4. Shape factor (parameter n)	$Q_{cal} = P \left(1 - \frac{1}{\left(1 + \left(\frac{P}{E_0} \right)^{R_4} \right)^{1/R_4}} \right)$ <p>i.e., by successive iterations:</p> $R_4 = - \frac{\ln \left(1 + \left(\frac{P}{E} \right)^{R_4} \right)}{\ln \left(1 - \frac{Q_{obs}}{P} \right)}$

Since an ungaged catchment has several neighbors that can be considered as donors, the value of the regionalized variable is computed as the weighted average of the same variable estimated on the neighboring gaged catchments. This regionalized variable is then applied to the target catchment.

Here, weights consist in the inverse squared distance between catchments:

$$R_{region} = \frac{\sum_{i=1}^p \left(R_{i\ neighbor} \times \frac{1}{d_i^2} \right)}{\sum_{i=1}^p \frac{1}{d_i^2}} \quad \text{Eq. 16}$$

with R_{region} the value of the regionalized variable to transfer to the ungaged catchment, $R_{i\ neighbor}$ the value of the locally estimated variable for neighbor i , and d_i the distance between the ungaged and the neighboring catchment.

Note that for the R4 solution, whatever the value of n , the Turc-Mezentsev formula generates values lying within the limits defined in Figure 28. From a regionalization point of view, this implies that nonconservative catchments cannot be used as donors. This is why distinguishing between conservative and nonconservative catchments is important.

Preliminary tests (not presented here) showed that:

- Considering the first 10 neighbors was sufficient to obtain good performance. Therefore, for the following the p in equation 16 will be set equal to 10.
- Using a distance combining distance to the outlet and distance to the centroid was beneficial to transfer information between catchments of different sizes. The distance considered in equation 16 is thus a "hydrological distance" defined as:

$$d = 0.2 \times d_{outlet} + 0.8 \times d_{centroid} \quad \text{Eq. 17}$$

- Note that several exponents were tested, a squared distance being identified as optimum.

4.3.3 Measuring regionalization performance

To evaluate the performance of the four different regionalization alternatives, the absolute error is computed for each catchment. Catchments are successively considered as the receiver and the distribution of the absolute errors of all basins will be presented in the Results section.

$$\text{Absolute error} = |Q_{cal} - Q_{obs}| \quad \text{Eq. 18}$$

with Q in mm/year.

4.3.4 Assessing regionalization robustness using the hydrometrical desert method

When comparing different regionalization methods, it is important to focus not only on the efficiency of the methods, but also on their robustness, i.e., on how their performance degrades when the hydrometric network density is reduced. Indeed, two methods could be very similar in a data-rich environment, and considerably different under conditions of limited data availability.

To assess the robustness of a spatial proximity-based regionalization method, the hydrometrical desert method is used [Boldetti, 2012], which varies the distance at which the first donor station is located. We start by using all available neighbors, and then the nearest ones located within a 10-, 20-, 30-, ... 100-km radius are progressively excluded. The aim of this test is to show the extent to which the efficiency of the regionalized Turc-Mezentsev formula decreases when regionalization information has to be transferred from greater distances.

4.4. Results

4.4.1 Comparison of the efficiencies for the four regionalization options

First, the efficiencies of the four regionalization options are compared. To implement this test, an average residual is computed based on the 10 nearest donor catchments, as presented in equation 16.

❖ Results obtained over the conservative catchments

Figure 31 shows that regionalization improves the performance of the formula for the four options defined in Table 4. With regionalization, the Turc-Mezentsev formula reaches an absolute error varying between 56.5 and 58.8 mm/year (versus 77.2 mm/year without regionalization). Although the four options yield close results, the one based on regionalizing the n parameter of the Turc-Mezentsev formula (R4) gives the best performance (absolute error, 56.5 mm/year). This improvement demonstrates that enough spatial correlation is present to warrant a successful information transfer from the donors to the receivers.

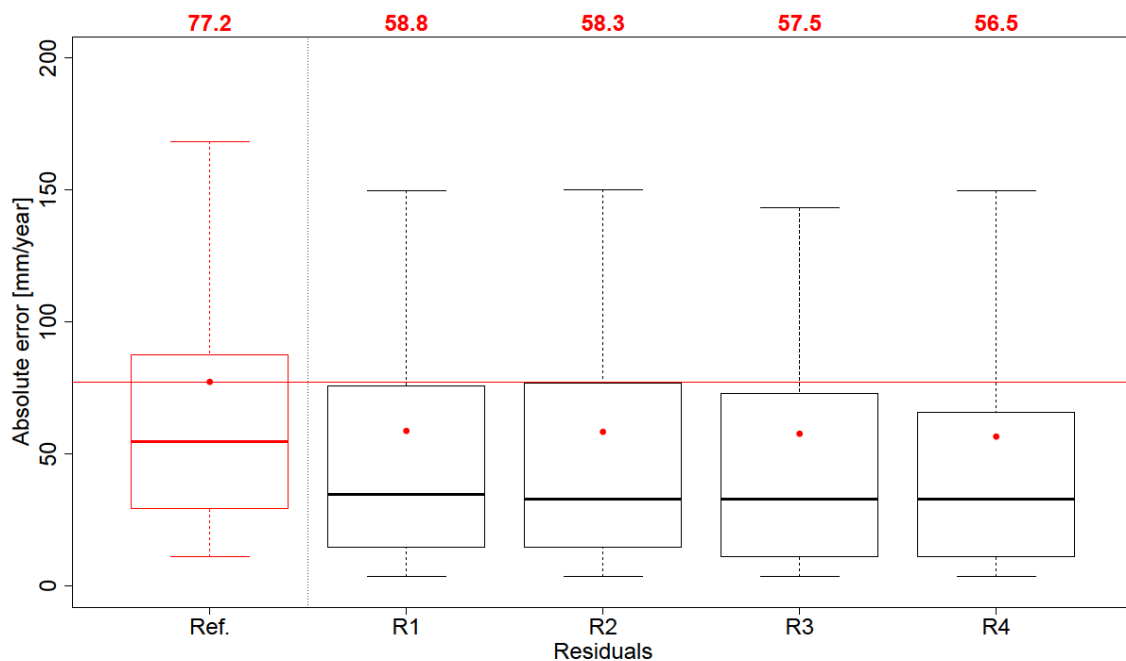


Figure 31. Efficiencies of the four regionalization options for the conservative part of the data set (490 catchments). The reference boxplot on the left corresponds to the nonregionalized formula. The average absolute error value for each method is given on the top of the figure. NB: on this conservative subdataset, all catchments serve as donors and receivers.

❖ Results obtained on the nonconservative part of the data set

It may seem strange to apply a formula that is strictly conservative to catchments that are not conservative. However, one must recognize that in an ungaged situation, it is most often impossible to tell whether a catchment will be conservative or if it will contribute water to a regional aquifer or a neighboring catchment. This is why we also applied the four methods to the nonconservative catchment subset.

Two options are possible: (a) using the 490 conservative catchments as potential donors or (b) using all 609 catchments as potential donors.

- (a) Figure 32-a compares the efficiencies of the four regionalizations for the nonconservative catchments. Here, only the 490 conservative catchments have been used as donors. In this case, absolute error values are much larger, and on average, regionalization deteriorates the results in comparison to when the residual is not applied. However, as for the conservative data set, method R_4 shows the more limited loss of efficiency (absolute error, 258.7 mm/year), with results approaching the reference value (absolute error, 248.9 mm/year).
- (b) In Figure 32-b, all neighbors serve as donors, whether or not they are conservative. This approach advantages the first three methods, which can benefit from 609 donors, while method R_4 (which can only compute parameter n over conservative catchments) must rely on 490 catchments only: its results remain unchanged in comparison with Figure 32-a.

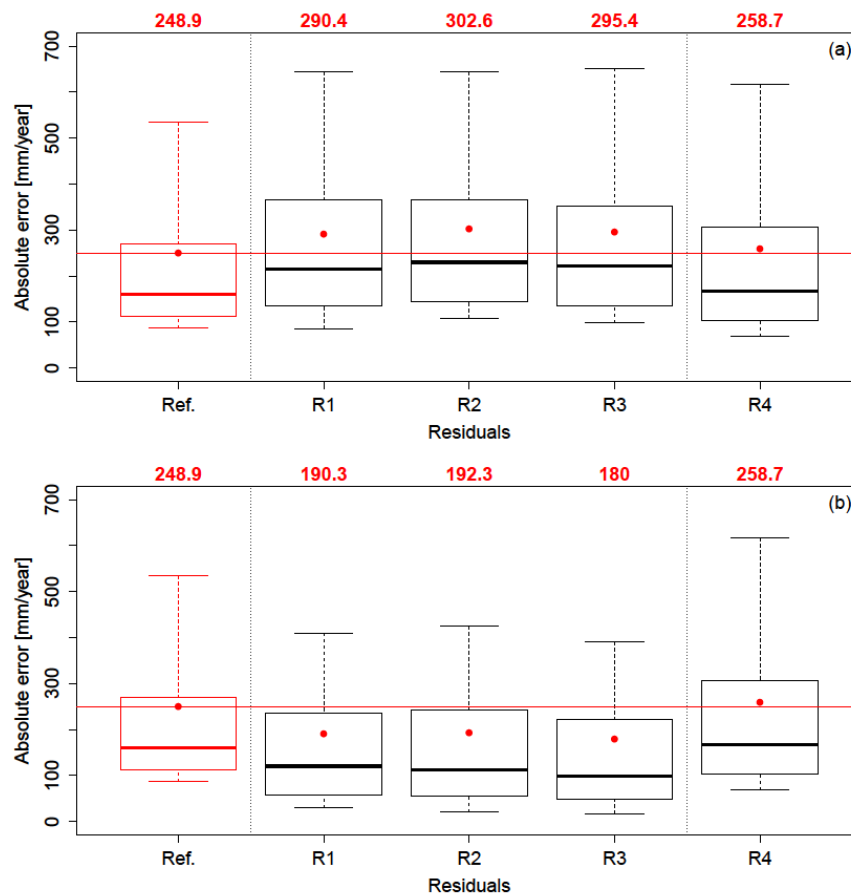


Figure 32. Regionalization of the Turc-Mezentsev formula (see Table 4 for a description of the four options) tested on a subset of nonconservative catchments. (a) Donors are chosen among the conservative catchments only; (b) donors are chosen from the whole data set (conservative and nonconservative catchments, whichever is closest). The results of method R_4 remain unchanged in (a) and (b), since its parameter can only be computed over conservative catchments.

❖ **Results obtained on the entire data set (i.e., conservative and nonconservative data sets combined)**

As mentioned above, in an ungaged situation, it is often impossible to tell whether or not a catchment will be conservative. To obtain general results, the four methods were applied to the entire data set and allowed all neighbors to serve as donors (except for method R4, which can only compute a residual over conservative catchments).

Figure 33 shows that when no distinction between conservative and nonconservative catchments is made, regionalization of the Turc-Mezentsev formula improves absolute error values for the four methods. Absolute error values are quite similar between methods (variation between 88.5 and 96 mm/year) even if method R3 seems to give the best results.

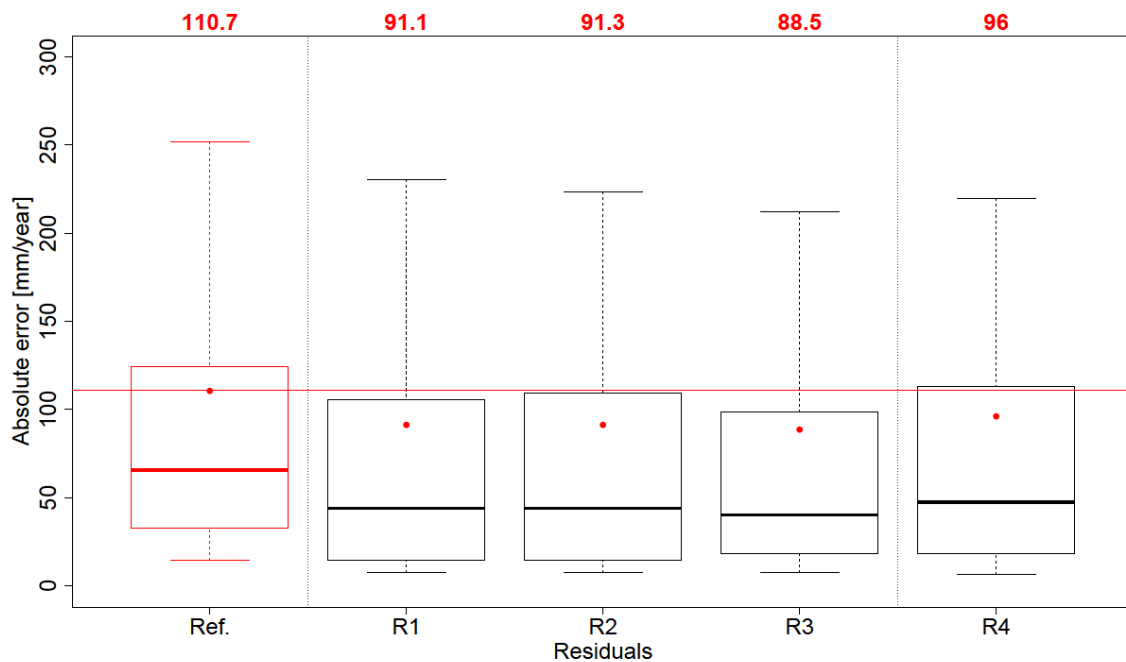


Figure 33. Efficiencies of the four options of residual transfer for the whole data set. The reference boxplot on the left corresponds to the nonregionalized formula. The absolute error value for each method is given on the top of the figure. In this case, donors are chosen from the whole data set (conservative and nonconservative catchments), except for method R4, since its residual can only be computed over conservative catchments.

Figure 34 presents the geographical distribution of the best option of residual transfer for each catchment. Options R1 and R4 give the best results for 37% and 34% of catchments, respectively. Option R3 dominates in the nonconservative catchments, which correspond mostly to the mountainous east margin of the country. The map also shows that there is no spatial pattern of the regionalization options, i.e., there is no region characterized by a particular option.

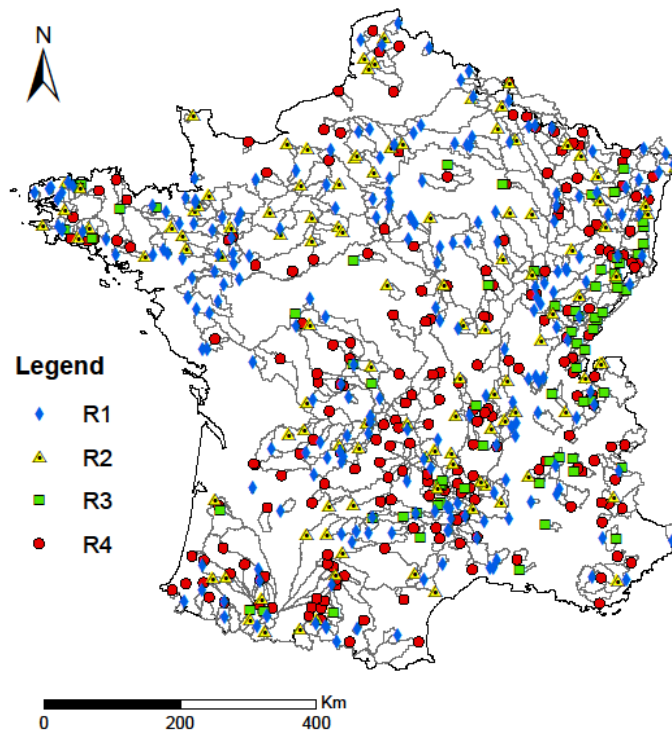


Figure 34. Geographical distribution of the best option of residual transfer for each catchment.

4.4.2 Robustness assessment

The hydrometrical desert method described in section 4.3.4 was tested on the conservative part of the data set for the four residuals, with limit distances of the closest donor of 0 (no distance limit), 10, 20, 30, ... 100 km.

Figure 35 presents the robustness assessment for the conservative part of the data set. The first result was expected: when the donor catchments used to compute the residual are located farther and farther from the ungauged catchment, the results of the regionalized formula deteriorate. However, for all the residuals, even if the donor catchments are more than 80 km from the target catchment, the regionalization of the residual is still beneficial (the results are better than if only the Turc-Mezentsev formula is used without regionalization). The best configuration is when the nearest neighbors are available.

Furthermore, method R4 (i.e., the method that regionalizes n , the shape parameter of the Turc-Mezentsev formula) gives the best results and appears to be the most robust: even when no neighbor closer than 100 km is allowed, it still manages to provide a slight improvement over the nonregionalized alternative.

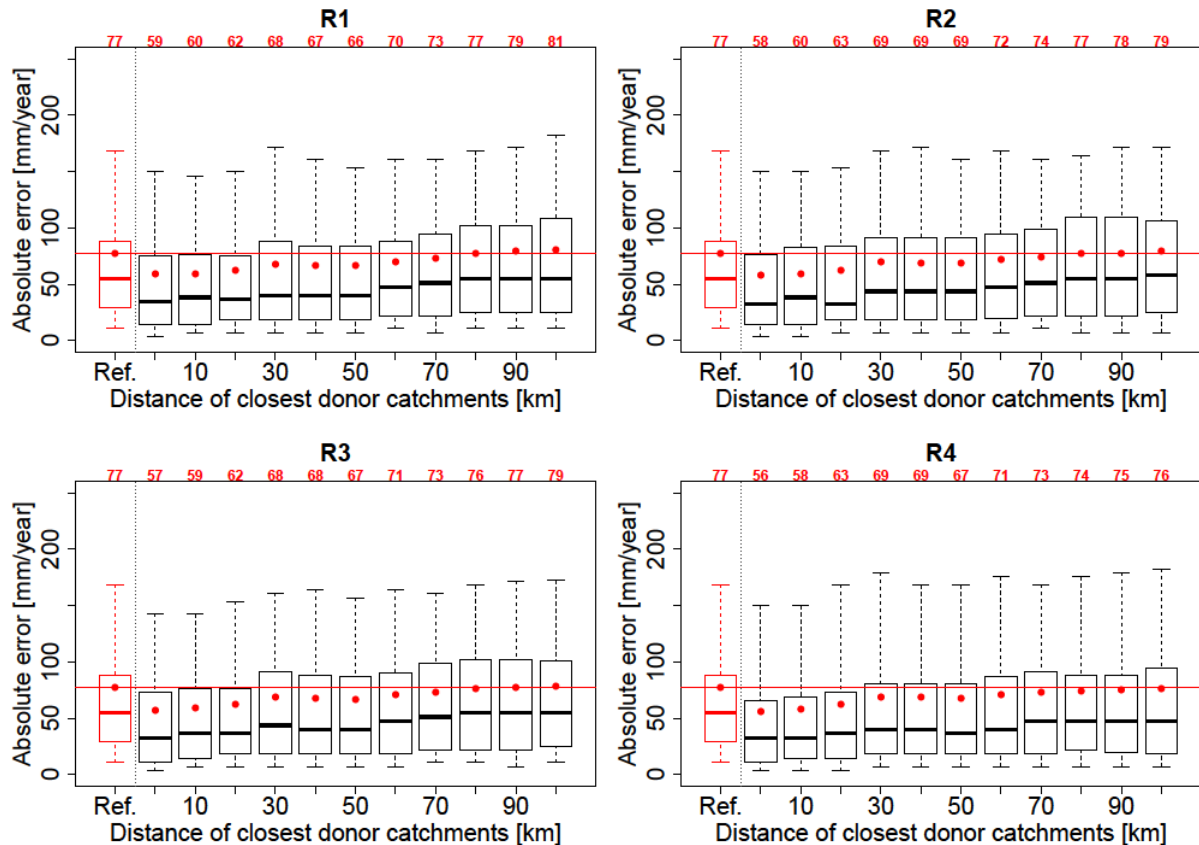


Figure 35. Impact of the distance of donor catchments used in regionalization on regionalization efficiency. The reference boxplot on the left corresponds to the nonregionalized formula. On the top of the figure, the average absolute error value for each method is given.

4.4.3 Discussion

These results underscore the difference existing between the conservative and nonconservative catchments. We believe the methods must be tested on these difficult cases as well, even if they yield more complex and less optimistic results [Andréassian *et al.*, 2010a]. Of course, the nonconservative catchments that leak or gain water do not respect the water balance, and the Turc-Mezentsev formula is not expected to represent these effects: this explains method R4's shortcomings (based on the regionalization of the shape factor n).

For the catchments that leak or gain water, the addition of these catchments to the donor data set greatly improves the results of the R1, R2 and R3 methods, which can then provide better results than the R4 method, for which only conservative catchments can be used: R1, R2, and R3 use the regionalized variable/residual to explore the forbidden areas of Figure 28.

It is not easy to conclude on which solution is preferable: on a truly ungauged catchment, it is difficult to tell in advance whether or not the catchment will be conservative:

- The most satisfying method (from a physical and water balance point of view) is method R4, where the shape parameter of the Turc-Mezentsev formula is regionalized;
- In mountainous regions, where the risk of having nonconservative catchments is high, method R3, which regionalizes a correction factor to the aridity index, is preferable;

- When the category the catchment belongs to is unknown, using method R3 is safer.

4.4.4 Does catchment size have an impact on the results?

This section analyzes the sensitivity of the results to the catchment area. Method R4, which regionalizes the shape parameter of the Turc-Mezentsev formula, is used. Figure 36 represents the performance of the regionalized Turc-Mezentsev formula for the two data sets as a function of catchment area. It shows that performance is much better for the large basins.

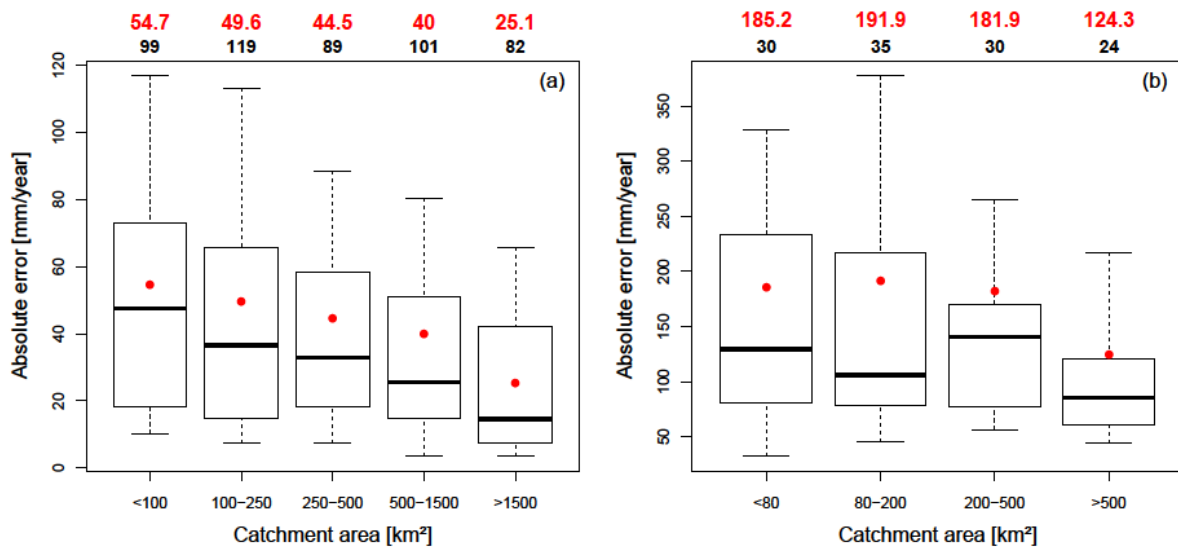


Figure 36. Influence of catchment area on the performance of the Turc-Mezentsev formula for (a) the conservative data set and (b) the catchments that leak or gain water. The mean absolute error value for each catchment size is given in red on the top of the figure. The number of catchments for each group is given in black under the mean absolute error.

4.5. Conclusion and perspectives

4.5.1 Conclusion

This paper aimed to compare four options for regionalizing the Turc-Mezentsev formula to improve its predictive efficiency: (i) the relative error, (ii) the arithmetic difference, (iii) the correcting factor of the aridity index and (iv) the shape factor (parameter n).

The results show that differences exist between the conservative catchments (i.e., those that fall strictly within the water and energy limits) and nonconservative catchments (i.e., those that either leak or gain water): for the former, the best results are obtained with the regionalization of the shape factor, while for the latter, the regionalization of the correcting factor of the aridity index is preferable. The regionalization of the shape factor does not work for these catchments because this method confines the residual within the catchment yield limits of Figure 28, while the correcting factor of the aridity index allows exploring the forbidden areas of Figure 28 (i.e., where the nonconservative catchments are located).

For a truly unengaged catchment, it is almost impossible to know beforehand whether the catchment is conservative or nonconservative. When the results of the two subparts of the data set are mixed, the

regionalization of the correcting factor provides the best result. This is supported by previous work reported by *Oudin et al.* [2008a].

4.5.2 Limits

The regionalization approach tested in this paper used a spatial proximity-based transfer of hydrological residuals. Obviously, other types of regionalization would be possible: many would agree that basing regionalization on physiographic characteristics would be more satisfying from a physical point of view. We limited our investigation to geographical regionalization because of the difficulties encountered in previous attempts (see e.g. *Oudin et al.* [2008b] and *Oudin et al.* [2010]). One of the reasons for these difficulties is that the most important physical determinants of model error (soil properties, geological properties) are not directly observable. Since spatial proximity provides a good 'proxy' for many of these unobservable physical properties, we believe that the results obtained here will still be of interest for other regionalization approaches.

4.5.3 Perspectives

This paper explored mathematical solutions for introducing regional knowledge into the Turc-Mezentsev formula. With an answer as to which mathematical solution is preferable, we could attempt to go beyond geographical regionalization into physical regionalization. Many options are possible. Interestingly, already in the conclusion of his 1954 paper, Lucien Turc had identified perspectives that deserve further research. He stressed that "the most urgent improvement should be to introduce the repartition of precipitation and the variation of temperatures in the course of the year." He also mentioned that the temperature-based formulation he used for E could be improved by accounting for incoming radiation and air moisture as well, and that in this case it would become necessary to readjust the exponent n .

Globally, a relation exists between hydrologic seasonality and mean annual water balance. Many authors have emphasized this [Milly, 1994; Dooge et al., 1999; Hickel and Zhang, 2006; Potter et al., 2005; Zhang et al., 2008; Yokoo et al., 2008; Williams et al., 2012]. In the future, it would be useful to identify a relation between the formula's shape parameter and a climatic index of hydrologic seasonality (see the attempt of *Mouelhi* [2003] with an index of rainfall seasonality and the analytical derivation of the Schreiber curve by *Gerrits et al.* [2009] on the basis of simple statistical assumptions on the time distribution of rainfall and potential evaporation). However, this should necessarily encompass a highly varied worldwide data set.

4.6. Acknowledgements

The authors wish to thank Dr. Irina Ginzburg for helping with the Russian-language papers.

We would like to thank the different institutions that allowed us to build the large catchment sets on which this paper is based: Météo-France for meteorological data and SCHAPI for hydrometrical data. ONEMA provided financial support for the first author.

4.7. Annex 1 – partial differential equations for streamflow as a function of P and E_0

The Turc-Mezentsev formula can be written as a function of streamflow (as in equation 15):

$$\frac{Q}{P} = 1 - \frac{1}{\left(1 + \left(\frac{P}{E_0}\right)^n\right)^{\frac{1}{n}}} \quad (\text{A1})$$

To compute its partial differentials, it is practical to rewrite it as:

$$Q = P - (P^{-n} + E_0^{-n})^{-\frac{1}{n}} \quad (\text{A2})$$

The partial differential equations for Q can then be directly deduced:

$$\frac{\partial Q}{\partial P} = 1 - \frac{1}{\left(1 + \left(\frac{P}{E_0}\right)^n\right)^{\frac{n+1}{n}}} \quad (\text{A3})$$

(which is the Bagrov's hypothesis as reparameterized by Mezentsev)

$$\frac{\partial Q}{\partial E_0} = - \frac{1}{\left(1 + \left(\frac{E_0}{P}\right)^n\right)^{\frac{n+1}{n}}} \quad (\text{A4})$$

Chapitre 5.

Analyse de la sensibilité des calculs hydrologiques à la densité spatiale des réseaux hydrométriques

5.1. Introduction

Ce chapitre a pour objectif d'évaluer l'impact de la densité spatiale du réseau hydrométrique sur divers calculs hydrologiques. Cette étude de sensibilité est effectuée sur l'échantillon de 609 bassins versants que nous avons présenté dans le chapitre 3. Nous avons utilisé dans ce chapitre le modèle GR4J couplé au module CemaNeige, la méthode de régionalisation basée sur la proximité spatiale, ainsi que la méthode du désert hydrométrique, discutés dans le chapitre 3, afin d'évaluer l'impact de la densité spatiale des réseaux hydrométriques sur ces calculs hydrologiques.

Dans ce chapitre, nous présenterons tout d'abord les différents calculs envisagés au sein de cette thèse. Nous nous sommes intéressés en particulier à la simulation de chroniques de débit journalier, à l'estimation du module, des débits de crues décennale et centennale et à certaines caractéristiques de débit d'étiage.

Enfin, les résultats concernant l'analyse de sensibilité des calculs hydrologiques à la densité spatiale des réseaux hydrométriques seront présentés.

5.2. Les calculs hydrologiques considérés

5.2.1 Simulation d'une chronique de débit journalier en mode non-jaugé avec le modèle GR4J

Pour la simulation de la chronique de débit journalier sur le bassin non jaugé, le modèle hydrologique utilisé est le modèle global pluie-débit GR4J, présenté dans le chapitre 3.

Le calage des paramètres des bassins voisins a été effectué sur la période 1997-2006, puis utilisé sur cette même période pour calculer une série moyenne de débit sur le bassin versant non jaugé à l'aide des données de pluie et d'ETP observées sur ce bassin (voir la méthode de régionalisation présentée dans le chapitre 3).

5.2.2 Calcul du module en mode non-jaugé

Pour le calcul du débit moyen pluri-annuel, appelé également module, de longues chroniques de débit sont nécessaires (au moins 18 années d'après *Sauquet et Catalogne* [2010]).

Ainsi, les paramètres des bassins voisins calés sur la période 1997-2006, ont été utilisés pour simuler une série moyenne de débit sur le bassin versant non jaugé sur la période 1977-2006 (voir la méthode de régionalisation présentée dans le chapitre 3). Le débit moyen à long terme est alors obtenu en moyennant les débits journaliers de cette chronique simulée sur 30 ans.

5.2.3 Calcul des débits de crues décennale et centennale en mode non-jaugé

La détermination des quantiles d'aléas hydrologiques (crue et étiage) nécessite également des chroniques de débit de longue durée (au moins 26 années d'après *Catalogne et Sauquet* [2011]). Des chroniques de trente ans simulées par le modèle GR4J ont également été utilisées pour les calculs de débits de crues décennale et centennale.

Afin de calculer les valeurs de débits de crue décennale et de crue centennale, la loi de Gumbel [Lang et Lavabre, 2007] a été utilisée. Les paramètres de cette loi ont été estimés avec la méthode des moments.

Ainsi, après avoir échantillonné les débits maximaux journaliers de chaque année, nous utilisons la loi de Gumbel pour déterminer les quantiles 0.9 (crue décennale) et 0.99 (crue centennale).

5.2.4 Calcul des caractéristiques d'étiage en mode non-jaugé

Dans cette étude, trois caractéristiques d'étiage ont été envisagées : le QMNA5 (débit minimal mensuel de période de retour 5 ans) et les quantiles 0.05 et 0.10 de la distribution des débits journaliers.

La méthode de référence utilisée pour estimer la valeur du QMNA5 est un ajustement statistique log-normal sur les valeurs minimales annuelles, en utilisant des simulations de trente années de GR4J comme précédemment (il ne s'agit pas cependant des mêmes simulations, voir ci-dessous).

5.3. Critère de calage des paramètres et d'évaluation du modèle

Selon l'objectif hydrologique, différentes fonctions objectif ont été utilisées pour le calage des paramètres du modèle GR4J :

- Débits journaliers et module : le critère de *Nash et Sutcliffe* [1970] calculé sur la racine carrée des débits qui donne autant de poids aux forts et faibles débits [Oudin et al., 2006b];

$$NS = 1 - \frac{\sum_{i=1}^n \left(\sqrt{Q_{sim}^i} - \sqrt{Q_{obs}^i} \right)^2}{\sum_{i=1}^n \left(\sqrt{Q_{obs}^i} - \sqrt{\overline{Q_{obs}}} \right)^2} \quad \text{Eq. 19}$$

avec i le pas de temps utilisé pour la simulation (ici journalier), Q_{obs}^i le débit observé au pas de temps i , Q_{sim}^i le débit simulé au pas de temps i et $\overline{Q_{obs}}$ le débit observé moyen.

- Débits d'étiage (QMNA5, quantiles de débit 0.05 et 0.10) : le critère de *Nash et Sutcliffe* [1970] calculé sur le logarithme des débits qui privilégie la qualité de simulation des bas débits [Oudin et al., 2006b];

$$NS = 1 - \frac{\sum_{i=1}^n \left(\ln(Q_{sim}^i + \varepsilon) - \ln(Q_{obs}^i + \varepsilon) \right)^2}{\sum_{i=1}^n \left(\ln(Q_{obs}^i + \varepsilon) - \ln(\overline{Q_{obs}}) \right)^2} \quad \text{Eq. 20}$$

avec i le pas de temps utilisé pour la simulation (ici journalier), Q_{obs}^i le débit observé au pas de temps i , Q_{sim}^i le débit simulé au pas de temps i , $\overline{Q_{obs}}$ le débit observé moyen et ε constante très faible ($\overline{Q_{obs}}/40$) ajoutée aux débits pour éviter les difficultés en cas de débit nul (impossible de calculer $\ln(0)$).

- Débits de crues décennale et centennale : critère KGE développé par *Gupta et al.* [2009]. Ces auteurs ont montré que l'utilisation du critère KGE à la place du critère de *Nash et Sutcliffe* [1970] comme fonction objectif permet de concilier différentes qualités attendues de la part

du modèle hydrologique (biais, variance et corrélation). Ils soulignent également que lorsque le critère de *Nash et Sutcliffe* [1970] est utilisé pour le calage des paramètres, les pics de crue ont tendance à être davantage sous-estimés qu'avec le critère KGE. C'est pourquoi nous choisissons d'utiliser le critère KGE comme fonction objectif lorsque l'on s'intéresse à l'estimation des débits de crue.

$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad \text{Eq. 21}$$

avec r le coefficient de corrélation linéaire entre les débits observés et simulés, α le ratio entre l'écart-type simulé et l'écart-type observé et β le ratio entre le débit moyen simulé et le débit moyen observé (soit le biais).

La Figure 37 présente les distributions de l'erreur relative obtenues pour le calcul de différentes gammes de quantiles de débit calculés à partir de chroniques de débit de 30 ans simulées avec un calage des paramètres réalisé en utilisant chacune des trois fonctions objectifs présentées ci-dessus (quantiles de non dépassement de débits journaliers de 5% (basses eaux) à 95% (hautes eaux)).

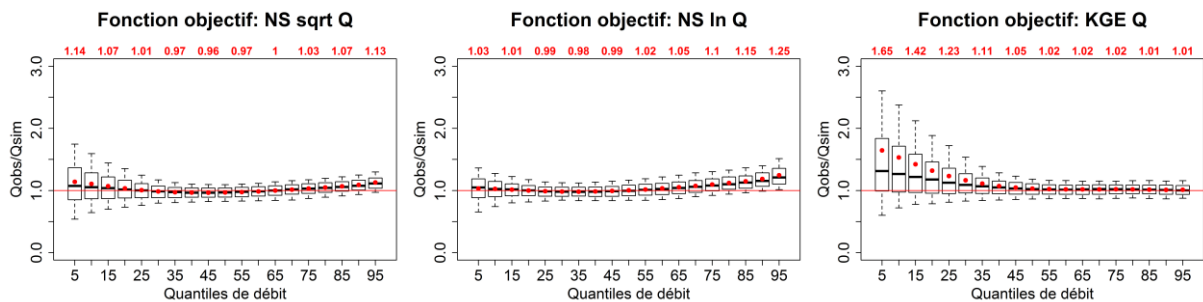


Figure 37. Distribution de l'erreur relative des différents quantiles de débit calculés à partir de chroniques de débit de 30 ans simulées avec un calage des paramètres effectué avec différentes fonctions objectifs.

Cette figure montre que lorsque le modèle est calé avec le critère de *Nash et Sutcliffe* [1970] calculé sur le logarithme des débits (Eq. 20), on obtient de bons résultats pour les bas et moyens débits alors que les hauts débits ont tendance à être biaisés. Dans le cas où le modèle est calé avec le critère de *Nash et Sutcliffe* [1970] calculé sur la racine carrée des débits (Eq. 19), les hauts débits sont moins biaisés, alors que les bas et les moyens débits restent peu biaisés en moyenne. Enfin, la dernière option a été de caler le modèle avec le critère KGE calculé sur les débits proposé par *Gupta et al.* [2009](Eq. 21). Dans ce cas, les faibles débits sont moins bien modélisés et on observe un fort biais, alors que les moyens et forts débits ne sont pas biaisés.

Pour évaluer les performances du modèle GR4J pour la simulation de la chronique de débit journalier, le critère de *Nash et Sutcliffe* [1970] calculé sur la racine carrée des débits a été utilisé. Afin d'évaluer les performances des autres estimations de débit (moyen à long-terme, de crue décennale ou centennale et d'étiage), l'erreur absolue (Eq. 22) ainsi que le biais (Eq. 23) entre débit régionalisé (Q_{reg}) et débit observé (Q_{obs}) ont été calculés pour chaque bassin.

$$\text{Erreur absolue} = |Q_{reg} - Q_{obs}| \quad \text{Eq. 22}$$

$$\text{Biais} = Q_{reg}/Q_{obs} \quad \text{Eq. 23}$$

Nous présenterons les résultats sous forme de distributions relatives à tous les bassins étudiés :

- distribution des valeurs du critère de Nash-Sutcliffe pour la simulation de la chronique journalière ;
- distribution de l'erreur absolue et du biais pour les autres calculs de débit.

5.4. Evaluation préalable suivant les gammes de débit

Nous avons dans un premier temps analysé les performances obtenues par le modèle calé et le modèle régionalisé pour différentes gammes de débits (quantiles de non dépassement de débits journaliers de 5% (basses eaux) à 95% (hautes eaux)).

La Figure 38 présente l'erreur relative obtenue pour le modèle calé et le modèle régionalisé. On observe d'une part que l'amplitude des erreurs relatives augmente lorsque l'on passe d'un modèle calé à un modèle régionalisé, ce qui était attendu. D'autre part, aussi bien pour le cas régionalisé que pour le cas jaugé, on observe que les débits d'étiage (quantiles de débit 5% à 15%) et les débits de hautes eaux (quantiles de débit 70% à 95%) sont sous-estimés. Entre ces deux extrêmes, les débits intermédiaires sont quant à eux un peu surestimés.

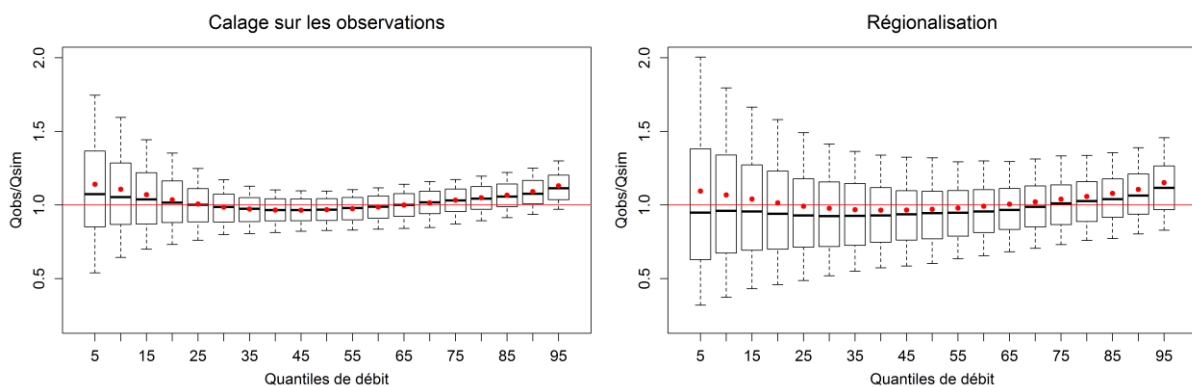


Figure 38. Performances des reconstitutions de quantiles sélectionnés en utilisant un calage de GR4J sur les données observées (à gauche) et en utilisant des chroniques régionalisées avec GR4J (à droite). Note : calage réalisé en NS sur la racine carrée des débits.

5.5. Evaluation de la robustesse de la régionalisation

Nous cherchons maintenant à évaluer la robustesse de la méthode de régionalisation pour chacun des calculs présentés ci-dessus, en appliquant la méthode du désert hydrométrique présenté dans le chapitre 3. Les Figure 39 à Figure 45 présentent la distribution des performances obtenues pour l'ensemble de l'échantillon de bassins versants en fonction de la distance des plus proches voisins considérés pour les différents calculs hydrologiques envisagés. La distribution « Cal. » présente le cas où les débits simulés ont été obtenus en validation sur la même période que pour la régionalisation mais avec les paramètres calés sur le bassin cible.

Les Figure 39 à Figure 45 montrent dans un premier temps, que pour tous les calculs de débits envisagés, l'utilisation de la régionalisation pour le calcul des débits du bassin non jaugé apporte de moins bons résultats que dans le cas de référence. Ce résultat était attendu puisque le cas de

référence correspond au cas « idéal » (lorsque les données de débit sur le bassin cible sont disponibles pour caler les paramètres du modèle).

On observe également que pour chacun des débits cibles, les performances en régionalisation se dégradent lorsque la distance au plus proche voisin augmente mais atteignent un palier à partir d'une distance de 150 km. Ainsi, lorsque les bassins voisins sont éloignés de plus de 150 km, le rôle de la proximité géographique des bassins voisins semble s'affaiblir. A partir de cette distance, la méthode engendre de trop fortes erreurs, la notion de proximité géographique comme indicateur de similarité physique perdant alors de son sens.

D'autre part, cette baisse d'efficacité en fonction de la distance des plus proches voisins, non négligeable, est variable entre les différents calculs (Tableau 6). Ce sont les bas débits qui ont une plus forte augmentation de l'erreur lorsque l'on passe d'un modèle calé à un modèle régionalisé ou d'un modèle régionalisé 0 km à un modèle régionalisé 150 km. Ceci s'explique par le fait que pour les faibles valeurs de débit, une moindre différence aura un impact relatif beaucoup plus important que pour des débits moyens ou des hauts débits. Nous pouvons également constater que la chute de performances lorsque la distance du plus proche voisin augmente de 0 à 150 km, est inférieure aux erreurs dues à la méthode de régionalisation elle-même pour tous les calculs.

Par ailleurs, les pertes de performances pour les calculs du quantile 0.05 et du QMNA5 sont quasi-identiques lorsque la distance des plus proches voisins augmente. Le comportement de ces deux calculs est très similaire.

Enfin, les graphiques des Figure 39 à Figure 45 pourraient être utilisés comme graphiques de référence lorsque l'on veut connaître les performances auxquelles on pourrait s'attendre dans le cas d'un bassin versant non jaugé. En effet, il est facile d'obtenir la distance séparant un bassin non jaugé de son plus proche voisin. A partir de cette distance et en utilisant ces graphiques, il devient facile d'estimer les performances moyennes attendues pour chacun de ces calculs (et leur incertitude).

Tableau 6. Erreur relative médiane des performances entre le modèle calé et les modèles régionalisés avec des distances limites des plus proches voisins de 0 km et 150 km et erreur relative des performances entre les modèles régionalisés 0 km et 150 km pour tous les calculs hydrologiques envisagés.

Calculs	Modèle calé et Mod. régionalisé 0km	Modèle calé et Modèle régionalisé 150 km	Modèles Régionalisés 0km et 150km
GR4J	-10%	-23%	-14%
Module	71%	157%	50%
Débit de crue décennale	67%	142%	45%
Débit de crue centennale	59%	123%	41%
Quantile de débit 0.05	150%	350%	80%
Quantile de débit 0.10	200%	400%	67%
QMNA5	150%	350%	80%

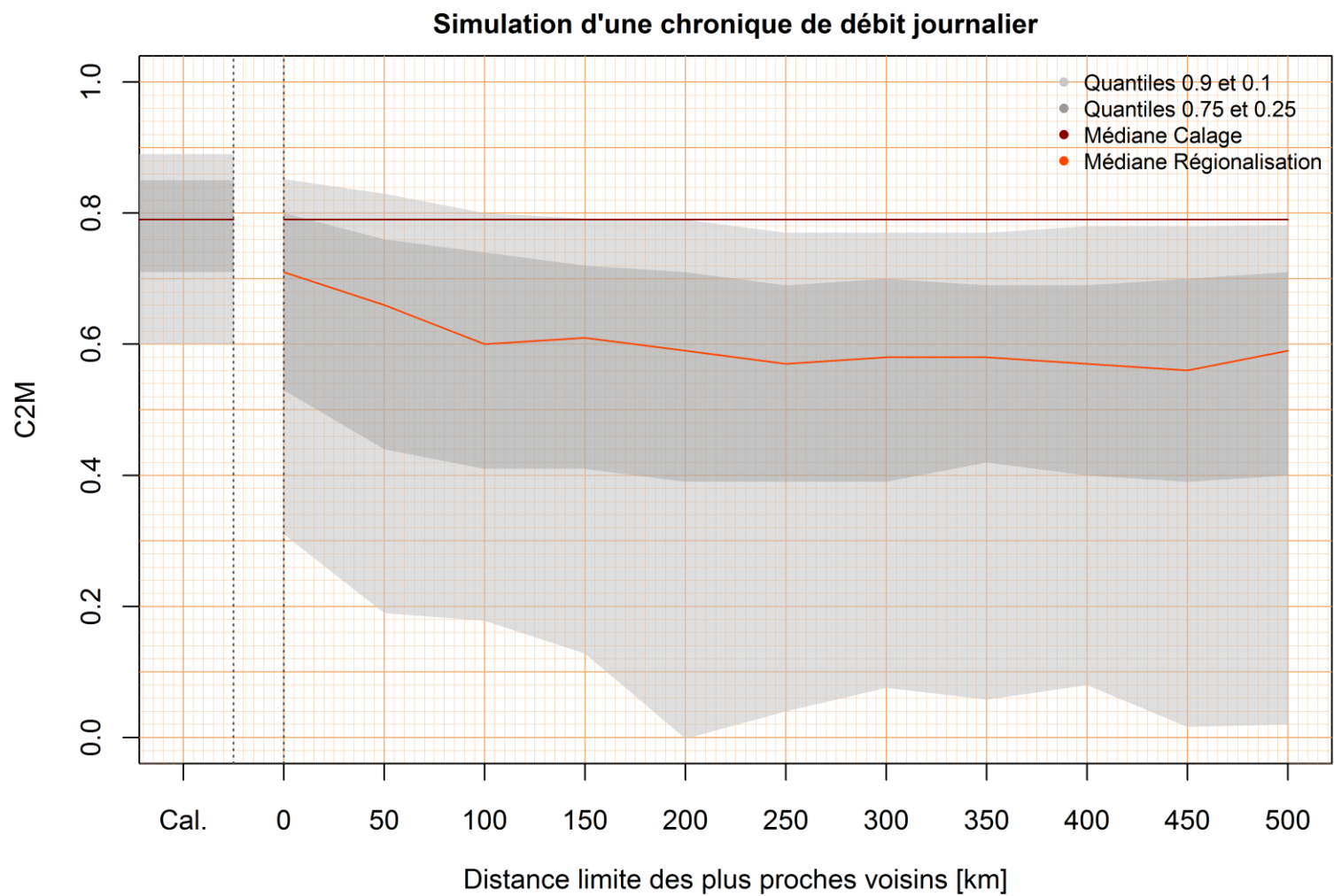


Figure 39. Influence de la distance à laquelle se situe le bassin voisin le plus proche sur l'efficacité de la régionalisation avec le modèle GR4J. « Cal. » présente le cas où les paramètres du modèle GR4J et les débits simulés ont été obtenus par calage à l'aide des données de débit observées sur le bassin cible. La distance 0 km correspond au cas où aucune limite de distance des plus proches voisins n'est appliquée.

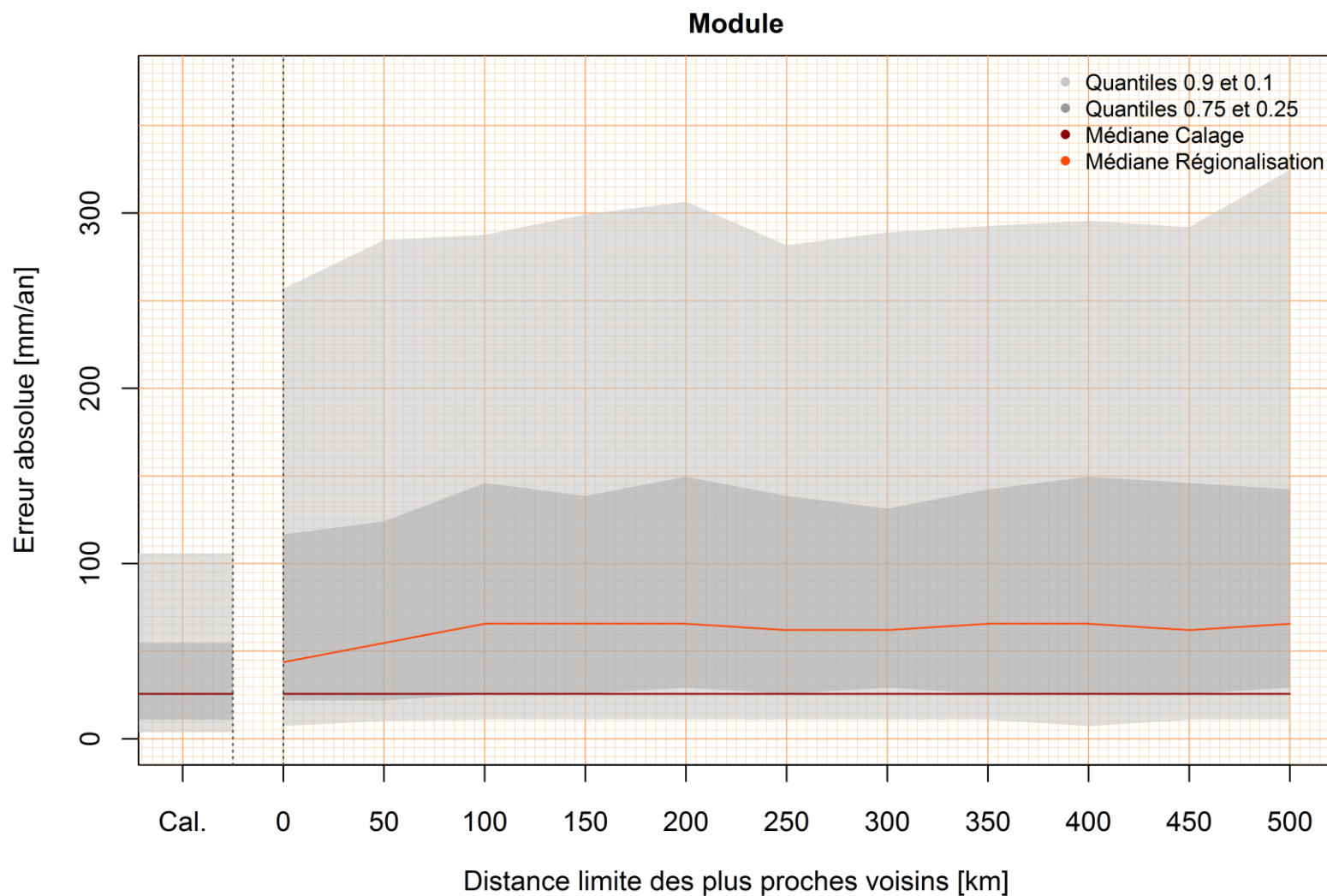


Figure 40. Influence de la distance à laquelle se situe le bassin voisin le plus proche sur l'efficacité de la régionalisation avec le modèle GR4J pour le calcul du module. « Cal. » présente le cas où les débits simulés ont été obtenus en validation sur la même période que pour la régionalisation mais avec les paramètres calés sur le bassin cible. La distance 0 km correspond au cas où aucune limite de distance des plus proches voisins n'est appliquée.

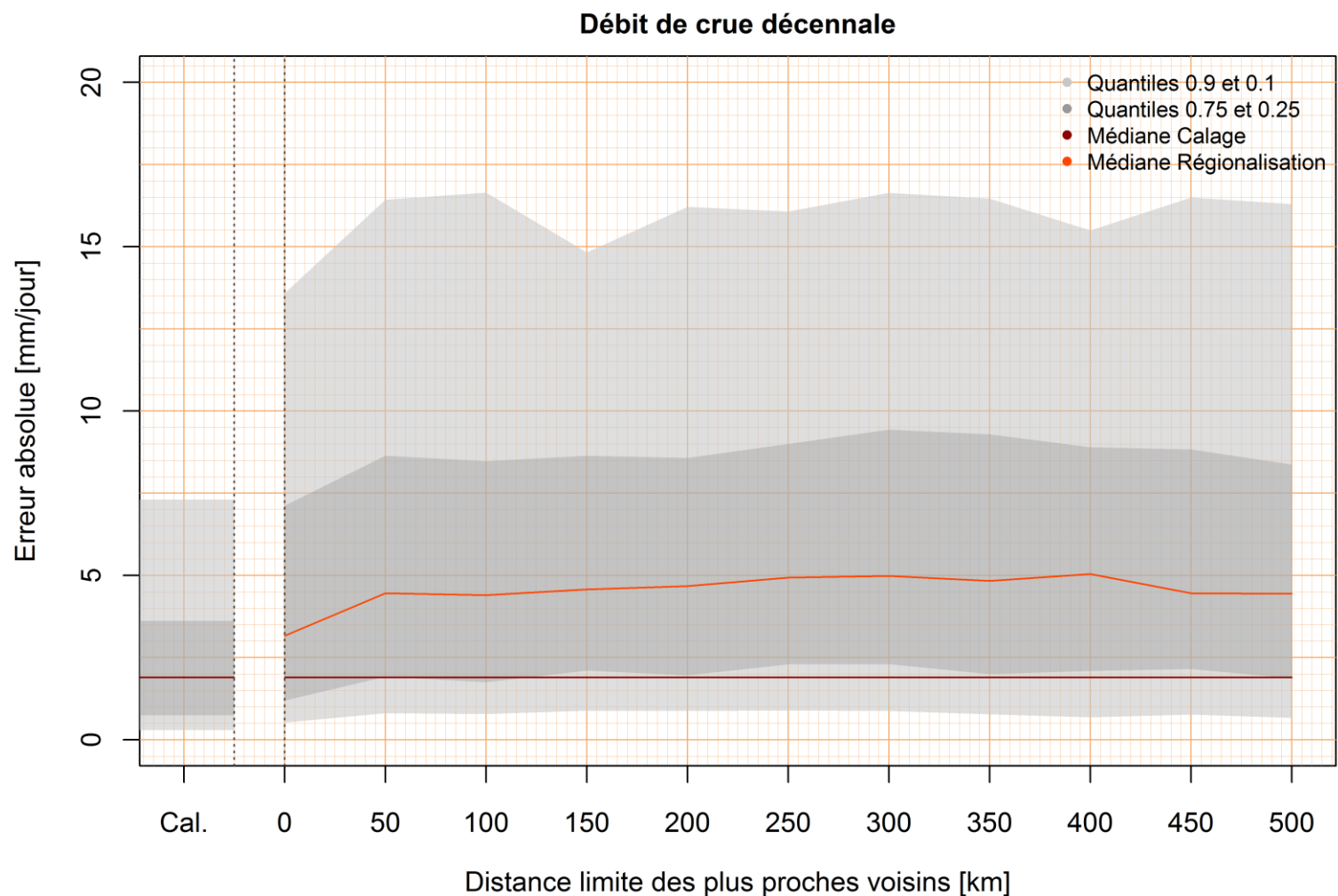


Figure 41. Influence de la distance à laquelle se situe le bassin voisin le plus proche sur l'efficacité de la régionalisation avec le modèle GR4J pour le calcul de la crue décennale. « Cal. » présente le cas où les débits simulés ont été obtenus en validation sur la même période que pour la régionalisation mais avec les paramètres calés sur le bassin cible. La distance 0 km correspond au cas où aucune limite de distance des plus proches voisins n'est appliquée.

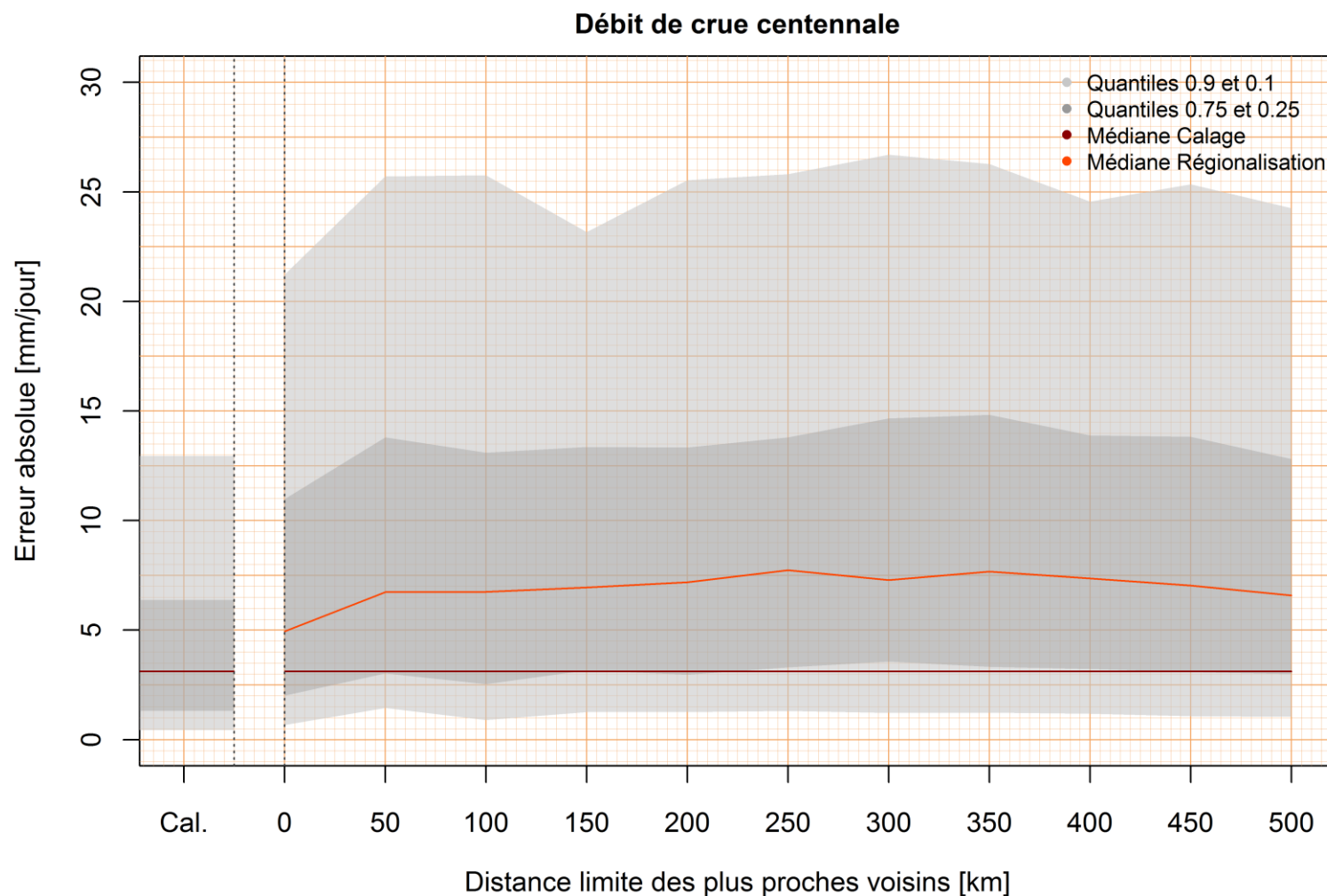


Figure 42. Influence de la distance à laquelle se situe le bassin voisin le plus proche sur l'efficacité de la régionalisation avec le modèle GR4J pour le calcul de la crue centennale. « Cal. » présente le cas où les débits simulés ont été obtenus en validation sur la même période que pour la régionalisation mais avec les paramètres calés sur le bassin cible. La distance 0 km correspond au cas où aucune limite de distance des plus proches voisins n'est appliquée.

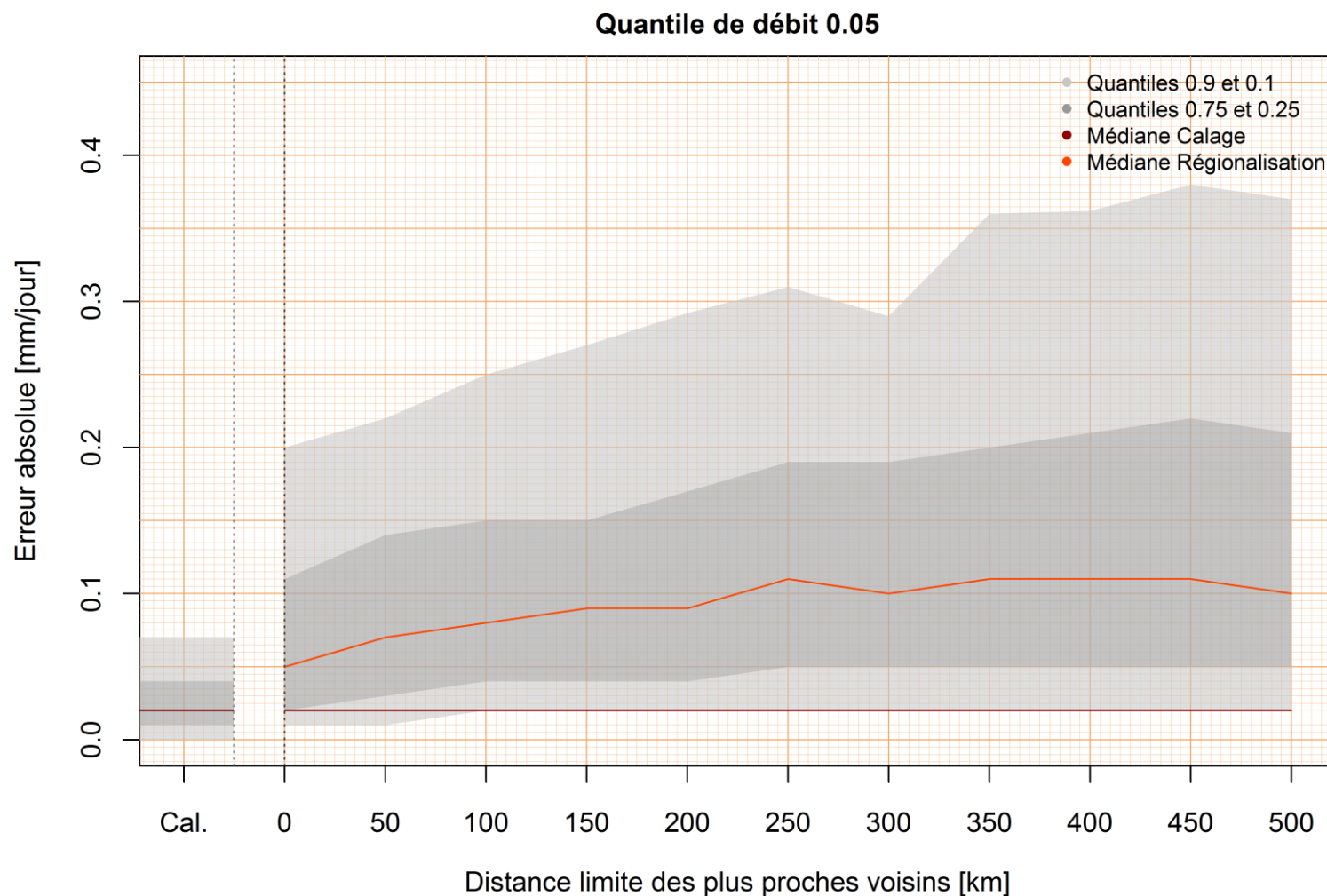


Figure 43. Influence de la distance à laquelle se situe le bassin voisin le plus proche sur l'efficacité de la régionalisation avec le modèle GR4J pour le calcul du quantile de débit 0.05. « Cal. » présente le cas où les débits simulés ont été obtenus en validation sur la même période que pour la régionalisation mais avec les paramètres calés sur le bassin cible. La distance 0 km correspond au cas où aucune limite de distance des plus proches voisins n'est appliquée.

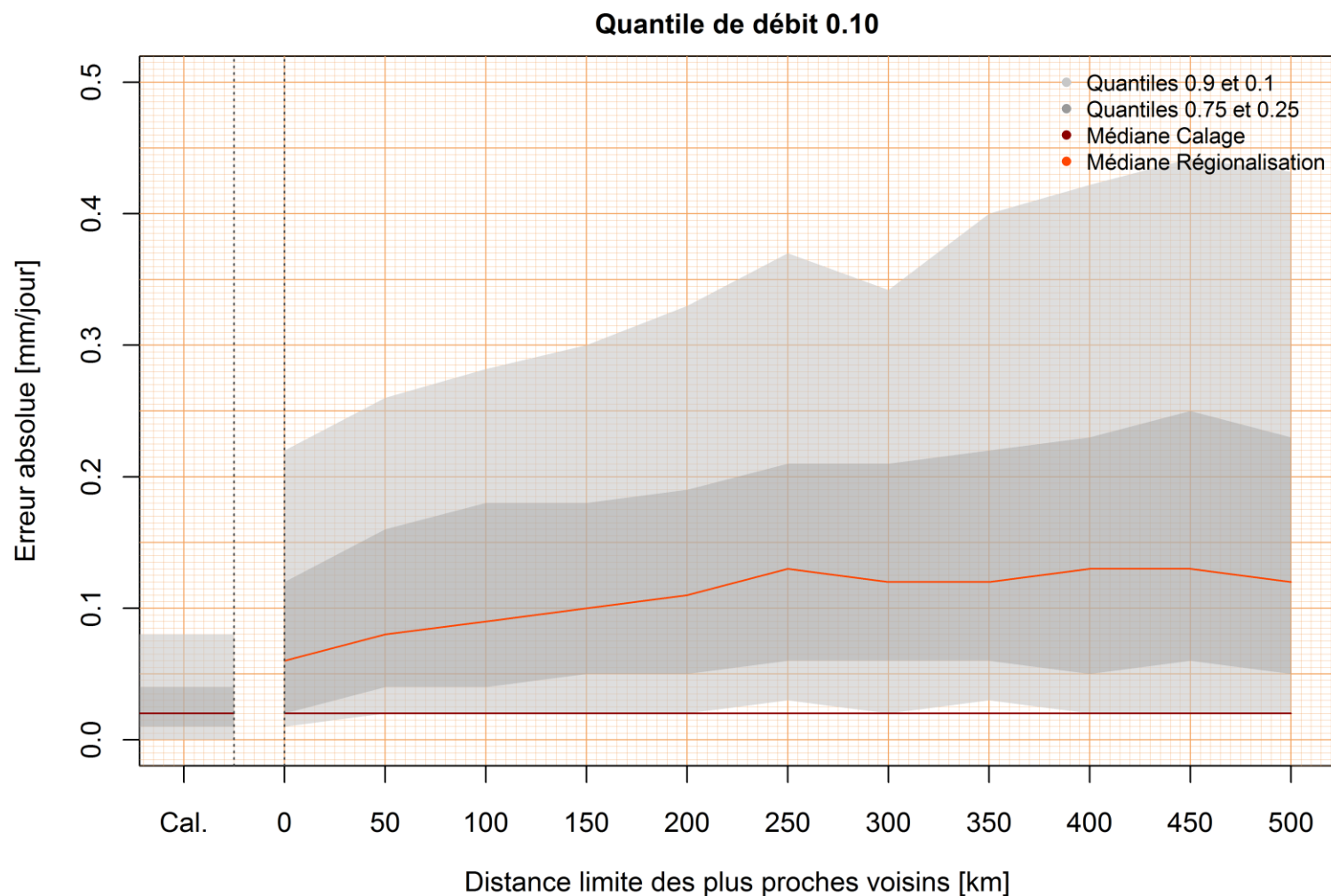


Figure 44. Influence de la distance à laquelle se situe le bassin voisin le plus proche sur l'efficacité de la régionalisation avec le modèle GR4J pour le calcul du quantile de débit 0.10. « Cal. » présente le cas où les débits simulés ont été obtenus en validation sur la même période que pour la régionalisation mais avec les paramètres calés sur le bassin cible. La distance 0 km correspond au cas où aucune limite de distance des plus proches voisins n'est appliquée.

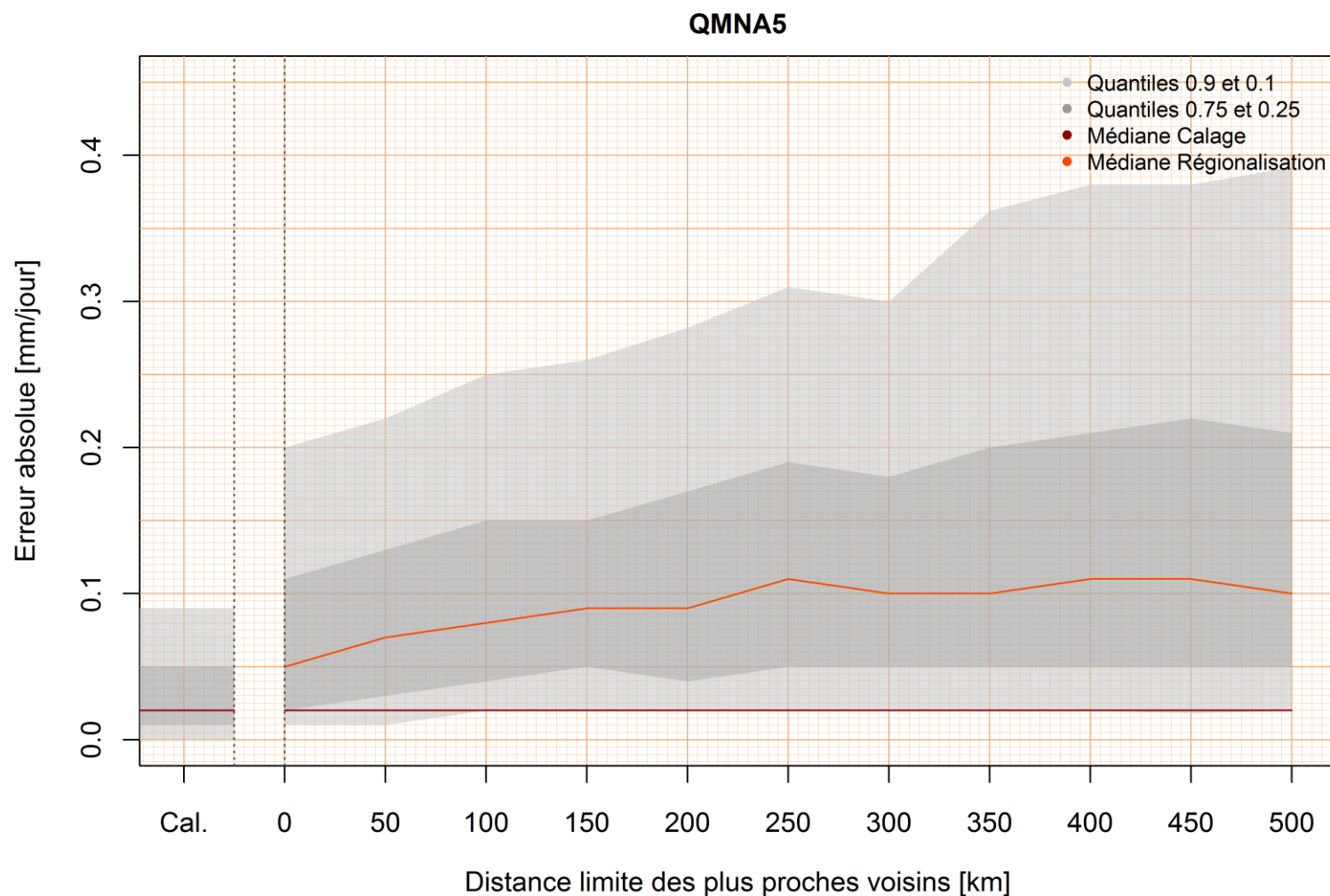


Figure 45. Influence de la distance à laquelle se situe le bassin voisin le plus proche sur l'efficacité de la régionalisation avec le modèle GR4J pour le calcul du QMNA5. « Cal. » présente le cas où les débits simulés ont été obtenus en validation sur la même période que pour la régionalisation mais avec les paramètres calés sur le bassin cible. La distance 0 km correspond au cas où aucune limite de distance des plus proches voisins n'est appliquée.

5.6. Analyse détaillée des performances obtenues en régionalisation

Les Figure 46 à Figure 52 présentent les mêmes résultats que précédemment, dans une perspective cartographique. Chaque point représente une station et la couleur du point représente la performance du modèle. Le biais obtenu en régionalisation est également présenté en fonction de la distance du plus proche voisin. Avec ces cartes, on cherche à comprendre:

- D'une part, s'il existe un caractère régional expliquant les performances des modèles;
- D'autre part, si la chute de performances lors du passage d'un modèle « calé » à un modèle régionalisé est liée au premier ordre à la proximité au plus proche voisin (on a montré auparavant que l'éloignement diminue les performances de la régionalisation, ce que l'on cherche à montrer c'est si cet effet explique les mauvais résultats dans certaines régions).

Globalement, les figures suivantes montrent que l'effet de la densité n'est pas du premier ordre: les performances des calculs ne semblent pas être liées à la distance du plus proche voisin.

Pour la simulation du débit journalier (Figure 46), c'est un effet régional qui est clairement apparent, aussi bien en situation jaugée que non jaugée. Le massif Armoricaïn par exemple se caractérise par des performances très élevées aussi bien au calage qu'en non-jaugé. En général, si le modèle GR4J a des difficultés à reproduire le débit, ces difficultés s'aggraveront en régionalisation.

Pour les autres caractéristiques de débit, aucune tendance ne semble se dégager clairement. Des bassins peuvent avoir des mauvaises performances pour une caractéristique mais pas obligatoirement pour les autres.

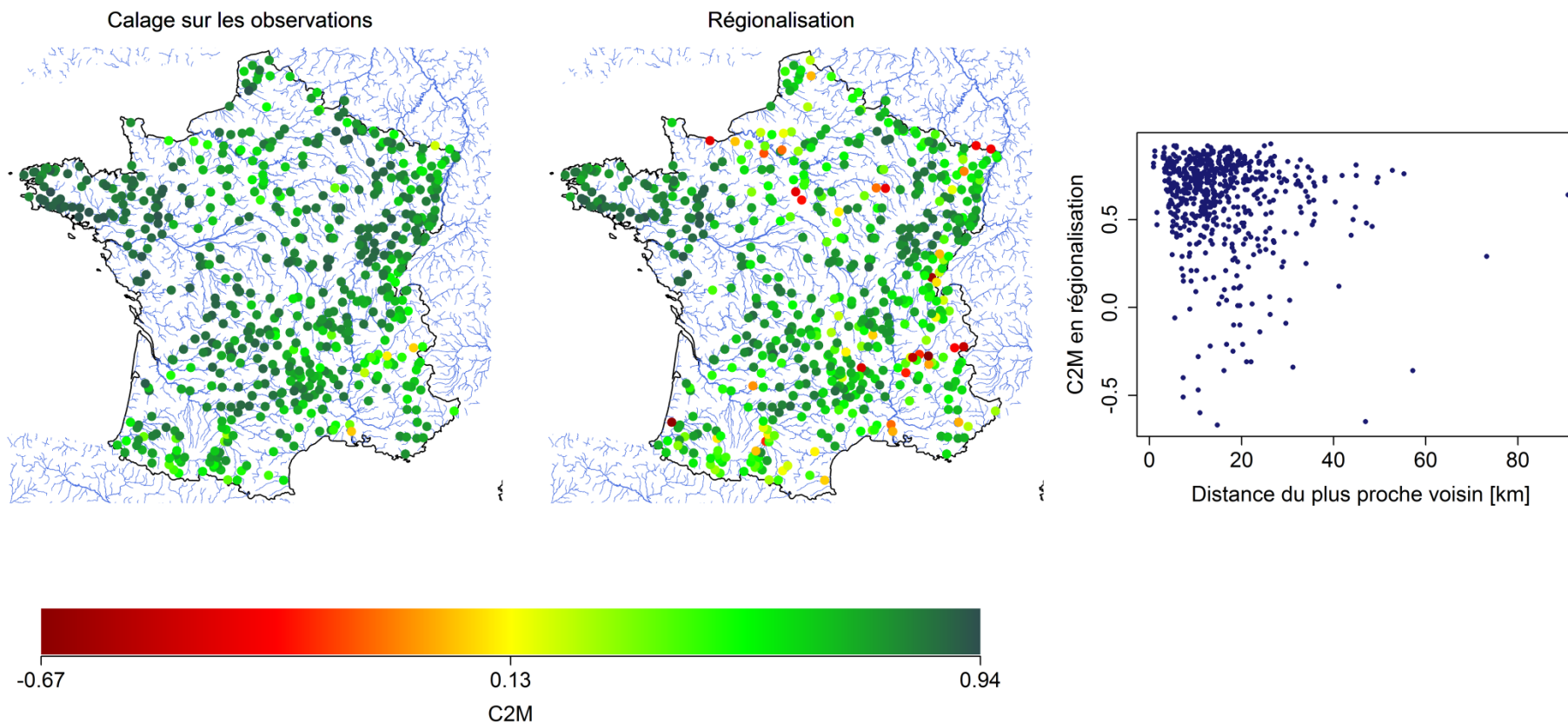


Figure 46. Cartographie des performances (ici le critère C2M) pour la simulation d'une chronique de débit journalier, dans le cas d'un calage des paramètres effectués sur les observations à gauche et dans le cas d'utilisation de la méthode de régionalisation sans limite de distance au milieu. A droite, graphique représentant les performances obtenues en régionalisation en fonction de la distance du plus proche voisin.

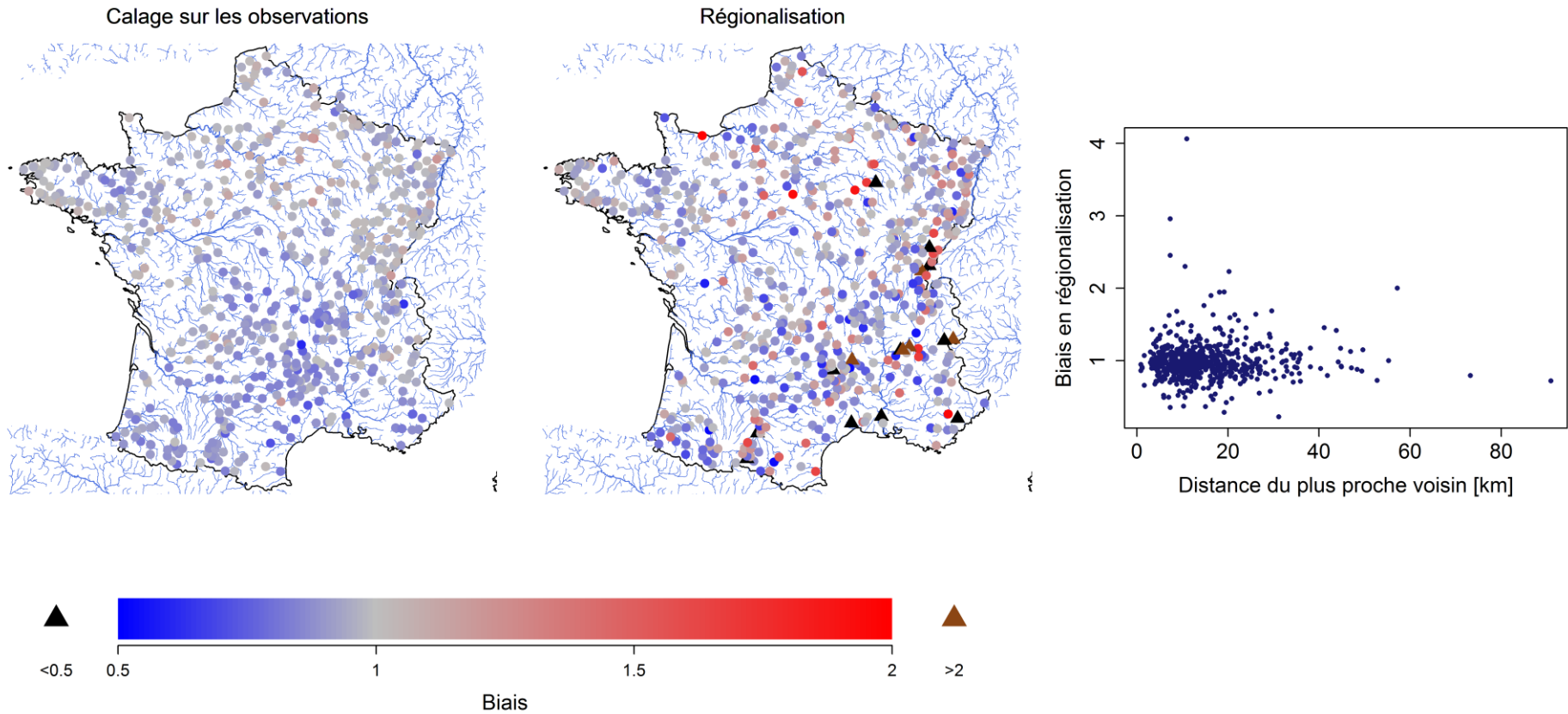


Figure 47. Cartographie des performances (ici le biais (Q_{sim}/Q_{obs})) pour l'estimation du module, dans le cas d'un calage des paramètres effectués sur les observations à gauche et dans le cas d'utilisation de la méthode de régionalisation sans limite de distance au milieu. A droite, graphique représentant le biais en fonction de la distance du plus proche voisin.

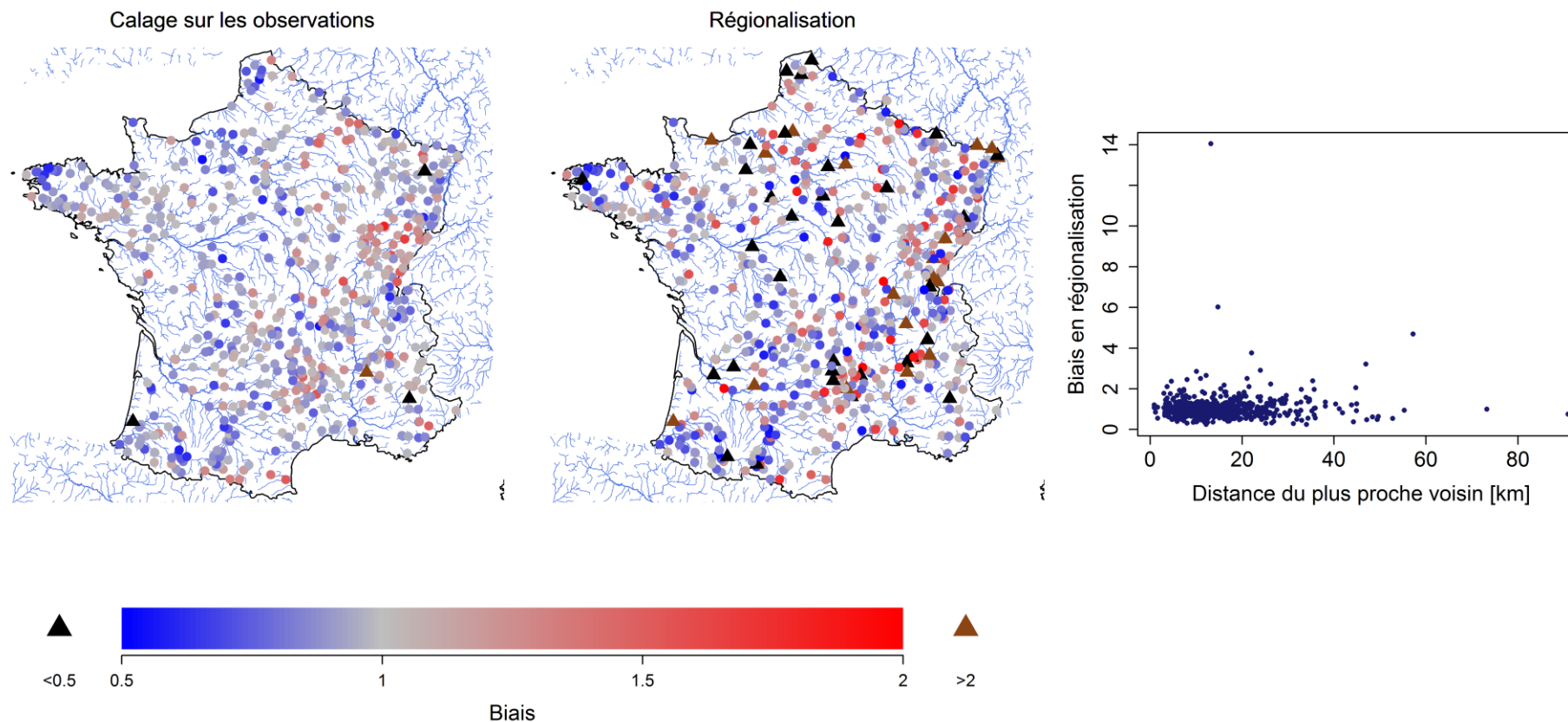


Figure 48. Cartographie des performances (ici le biais (Q_{sim}/Q_{obs})) pour l'estimation du débit de crue décennale, dans le cas d'un calage des paramètres effectués sur les observations à gauche et dans le cas d'utilisation de la méthode de régionalisation sans limite de distance au milieu. A droite, graphique représentant le biais en fonction de la distance du plus proche voisin.

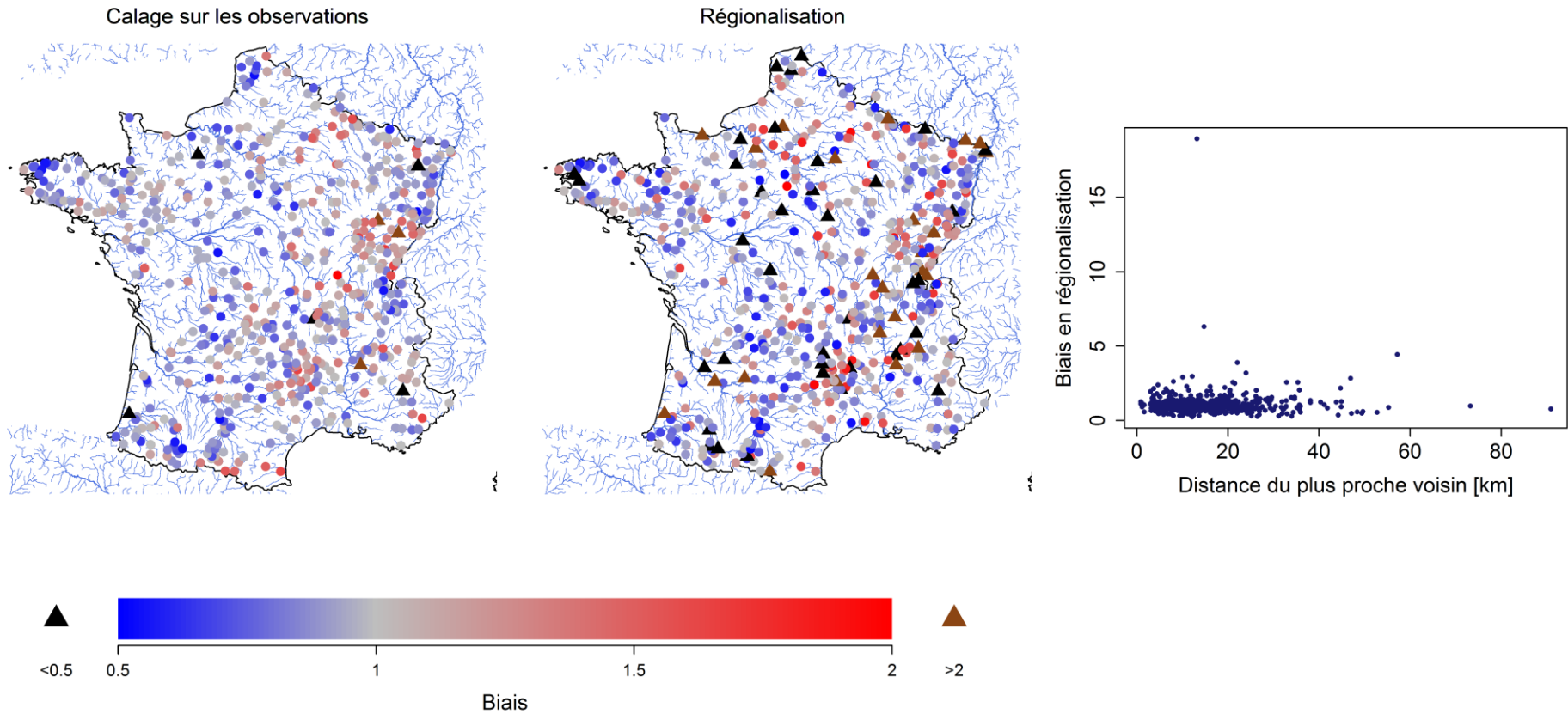


Figure 49. Cartographie des performances (ici le biais (Q_{sim}/Q_{obs})) pour l'estimation du débit de crue centennale, dans le cas d'un calage des paramètres effectués sur les observations à gauche et dans le cas d'utilisation de la méthode de régionalisation sans limite de distance au milieu. A droite, graphique représentant le biais en fonction de la distance du plus proche voisin.

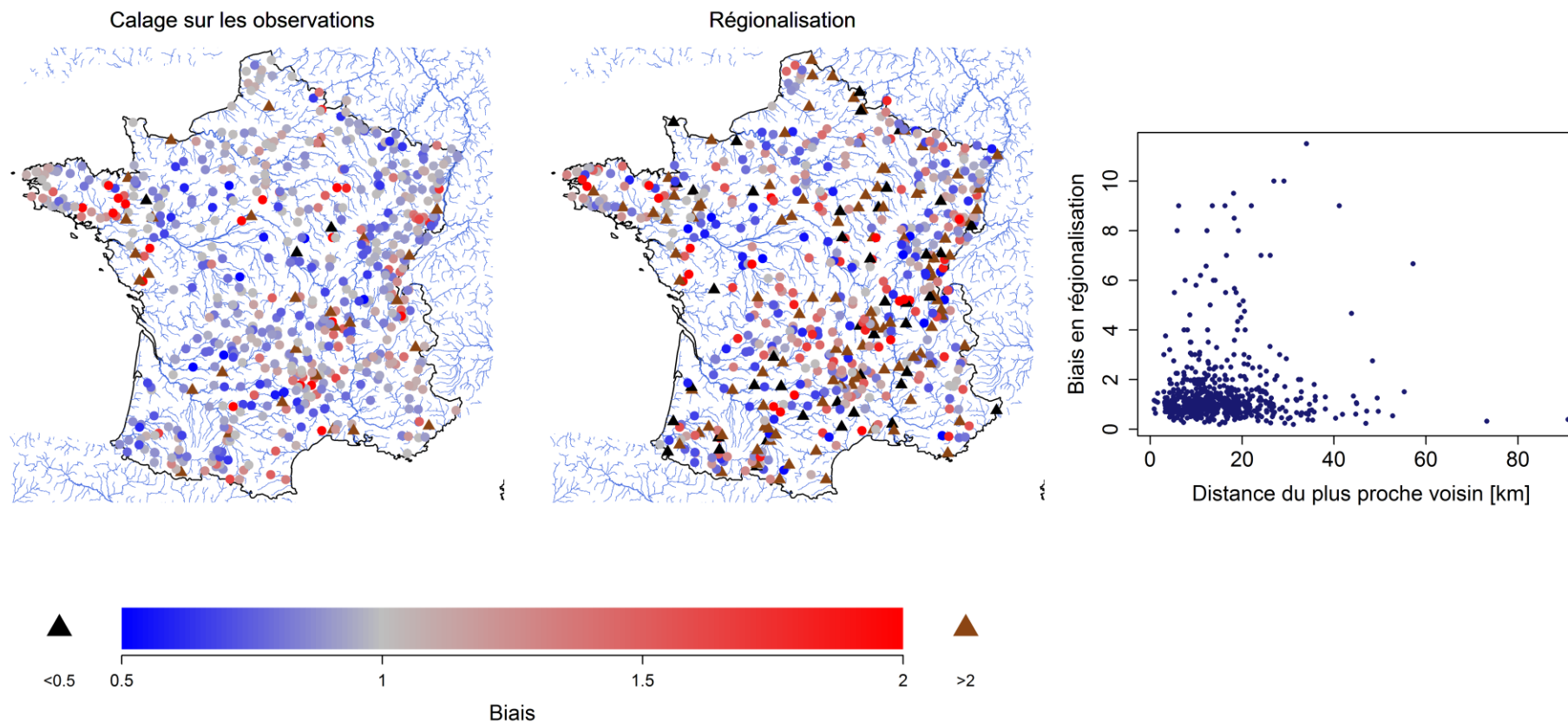


Figure 50. Cartographie des performances (ici le biais (Q_{sim}/Q_{obs})) pour l'estimation du quantile de débit 0.05, dans le cas d'un calage des paramètres effectués sur les observations à gauche et dans le cas d'utilisation de la méthode de régionalisation sans limite de distance au milieu. A droite, graphique représentant le biais en fonction de la distance du plus proche voisin.

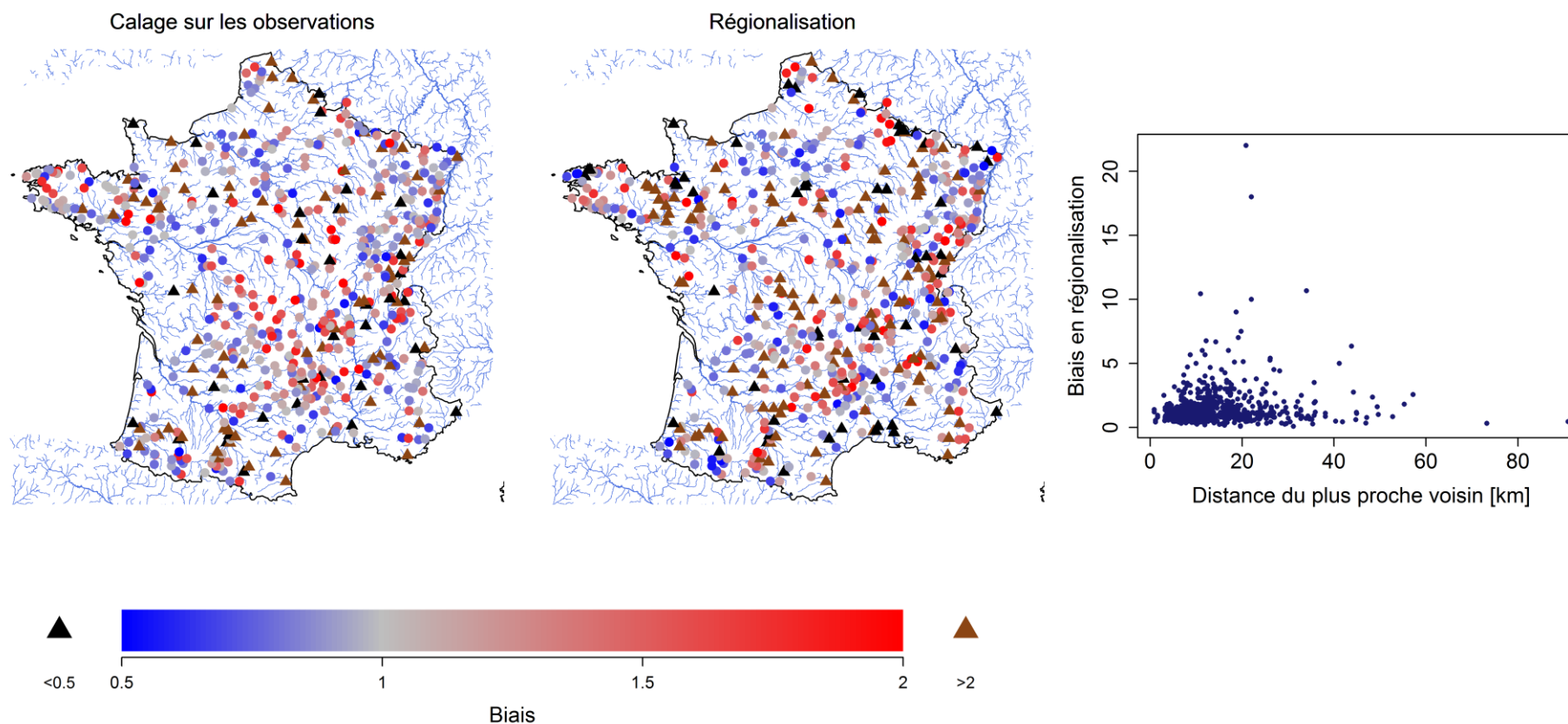


Figure 51. Cartographie des performances (ici le biais (Q_{sim}/Q_{obs})) pour l'estimation du quantile de débit 0.10, dans le cas d'un calage des paramètres effectués sur les observations à gauche et dans le cas d'utilisation de la méthode de régionalisation sans limite de distance au milieu. A droite, graphique représentant le biais en fonction de la distance du plus proche voisin.

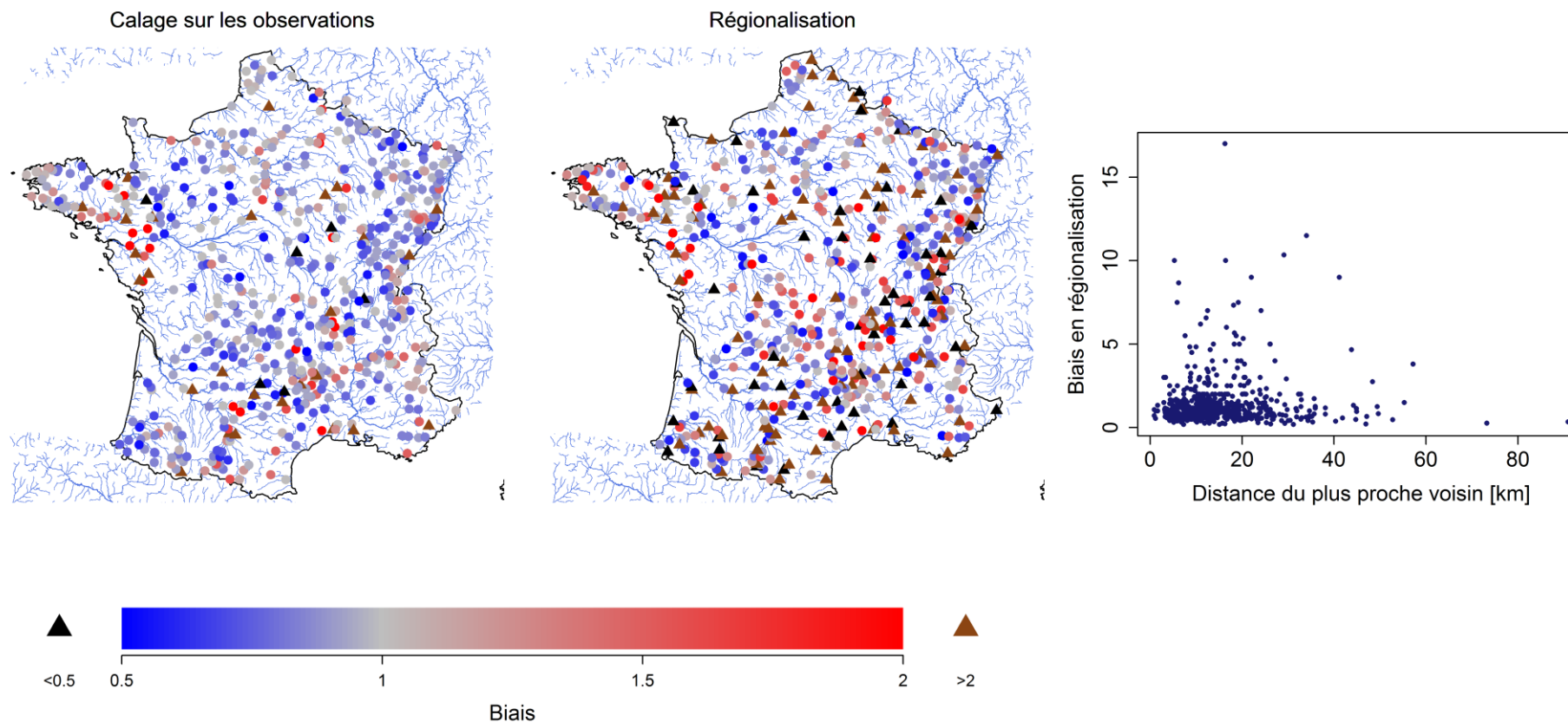


Figure 52. Cartographie des performances (ici le biais (Q_{sim}/Q_{obs})) pour l'estimation du QMNA5, dans le cas d'un calage des paramètres effectués sur les observations à gauche et dans le cas d'utilisation de la méthode de régionalisation sans limite de distance au milieu. A droite, graphique représentant le biais en fonction de la distance du plus proche voisin.

Les cartes ci-dessus ne présentant pas de tendance claire pour l'ensemble de l'échantillon, nous allons maintenant nous intéresser à deux bassins en particulier et essayer de comprendre les résultats obtenus. Ces deux bassins sont le Jabron à Comps-sur-Artuby (code hydro : X2305010) et la Siagne à Callian (code hydro : Y5514040).

Ces bassins ont un comportement un peu particulier, notamment pour le calcul régionalisé du module (Figure 53). En effet, ces deux bassins voisins ont dans un cas un biais inférieur à 0.5 (Y5514040) et dans l'autre un biais égal à 1.94, soit très proche de 2 (X2305010). Cependant, en raison de leur grande proximité, la méthode de régionalisation les utilise comme donneurs l'un pour l'autre.

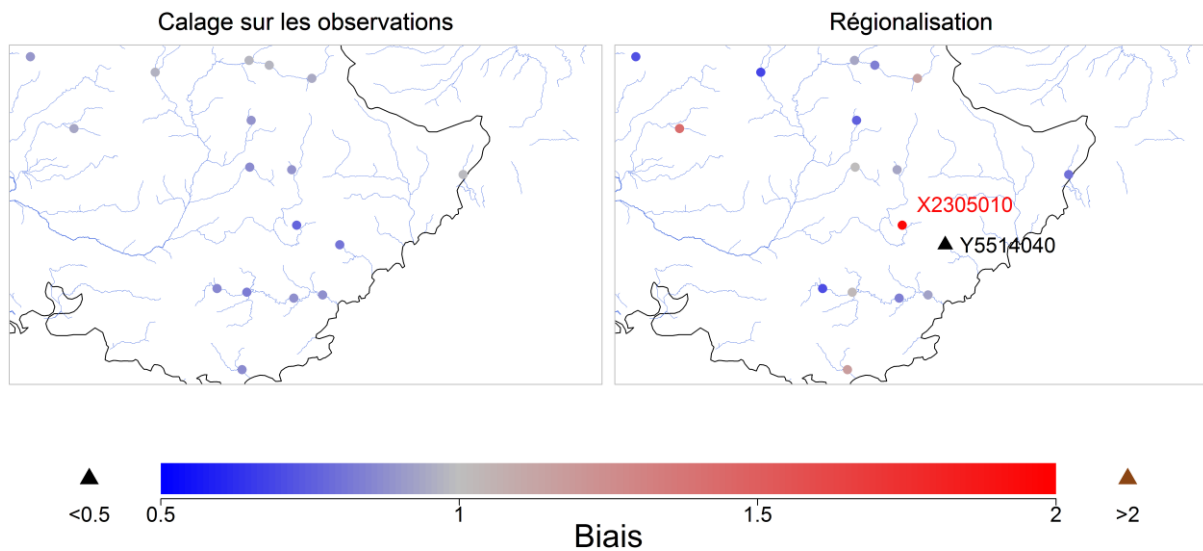


Figure 53. Cartographie des performances pour l'estimation du module, dans le cas d'un calage des paramètres effectués sur les observations à gauche et dans le cas d'utilisation de la méthode de régionalisation sans limite de distance à droite. Zoom sur les bassins du Jabron à Comps-sur-Artuby (code hydro : X2305010) et de la Siagne à Callian (code hydro : Y5514040).

L'existence de deux bassins contigus mais présentant un biais opposé est surprenante. Pour essayer de pousser le diagnostic, il est pratique de représenter chaque bassin dans le graphe adimensionnel de Turc-Budyko : ce graphique relie le coefficient d'écoulement (Q/P) à l'indice d'aridité (P/ETP) (Figure 54). Nous observons que le bassin de la Siagne à Callian (Y5514040) se situe au-dessus de la limite du réalisme du bilan en eau, c'est-à-dire dans la zone où les bassins « gagnent » nécessairement de l'eau ($Q > P$). A l'inverse, nous constatons que le bassin du Jabron à Comps-sur-Artuby (X2305010) se situe en dessous de la limite du réalisme du bilan énergétique, c'est-à-dire dans la zone où les bassins « perdent » de l'eau ($Q < P - ETP$: un bassin ne peut pas perdre plus d'eau dans l'atmosphère qu'il n'y a d'énergie pour l'évaporer).

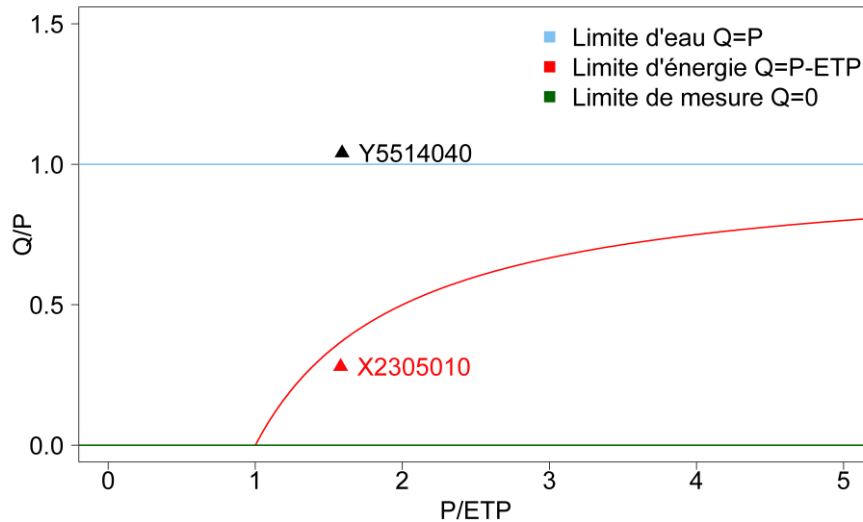


Figure 54. Localisation des bassins versants du Jabron à Comps-sur-Artuby (code hydro : X2305010) et de la Siagne à Callian (code hydro : Y5514040) sur le graphique adimensionnel reliant le coefficient de production (Q/P) à l'indice d'aridité (P/ETP).

Mais la méthode de régionalisation des paramètres ne permet pas de rattraper de telles différences de comportement : comment corriger un bassin qui « perd » de l'eau avec un voisin qui en « gagne » et inversement ? Ceci explique les mauvais résultats obtenus pour ces deux bassins pour le calcul du module en régionalisation.

Nous allons maintenant regarder plus en détail les autres variables hydrologiques calculés pour ces deux bassins (Figure 55).

Nous pouvons observer pour le bassin du Jabron à Comps-sur-Artuby (X2305010) que toutes les variables hydrologiques calculées sont surestimées par le modèle, comme pour le calcul du module.

Par contre, pour le bassin de la Siagne à Callian (Y5514040), la situation est moins tranchée. Pour les calculs de débit de crue, le biais est dans les deux cas inférieur à 1. Les débits de crue sont alors sous estimés, comme pour le calcul du module. Alors que pour les caractéristiques d'étiage, nous pouvons constater une surestimation de ces débits.

Ce que l'on peut observer de manière générale pour toutes les variables sur la Figure 55, c'est que le biais du bassin du Jabron à Comps-sur-Artuby (X2305010) est toujours supérieur à celui du bassin de la Siagne à Callian (Y5514040).

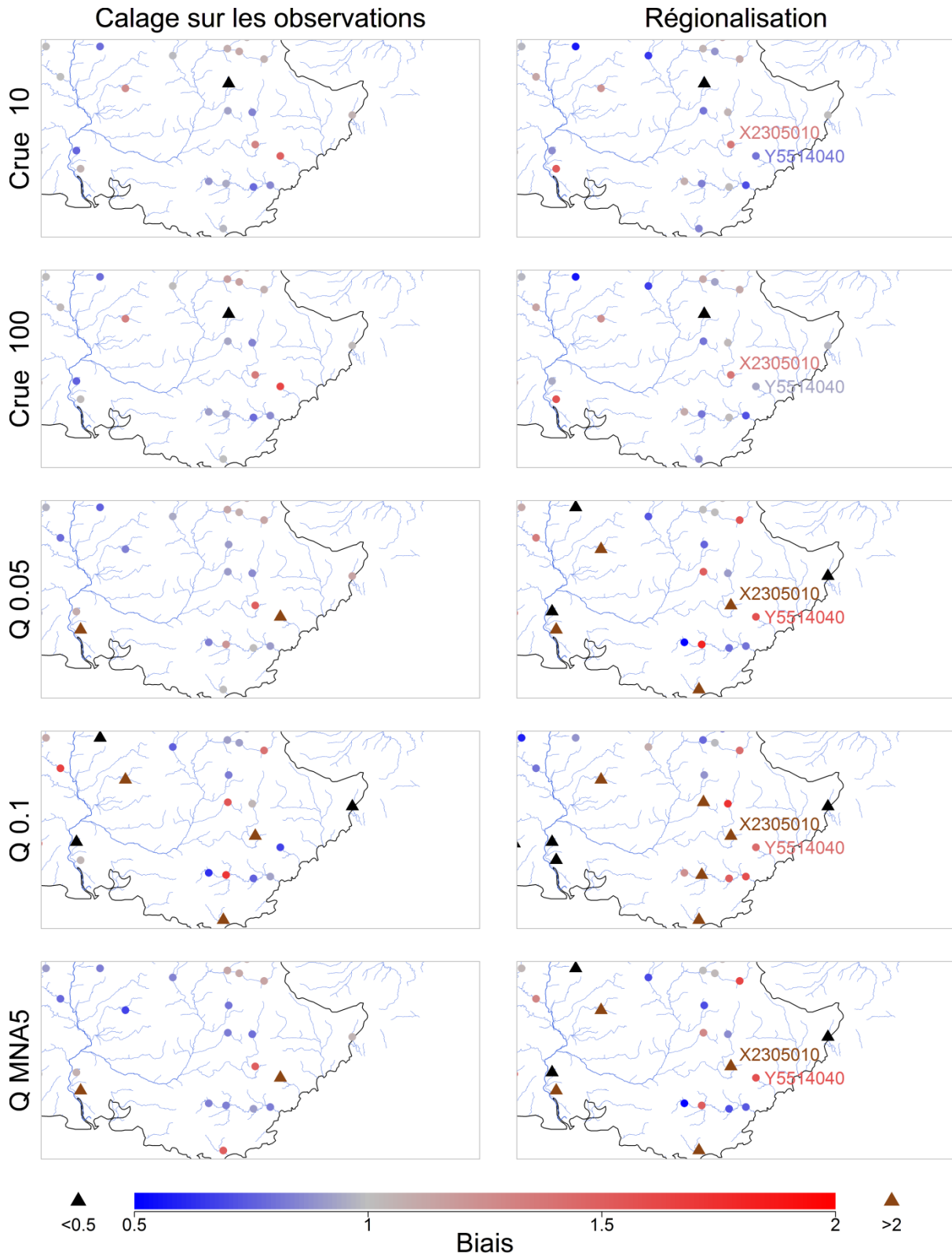


Figure 55. Cartographie des performances pour l'estimation des débits de crue et des caractéristiques de débit d'étiage, dans le cas d'un calage des paramètres effectués sur les observations à gauche et dans le cas d'utilisation de la méthode de régionalisation sans limite de distance à droite. Zoom sur les bassins du Jabron à Comps-sur-Artuby (code hydro : X2305010) et de la Siagne à Callian (code hydro : Y5514040).

La Figure 56 illustre à l'aide d'hydrogrammes observés et régionalisés, les résultats obtenus précédemment. La première chose que l'on peut constater pour les deux bassins versants est que peu importe la fonction objectif utilisée pour caler les paramètres des bassins voisins, la régionalisation peine à bien modéliser le débit de manière générale (le modèle est tout aussi mauvais en basses eaux qu'en hautes eaux). Les mauvais résultats obtenus pour ces bassins ne sont donc pas expliqués par le choix de la fonction objectif utilisée pour le calage des voisins. D'autre part, les hydrogrammes confirment que pour le bassin du Jabron à Comps-sur-Artuby (X2305010), le débit est systématiquement surestimé par les modèles régionalisés alors que pour le bassin de la Siagne à Callian (Y5514040), les moyens et hauts débits sont sous-estimés et les bas débits ont tendance à être surestimés.

Par ailleurs, on peut facilement supposer que des échanges souterrains d'eau ont lieu entre ces deux bassins voisins.

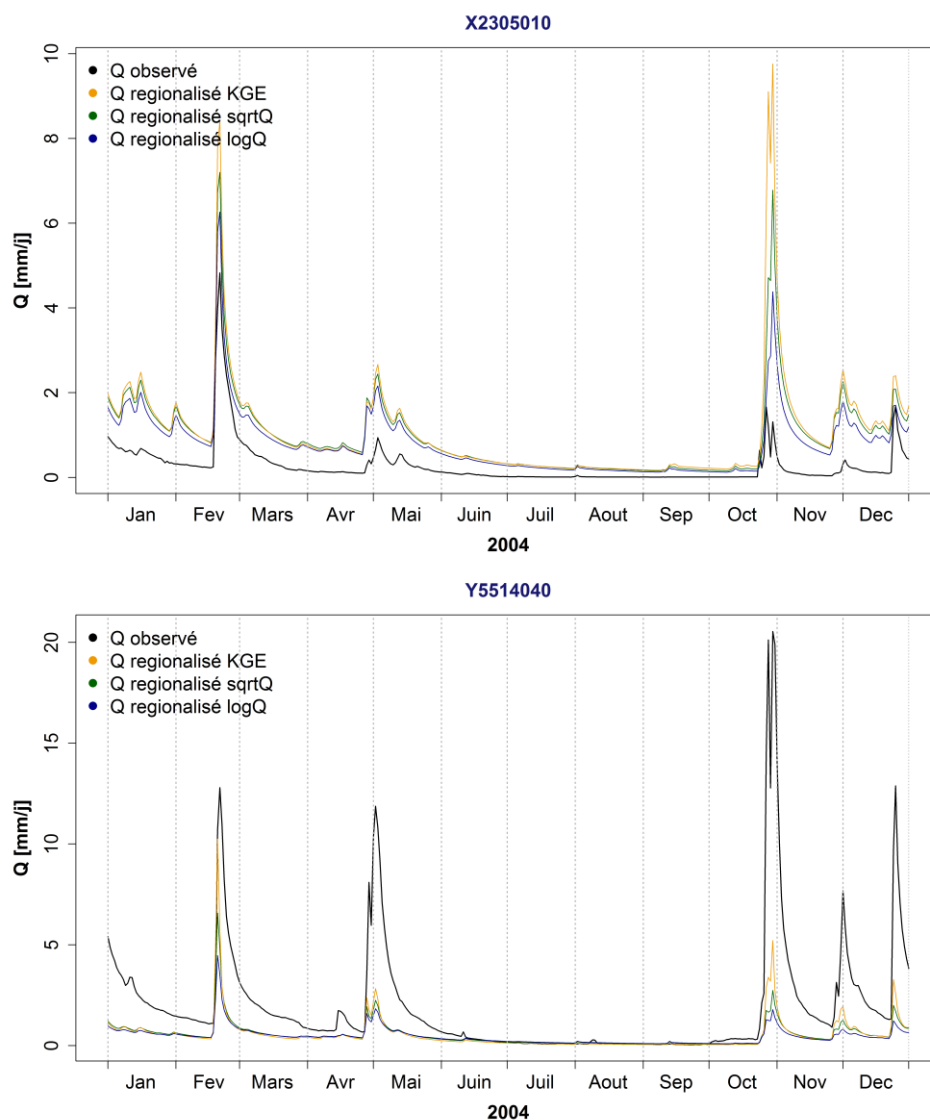


Figure 56. Hydrogrammes observés (en noir) et régionalisés sans limite de distance des plus proches voisins sur le bassin du Jabron à Comps-sur-Artuby (code hydro : X2305010) et de la Siagne à Callian (code hydro : Y5514040) pour l'année 2004. Chroniques régionalisées à partir de paramètres voisins calés avec Nash sqrtQ (en vert), Nash logQ (en bleu) et KGE Q (en jaune).

5.7. Synthèse

Le but de ce chapitre était d'évaluer l'impact de la densité spatiale du réseau hydrométrique sur des calculs de débit réalisés sur des bassins non jaugés. Nous avons abordé la simulation de chroniques de débit journalier, l'estimation du module, des débits de crues décennale et centennale et de trois caractéristiques de débit d'étiage dont le QMNA5. Nous avons appliqué la méthode du désert hydrométrique afin d'évaluer la robustesse de la méthode de régionalisation lorsque celle-ci est appliquée au sein d'un réseau hydrométrique de moins en moins dense.

Nos résultats montrent que pour tous ces calculs, la performance du processus de régionalisation diminue lorsque le réseau de bassins voisins devient moins dense (lorsque l'on supprime les bassins voisins les plus proches). Toutefois, cette chute d'efficacité liée à l'augmentation de la distance des plus proches voisins est moindre en comparaison de celle attribuée à la méthode de régionalisation elle-même. D'autre part, la quantification de la perte de performances est variable entre les différentes variables et aucune explication spatiale (proximité du plus proche voisin, zone géographique de la zone d'étude) ne semble se démarquer.

Par ailleurs, pour une application opérationnelle sur un bassin versant non-jaugé donné, en connaissant la distance à laquelle se trouve le plus proche voisin, les courbes obtenues aux Figure 39 à Figure 45 donnent une estimation des gammes de performances auxquelles on peut s'attendre sur un bassin versant non jaugé et pour divers calculs hydrologiques.

En conclusion, la densité spatiale du réseau hydrométrique a un impact sur les performances des calculs hydrologiques en situation non jaugée. En effet, l'information des bassins voisins est essentielle au bassin non jaugé pour la simulation du débit en mode non jaugé.

Chapitre 6.
**Stratégie de récolte de jaugeages ponctuels
sur des bassins versants non-jaugés : au sujet
de la redondance de l'information
hydrologique**

6.1. Introduction

Il existe beaucoup d'endroits dans le monde où il est difficile de maintenir des stations de mesure de débit à long terme et où seulement quelques mesures sont disponibles [Bhatt et Tiwari, 2008; Kim et Kaluarachchi, 2009; Konz et al., 2007]. Par ailleurs, dans le cas de bassins versants non jaugés, les hydrologues n'ont souvent ni le temps ni les ressources financières pour installer une station de jaugeage pérenne, mais ils ont toujours la possibilité d'effectuer ponctuellement une campagne de jaugeages, leur permettant de récolter quelques mesures de débit.

Dans ce chapitre, nous allons dans un premier temps présenter un article en cours de correction afin d'être soumis une seconde fois. Cet article basé sur les résultats de la thèse de Rojas-Serna [2005], met en place une méthode permettant de paramétrer un modèle pluie-débit sur des bassins très peu jaugés. Pour cela, la méthode combine une information partielle fournie par les quelques mesures de débit du bassin cible, avec une information régionale, fournie par les paramètres des bassins voisins. Cette étude vise également à déterminer les poids attribués à chacune des sources d'information de cette combinaison, selon le nombre de mesures aléatoires disponibles sur le bassin cible.

Dans la suite du chapitre, nous avons repris cette méthode mais en essayant de la pousser plus loin, afin d'évaluer quelles étaient les gammes de débit les plus informatives du point de vue du gestionnaire. En pratique, si nous devons lancer une campagne de jaugeage, quel quantile de débit devrions-nous mesurer de façon à être le plus efficace pour la détermination des paramètres du modèle à transférer ?

6.2. Sélection de jaugeages aléatoires (SA) - Résumé de l'article présenté en annexe B

Cet article fait suite à la thèse de Rojas-Serna [2005]. Suite à son rejet, nous avons choisi de le reprendre dans le cadre de ces travaux de thèse, pour en confirmer les résultats et apporter des résultats complémentaires sur la base de notre échantillon de données. Les nouveaux calculs ont donc été effectués sur l'échantillon de bassins versants présentés dans le chapitre 3. Je ne présente ici qu'un résumé de la méthode, pour permettre au lecteur de suivre ce que j'ai fait d'original par rapport au travail initial.

6.2.1 La méthode

La méthode vise à paramétrer un modèle pluie-débit (ici les modèles GR4J et TOPMO) sur des bassins très peu jaugés. Elle consiste à combiner une information partielle fournie par les quelques mesures de débit du bassin cible, avec une information régionale, fournie par une méthode de régionalisation (ici, la proximité géographique). Les quelques mesures sont choisies aléatoirement (ce ne sera plus le cas dans les tests de la section suivante).

L'approche combinatoire consiste en une combinaison de rangs. La base de données de bassins jaugés permet de proposer des jeux de paramètres, que l'on peut classer :

- soit en fonction de la proximité géographique des bassins considérés avec le bassin cible non jaugé (rang régional r_j^{reg});

- soit en fonction de l'erreur de simulation (que l'on ne peut bien sûr calculer que sur le petit nombre de jaugeages) : les jeux de paramètres sont classés suivant cette information locale (rang local r_j^{loc}).

Le classement des jeux de paramètres candidats se fait sur la base d'un rang combiné r_j :

$$r_j = \alpha \cdot r_j^{reg} + (1-\alpha) \cdot r_j^{loc} \quad \text{Eq. 24}$$

avec α , le coefficient pondérateur variant entre 0 et 1 exprimant l'importance relative de l'information régionale par rapport à l'information locale.

Quand α est égal à 1, la méthode équivaut à une simple approche de régionalisation et quand α est égal à 0, la méthode utilise seulement les quelques mesure ponctuelles obtenues sur le bassin cible pour sélectionner les jeux de paramètres.

Enfin, les M jeux de paramètres les plus proches en termes de distance combinée r_j sont sélectionnés. M simulations sont réalisées et une simulation moyenne est produite.

6.2.2 Résultats

Les résultats peuvent être illustrés par deux types de graphes :

- un graphe présentant la performance moyenne en fonction du nombre de jaugeages réalisés et de la valeur de α (Figure 57);
- un graphe présentant la valeur de α à sélectionner selon le nombre de jaugeages réalisés (Figure 58). Une relation entre α optimale et le nombre de jaugeages réalisés peut ainsi être obtenue pour chaque critère d'évaluation.

La Figure 59 présente les performances finalement obtenues en appliquant la méthode, en comparaison de celles du modèle GR4J obtenues dans le cas où les bassins versants sont totalement jaugés et le cas où ils ne le sont pas du tout. Ainsi, la méthode permet d'obtenir des résultats intermédiaires entre le cas d'un bassin totalement jaugé et celui d'un bassin totalement non jaugé. De meilleures performances sont obtenues par rapport à la simple utilisation de la méthode de régionalisation.

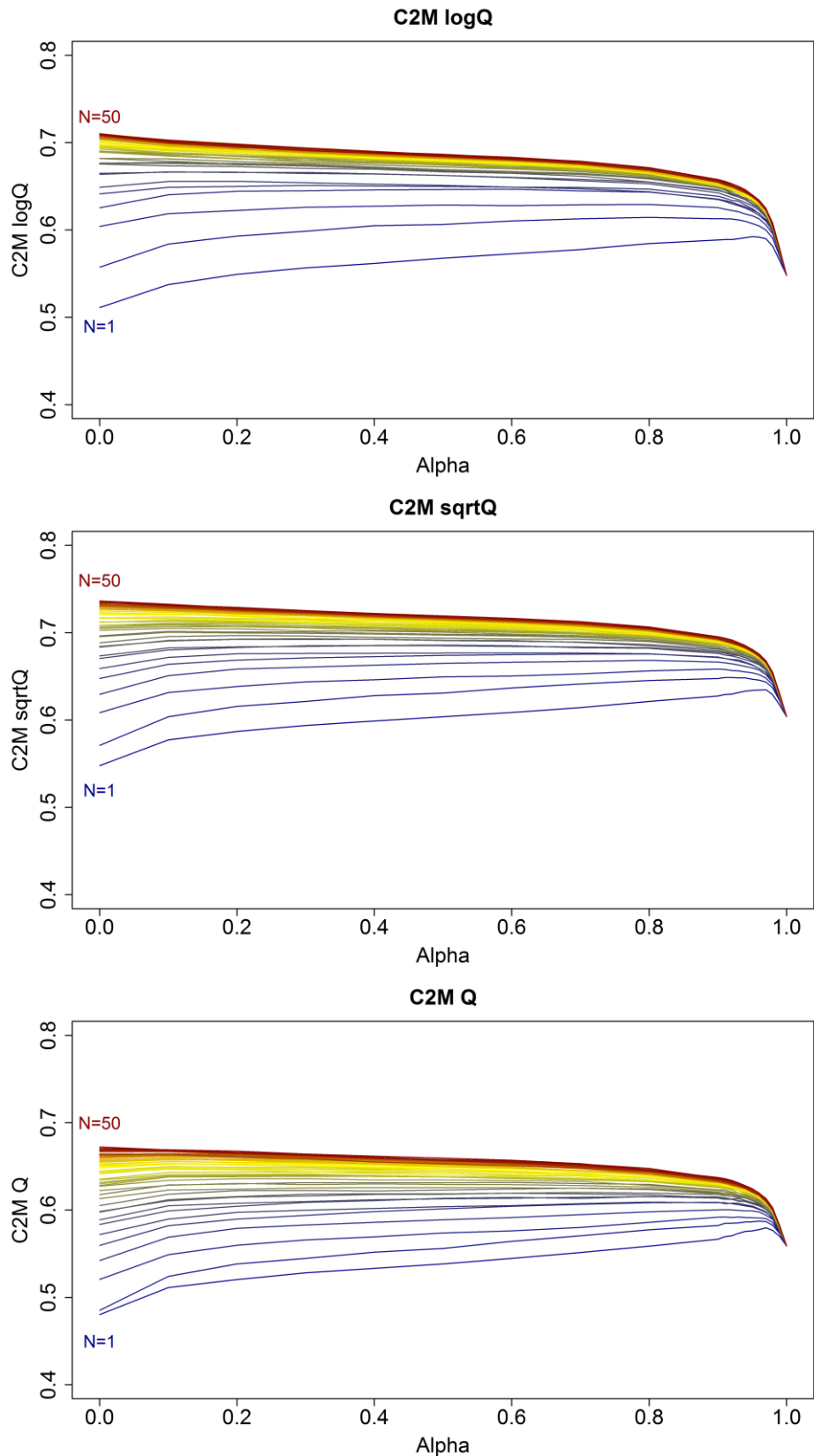


Figure 57. Performances moyennes du modèle GR4J en fonction de α et du nombre de jaugeages réalisés, pour trois critères d'évaluation (C2M calculé sur le logarithme des débits, C2M calculé sur la racine carrée des débits et C2M calculé sur les débits).

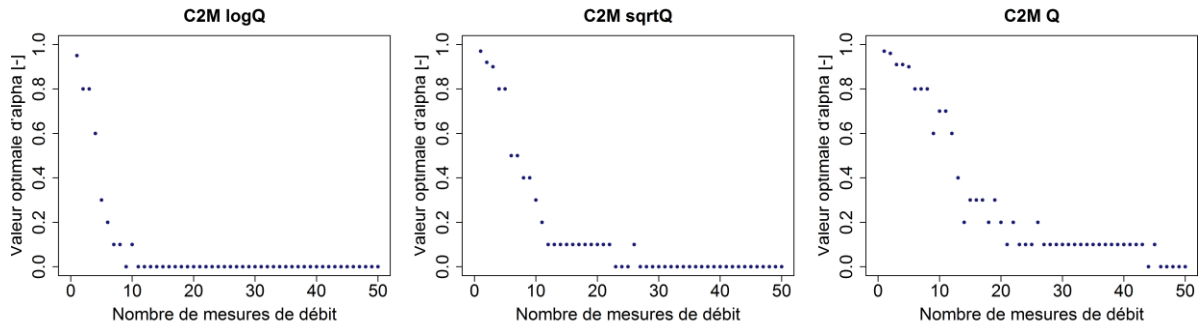


Figure 58. Valeur optimale de α en fonction du nombre de jaugeages réalisés pour le modèle GR4J et pour trois critères d'évaluation (C2M calculé sur le logarithme des débits, C2M calculé sur la racine carrée des débits et C2M calculé sur les débits).

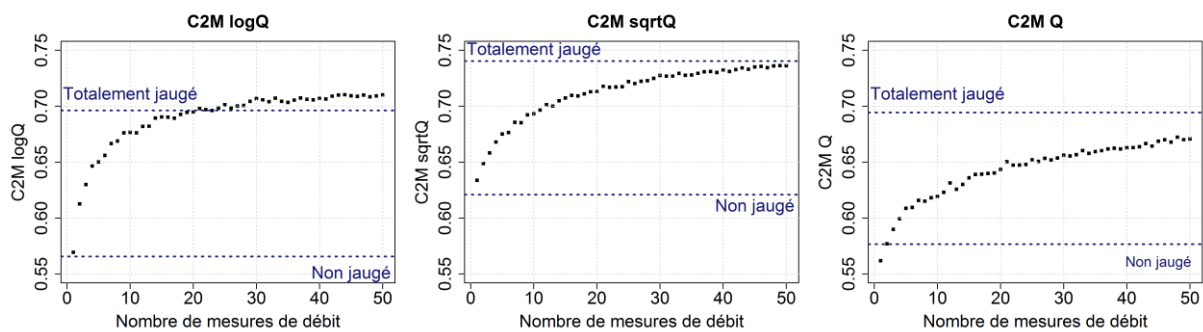


Figure 59. Performances moyennes du modèle GR4J après l'application de la méthode présentée, pour les trois critères d'évaluation. Les traits bleus pointillés représentent les performances moyennes du modèle lorsque les bassins versants sont totalement jaugés et non jaugés.

6.3. Sélection du débit observé selon la gamme de débit classé (sélection redondante (SR))

6.3.1 Introduction

D'un point de vue opérationnel, nous venons de voir que quelques mesures suffisent pour améliorer la méthode de régionalisation basée sur la proximité spatiale des bassins. Sur un bassin versant totalement non jaugé, il « suffirait » ainsi de pouvoir réaliser quelques jaugeages pour récolter un nombre suffisant de mesures de débit.

Imaginons à présent un bassin versant sur lequel nous disposons d'une chronique de hauteur d'eau mais sur lequel aucune courbe de tarage n'a été établie. Si un jaugeage de débit était effectué sur ce bassin, nous obtiendrions un débit mesuré pour une hauteur d'eau donnée. En supposant que la relation hauteur-débit soit restée stable dans le temps et que la relation débit-hauteur est univoque (pas de phénomène d'hystérésis par exemple), à chaque fois que cette hauteur d'eau est rencontrée dans la chronique, on peut lui associer le débit mesuré lors de l'unique jaugeage effectué. On se retrouve ainsi avec une chronique de débit composée de N' mesures de débit égales sur laquelle on se propose d'appliquer la méthode établie précédemment. Pour la suite de l'étude, nous appellerons cette méthode de sélection, la sélection redondante des débits (SR). Cette étude est complémentaire aux travaux de *Rojas-Serna* [2005] portant sur la sélection aléatoire des débits (SA).

Les questions que l'on se pose sont alors les suivantes :

- quelle gamme de débit est-il préférable de mesurer lors de cet unique jaugeage ? Vaudrait-il mieux viser les bas débits, les débits moyens ou les hauts débits pour pouvoir exploiter au mieux la méthode dans ce contexte ?
- quelle différence existe-t-il entre une information aléatoire (cf. les tests de *Rojas-Serna* [2005]) et une information répétitive ? Quelle stratégie est préférable ?

Nous allons essayer de répondre maintenant à ces questions. Nous comparerons les résultats avec ceux de la méthode originale, qui s'appuie sur N mesures déterminées de façon aléatoire.

6.3.2 Méthodologie

Globalement, la méthodologie adaptée est la même que celle exposée précédemment dans la partie 6.2. Ici, seul le modèle GR4J a été utilisé et la fonction objectif reste le critère de *Nash et Sutcliffe* [1970] calculé sur la racine carrée des débits. Les critères d'évaluation restent également les C2M calculés sur les débits, la racine carrée des débits et le logarithme des débits [Mathevet et al., 2006]. Il est important de souligner que pour deux cas d'évaluation, le critère de calage n'a pas été adapté selon l'objectif hydrologique (le critère d'évaluation étant alors différent du critère de calage).

Le principal changement concerne la sélection des valeurs de débit à partir desquelles les bassins vont être classés selon leur efficacité à bien simuler ces quelques débits. Auparavant, cette sélection était aléatoire. A présent, nous retenons dans notre chronique de débit de chaque sous-période, les débits correspondant successivement aux quantiles 0.10, 0.20, ..., 0.90 de la courbe des débits classés. Ces derniers correspondront à l'unique débit ponctuel mesuré sur le bassin cible. Afin d'obtenir suffisamment de mesures pour chaque test effectué, nous avons sélectionné les débits ayant un écart de débit de moins de 10% par rapport au quantile testé.

Une fois ces débits sélectionnés, nous classons les jeux de paramètres selon la distance combinée r_j et nous reprenons la procédure présentée précédemment.

6.3.3 Sensibilité par rapport au critère d'évaluation et au paramètre α

La Figure 60 présente les performances du modèle en fonction de la valeur α et des quantiles de débit présents dans la chronique de débit observé pour les trois critères d'évaluation présentés précédemment.

Le critère d'évaluation C2M calculé sur le logarithme des débits (C2M logQ)

On s'intéresse ici à un critère numérique d'évaluation qui privilégie la qualité de simulation des bas débits [Oudin et al., 2006b].

Pour le critère d'évaluation C2M calculé sur le logarithme des débits, nous pouvons constater que les meilleures performances sont obtenues lorsque les quantiles de bas débits sont sélectionnés comme unique débit ponctuel à mesurer, alors que l'utilisation des quantiles de hauts débits fournit les moins bonnes performances. A noter toutefois que ce ne sont pas les quantiles des plus bas débits (0.1) qui sont les plus informatifs, mais des débits intermédiaires entre les débits médians et les bas

débites. Ceci est compréhensible car le critère d'évaluation calculé sur le logarithme des débits ne se focalise pas uniquement sur les plus bas débits (comme le fait par exemple un critère utilisant l'inverse des débits) mais sur une gamme plus large de bas débits [Pushpalatha et al., 2012].

La valeur optimale du paramètre α varie selon le quantile de débit utilisé. En effet, lorsque des quantiles de haut débit sont utilisés, la valeur optimale d' α se rapproche de 1, donnant ainsi peu de poids à l'utilisation de ces quantiles par rapport à l'utilisation de la méthode de régionalisation. Pour les bas débits, la valeur optimale d' α est plus équilibrée, autour de 0.5, traduisant ainsi un meilleur équilibre entre les deux sources d'information.

Le critère d'évaluation C2M calculé sur la racine carrée des débits (C2M sqrtQ)

On s'intéresse ici à un critère numérique d'évaluation qui privilégie la qualité de simulation des débits moyens.

Pour le critère d'évaluation C2M calculé sur la racine carrée des débits, nous pouvons constater que les meilleures performances sont obtenues lorsque les quantiles de débit 0.5 et 0.6 sont sélectionnés. Ce résultat était attendu car le critère calculé sur la racine carrée des débits traduit la performance « moyenne » (au sens des débits) du modèle [Oudin et al., 2006b; Pushpalatha et al., 2012]. La différence de performances entre les quantiles est toutefois moins étendue que pour le critère C2M logQ.

De plus, les performances évoluent toujours en fonction du paramètre α . Nous pouvons constater que la valeur optimale d' α pour les quantiles 0.5 et 0.6 est de 0.8, privilégiant ainsi l'utilisation de la régionalisation pour le classement des bassins donneurs.

Le critère d'évaluation C2M calculé sur les débits (C2M Q)

On s'intéresse ici à un critère numérique d'évaluation qui privilégie la qualité de simulation des hauts débits.

Pour le critère d'évaluation C2M calculé sur les débits, les meilleures performances sont obtenues lorsque les quantiles de hauts débits sont sélectionnés. De la même manière, ces résultats étaient attendus car le C2M calculé sur les débits a tendance à privilégier les hauts débits [Oudin et al., 2006b; Pushpalatha et al., 2012].

Quand α est égal à 0, c'est la sélection des quantiles de hauts débits qui apporte les meilleures performances. Par ailleurs, les performances obtenues pour les différents quantiles évoluent toujours en fonction de α . Pour le quantile de hauts débits 0.9, la valeur optimale d' α est de 0.8, alors que pour les quantiles de bas débits, elle est autour de 0.95, donnant ainsi moins de poids à l'information locale que pour les quantiles de hauts débits.

Ainsi, la méthode présentée en 6.2 semble sensible au choix de quantiles de débit sélectionnés. Les quantiles optimaux varient en fonction du critère d'évaluation envisagé. Les quantiles de hauts débits sont plus informatifs lorsque l'on s'intéresse en particulier à la simulation des hauts débits. A l'inverse, les quantiles de bas débits fournissent d'avantage d'information lorsque l'on cherche à être performant pour la simulation des bas débits. Enfin, les débits médians sont plus informatifs pour la simulation de tous types de débit. Tout cela semble extrêmement logique.

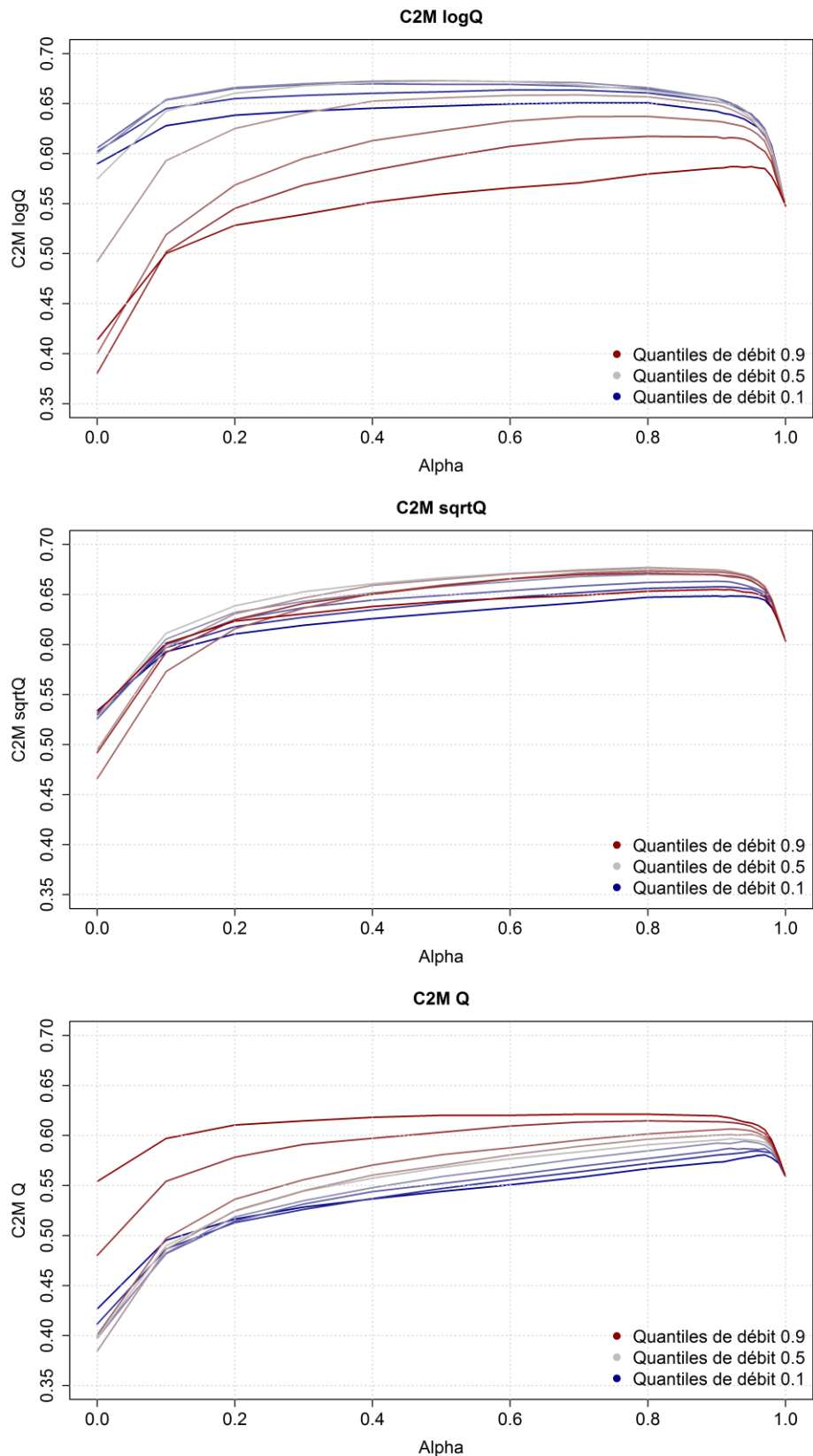


Figure 60. Performances moyennes du modèle en fonction de la valeur α et des quantiles de débit présents dans la chronique de débit observé, obtenues pour l'ensemble des 609 bassins versants. Critères d'évaluation utilisés : du haut vers le bas, le C2M calculé sur le log des débits, le C2M calculé sur la racine carrée des débits et le C2M calculé sur les débits.

6.4. Doit-on privilégier la quantité et la redondance ou la diversité des mesures?

Après cette analyse de la sensibilité du quantile optimum par rapport au critère d'évaluation utilisé et en fonction du paramètre α , nous allons déterminer si le caractère « qualitatif » et redondant des quantiles de débit apporte davantage d'informations que la diversité des N mesures sélectionnées aléatoirement dans la partie 6.2. En effet, dans la sélection de quantiles de débit (SR), un nombre important de débits de valeurs similaires peut être utilisé. Alors que pour la sélection aléatoire (SA), un nombre limité de N mesures de débit est sélectionné, mais la nature des débits sélectionnés peut être très variée, comme en témoigne la Figure 61. De ce fait, vaut-il mieux avoir beaucoup de débits de même valeur ou bien peu de débits variés ?

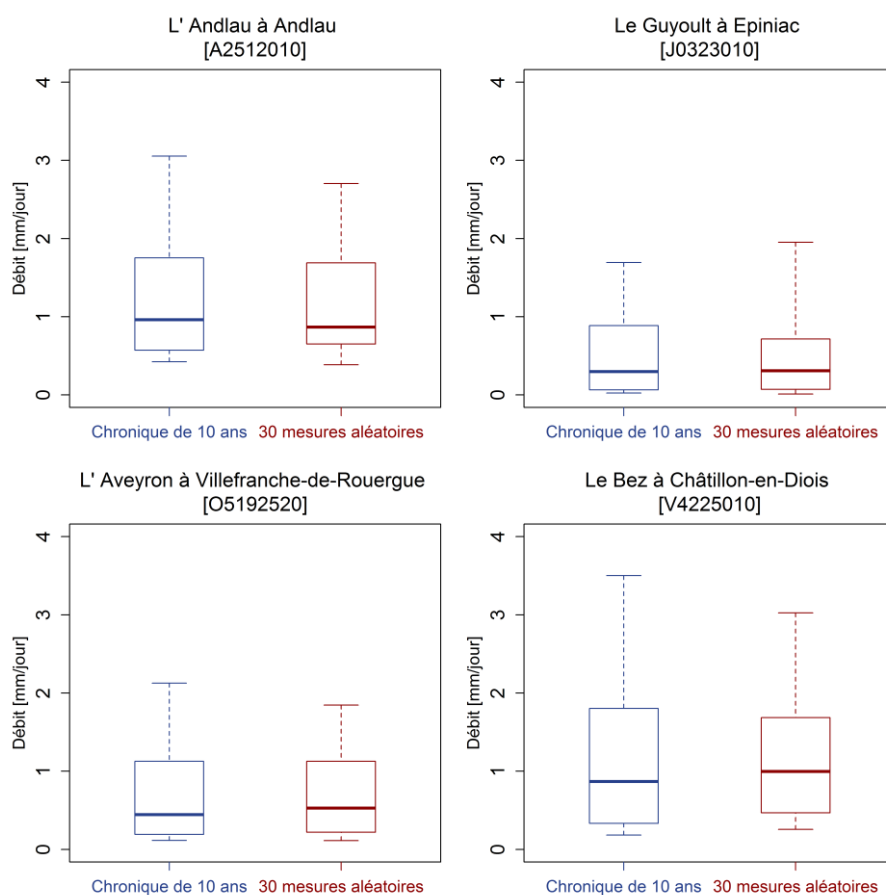


Figure 61. Distribution des débits observés sur la chronique de 10 ans (en bleu) et sur les trente mesures sélectionnées aléatoirement (en rouge) pour quatre bassins versants de notre échantillon.

Les Tableau 7 à Tableau 9 fournissent pour chaque quantile le nombre de mesures de débit disponibles dans la chronique correspondant au quantile donné, ainsi que le C2M maximal obtenu pour chaque quantile (avec la valeur α optimale). Dans ces tableaux, est également indiqué le nombre de mesures aléatoires à partir duquel la sélection de débit aléatoire est préférable à la sélection d'un quantile de débit particulier.

Le Tableau 7 montre que la connaissance des 47 mesures de débit correspondant au quantile 0.4 est équivalente à l'utilisation de 9 mesures de débit sélectionnées aléatoirement.

Par ailleurs, le Tableau 8 montre que l'utilisation des 44 mesures de débit correspondant au quantile 0.6 est équivalente à l'utilisation de 7 mesures de débit sélectionnées au hasard.

Enfin, le Tableau 9 montre que l'utilisation des 23 mesures de débit correspondant au quantile 0.9 est équivalente à l'utilisation de 11 mesures de débit sélectionnées aléatoirement.

Tableau 7. Tableau récapitulatif du nombre de mesures de débit disponibles dans la chronique correspondant au quantile donné, ainsi que le C2M maximal obtenu pour chaque quantile avec la valeur α optimale. Le nombre de mesures aléatoires à partir duquel la sélection de débit aléatoire (SA) est préférable à la sélection d'un quantile spécifique (SR) est également indiqué. En rouge, le quantile donnant les meilleures performances. Résultats pour le critère d'évaluation C2M logQ. Ces valeurs représentent les valeurs médianes de l'ensemble de l'échantillon de 609 bassins versants.

Quantiles	Nombre de mesures du même débit (SR)	C2M logQ maximal	Nombre de mesures de débits pris au hasard (SA)
0.1	33	0.6509	5-6
0.2	41	0.6638	6-7
0.3	45	0.6698	8-9
0.4	47	0.6731	8-9
0.5	46	0.6727	8-9
0.6	44	0.6587	6-7
0.7	40	0.6373	3-4
0.8	33	0.6173	2-3
0.9	23	0.5871	1-2

Tableau 8. Tableau récapitulatif du nombre de mesures de débit disponibles dans la chronique correspondant au quantile donné, ainsi que le C2M maximal obtenu pour chaque quantile avec la valeur α optimale. Le nombre de mesures aléatoires à partir duquel la sélection de débit aléatoire (SA) est préférable à la sélection d'un quantile spécifique (SR) est également indiqué. En rouge, le quantile donnant les meilleures performances. Résultats pour le critère d'évaluation C2M sqrtQ. Ces valeurs représentent les valeurs médianes de l'ensemble de l'échantillon de 609 bassins versants.

Quantiles	Nombre de mesures du même débit (SR)	C2M sqrtQ maximal	Nombre de mesures de débits pris au hasard (SA)
0.1	33	0.6486	1-2
0.2	41	0.6579	2-3
0.3	45	0.6634	3-4
0.4	47	0.6701	4-5
0.5	46	0.6751	5
0.6	44	0.6771	6-7
0.7	40	0.6736	4-5
0.8	33	0.6714	4-5
0.9	23	0.6551	2-3

Tableau 9. Tableau récapitulatif du nombre de mesures de débit disponibles dans la chronique correspondant au quantile donné, ainsi que le C2M maximal obtenu pour chaque quantile avec la valeur α optimale. Le nombre de mesures aléatoires à partir duquel la sélection de débit aléatoire (SA) est préférable à la sélection d'un quantile spécifique (SR) est également indiqué. En rouge, le quantile donnant les meilleures performances. Résultats pour le critère d'évaluation C2M Q. Ces valeurs représentent les valeurs médianes de l'ensemble de l'échantillon de 609 bassins versants.

Quantiles	Nombre de mesures du même débit (SR)	C2M Q maximal	Nombre de mesures de débits pris au hasard (SA)
0.1	33	0.5806	2-3
0.2	41	0.5845	2-3
0.3	45	0.5872	2-3
0.4	47	0.5943	3-4
0.5	46	0.5968	3-4
0.6	44	0.6011	4-5
0.7	40	0.6066	4-5
0.8	33	0.6146	6-7
0.9	23	0.6214	10-11

La Figure 62 présente pour chaque quantile, le rapport du nombre de mesures nécessaires pour chacune des deux méthodes (SR et SA) afin d'obtenir des performances équivalentes. Cette figure traduit la valeur de la redondance de l'information hydrologique contenue dans ces données. Nous pouvons constater que pour le critère valorisant les hauts débits (C2M Q), il faut seulement deux fois plus de données de quantiles de haut débit par rapport aux mesures aléatoires (SA) pour atteindre le même niveau de performances. Alors que pour l'utilisation des bas débits avec ce critère d'évaluation, il en faut au moins dix fois plus. Inversement, pour le critère d'évaluation valorisant les bas débits (C2M logQ), seulement cinq fois plus de données de quantiles de bas débit par rapport au nombre de mesures issues de la SA sont nécessaires pour atteindre des performances équivalentes, alors que ce rapport double quand des quantiles de hauts débits sont utilisés.

Ainsi, la valeur hydrologique des débits répétés diminue très fortement lorsque l'on s'éloigne des quantiles de débit ciblés, traduisant ainsi la baisse de la valeur de l'information redondante contenue dans ces données.

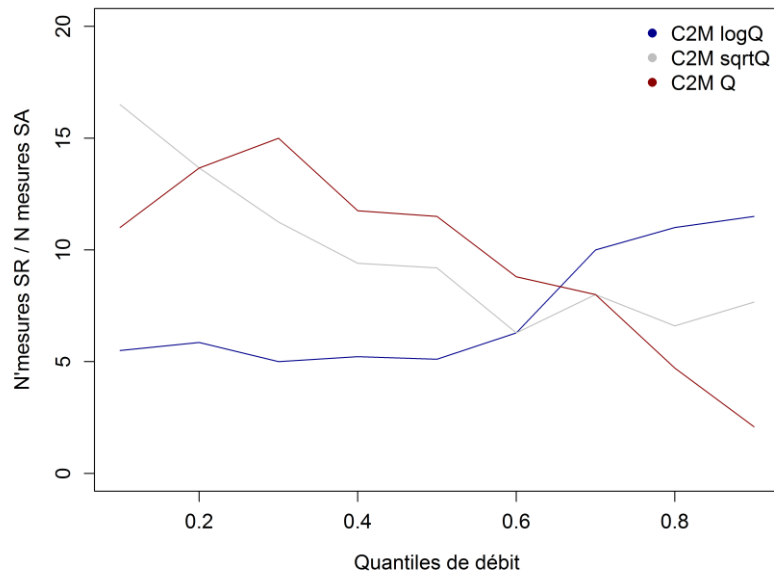


Figure 62. Une mesure de la valeur de la redondance de l'information hydrologique contenue dans un hydrogramme. Rapport du nombre de mesures N' de la sélection redondante (SR) de chaque quantile en fonction du nombre de mesures N de la sélection aléatoire (SA) à partir duquel la sélection aléatoire est meilleure que la sélection redondante.

6.5. Synthèse

Le but de ce chapitre était de présenter une méthode permettant de paramétrer un modèle pluie-débit sur des bassins très peu jaugés, issue des travaux de thèse de *Rojas-Serna* [2005], afin de pouvoir explorer plus avant le domaine des jaugeages ponctuels, qui peut s'avérer pour les gestionnaires de réseau comme un moyen très économique de réduire l'incertitude en situation non jaugée. La méthode combinait une information partielle fournie par les quelques mesures de débit du bassin cible, avec une information régionale, fournie par les paramètres des bassins voisins. Cette étude montre que la méthode proposée peut efficacement exploiter les deux sources d'information, ce qui est particulièrement intéressant quand peu de mesures sont disponibles sur un bassin versant. En effet, même lorsqu'une seule mesure de débit est disponible, des améliorations de performances peuvent être obtenues en comparaison de la méthode de régionalisation utilisée sous sa forme la plus simple.

La deuxième partie de ce chapitre s'est portée sur l'évaluation des gammes de débit les plus informatives selon l'objectif hydrologique envisagé (estimation des bas, moyens ou hauts débits). Pour cela, nous nous sommes placés dans le cas particulier d'un bassin versant disposant d'une chronique de hauteur d'eau mais sur lequel aucune courbe de tarage n'a été établie. Ainsi, si un jaugeage de débit était effectué sur ce bassin, nous obtiendrions un débit mesuré pour une hauteur d'eau donnée. En supposant que la relation hauteur-débit est restée stable et univoque dans le temps, à chaque fois que cette hauteur d'eau est rencontrée dans la chronique, on peut lui associer le débit mesuré lors de l'unique jaugeage effectué. Les résultats ont montré que le choix optimal de la gamme de débit dépendait essentiellement de l'objectif hydrologique. En effet, pour la simulation de bas débits, il était préférable de sélectionner des quantiles de bas débit ; pour la simulation de hauts débits, des quantiles de haut débit et pour la simulation de tous types de débits, des débits médians.

Dans ce chapitre, nous avons également essayé de déterminer s'il fallait « mettre tous ses œufs dans le même panier » lors de la « récolte » des jaugeages, en comparant les deux types d'approche présentés précédemment. Nous souhaitions répondre à la question : Vaut-il mieux avoir beaucoup de débits de même valeur ou bien peu de débits de natures variées ? Les résultats ont montré que la sélection aléatoire des débits était préférable lorsque plus d'une dizaine de mesures aléatoires de débit étaient disponibles sur le bassin versant. En dessous de dix mesures disponibles, si une chronique de hauteur d'eau est disponible à l'exutoire du bassin, mieux vaut privilégier l'unique mesure de débit ponctuelle associée à une valeur de hauteur d'eau, afin d'obtenir une sélection redondante de mesures. La redondance de l'information fournie par l'unique jaugeage ciblé pour un objectif hydrologique bien précis, sera alors préférable.

Il resterait cependant à imaginer la stratégie d'échantillonnage la plus efficace possible pour la mise en pratique de la collecte d'informations en contexte opérationnel. La question d'une stratégie évolutive, qui serait fonction de l'information déjà acquise se poserait alors : sachant que trois crues ont déjà été observées, quelle gamme de débit serait-il le plus bénéfique de viser ?

Partie II

Impact de la densité du réseau pluviométrique sur les calculs hydrologiques

Chapitre 7.

Impact de la densité du réseau pluviométrique en hydrologie : état de l'art

7.1. Introduction

La pluie est un élément essentiel pour les études hydrologiques. Cependant, il s'agit d'une grandeur très variable spatialement et temporellement, même à l'échelle d'un bassin versant [Krajewski et al., 2003]. Dans un contexte de modélisation pluie-débit, la variabilité spatiale de la pluie peut ainsi avoir un impact important sur l'estimation du débit et cette question a déjà été largement débattue dans la littérature scientifique [Arnaud et al., 2011; Bell et Moore, 2000; Faures et al., 1995; Lobligeois et al., 2014; Obléd et al., 1994]. Il est donc primordial de pouvoir estimer la pluie de manière fiable afin de limiter les incertitudes dans son utilisation au sein d'applications hydrologiques.

Toutefois, cette pluie de bassin utilisée dans les modèles hydrologiques, est majoritairement interpolée à partir de mesures effectuées au niveau de postes ponctuels situés au sol, appelés pluviomètres. On peut facilement imaginer que la densité et la configuration du réseau de ces postes de mesure va jouer un rôle important sur la précision de l'estimation de la pluie de bassin et la façon dont la variabilité spatiale de la pluie sera plus ou moins bien représentée.

Dans ce chapitre, nous dresserons un état de l'art de l'impact de la densité spatiale du réseau pluviométrique en hydrologie et de l'impact de la longueur des séries disponibles.

7.2. Impact de la densité spatiale du réseau pluviométrique en hydrologie

La question de la dépendance des résultats des calculs hydrologiques à la densité spatiale des réseaux pluviométriques n'est pas nouvelle. En 1972, l'Organisation Météorologique Mondiale [OMM, 1972] publiait un recueil de techniques et de principes concernant la planification des réseaux de mesure de la pluie. Ce rapport propose notamment des normes concernant leurs densités selon les régions et recommande également un nombre minimal de pluviomètres pour chaque station de jaugeage des rivières à des fins de prévision. Pour citer quelques chiffres, des densités de réseaux variant entre 1 à 100 stations pour 10 000 km² selon les régions, sont recommandées, ainsi qu'un minimum de deux stations pluviométriques par station hydrométrique.

En 1981, Schaake [1981] recommande également un nombre minimal de stations pluviométrique par bassin versant en fonction de la superficie du bassin pour des applications de prévisions de débit (Eq. 25). Cette équation a été développée à partir d'un réseau pluviométrique très dense (45km² par station). Il précise cependant qu'il est raisonnable d'appliquer cette équation pour des bassins dont la superficie varie entre 200 et 20 000 km² et pour une modélisation hydrologique avec un pas de temps équivalent au quart du temps de réponse du bassin [Schaake et al., 2000].

$$N = 0.6 \times A^{0.3} \quad \text{Eq. 25}$$

Toutefois, l'objectif des études liées aux densités spatiales des stations pluviométriques était, dans un premier temps, d'évaluer l'impact de la densité pour l'estimation des champs pluvieux et des pluies moyennes de bassin. Ce n'est que plus tard, que les auteurs ont commencé à s'intéresser à cet impact sur les applications hydrologiques.

7.2.1 Impact sur la qualité des estimations de la pluie

A partir des années 1970, un certain nombre d'études ont été réalisées concernant la densité spatiale des réseaux de pluie, mais sans réelle application hydrologique. Ces études portaient

davantage sur l'évaluation de la capacité des réseaux à représenter plus ou moins bien un champ pluvieux, et non sur la sensibilité des calculs hydrologiques aux variations de densité des réseaux :

- *Huff [1970]* a utilisé des données issues de deux réseaux de mesure de la pluie pour évaluer l'erreur d'échantillonnage dans la pluie moyenne de bassin pour des événements orageux. Les résultats ont indiqué que les erreurs d'échantillonnage diminuaient quand le réseau était plus dense.
- L'étude de *Rodriguez-Iturbe et Mejia [1974]* ainsi que celle de *Bras et Colon [1978]* ont révélé que, lors de l'établissement d'un réseau de stations de mesures de pluie, il était important de tenir compte entre autres de la géométrie du réseau et du nombre de stations afin de pouvoir calculer la pluie moyenne de bassin à long terme.
- *Seed et Austin [1990]* ont utilisé une grande quantité de données radar pour déterminer l'erreur de mesure pour les pluies journalières et mensuelles d'une grande surface, liée à l'organisation et la densité du réseau. Ils ont alors trouvé que le réseau régulier était un peu meilleur que le réseau de pluie sélectionné au hasard. La différence devient cependant plus prononcée pour des réseaux très dispersés.
- *Tsintikidis et al. [2002]* appliquent des méthodes statistiques pour quantifier l'incertitude associée à l'estimation des précipitations pour un réseau pluviométrique existant et essaient d'identifier les possibles sites de stations supplémentaires pour réduire les erreurs d'interpolation de la pluie.
- L'étude de *Hrachowitz et Weiler [2011]* a porté sur l'influence de la densité du réseau de mesures de pluie sur le calcul de la pluie moyenne de bassin. Leurs résultats ont montré que l'erreur absolue de l'estimation de la précipitation moyenne journalière du bassin (en comparaison avec la meilleure estimation de la pluie moyenne de bassin disponible calculée avec la totalité des stations) décroît exponentiellement avec l'augmentation du nombre de stations de pluie.

En France, une étude concernant l'impact de la densité du réseau pluviométrique sur l'estimation de la pluie de bassin, a déjà été menée sur un bassin versant située dans les Cévennes, région où les précipitations peuvent être très intenses [*Lebel et al., 1987*]. Les auteurs ont testé trois méthodes d'estimation de la pluie de bassin en faisant varier la densité du réseau de mesure. L'étude montre que la précision de l'interpolation de la pluie de bassin par le krigeage diminue lorsque la densité du réseau diminue également.

D'autre part, *Adhikary et al. [2014]* ont développé une méthode permettant d'obtenir un réseau pluviométrique optimal pour l'estimation de la pluie de bassin. Ils ont montré qu'une réduction significative de l'erreur d'estimation de la pluie de bassin pouvait être obtenue avec un positionnement spatial optimal d'une nouvelle station ou avec un repositionnement d'une station redondante.

Dans un autre registre, *Hendrick et Comer [1970]* ont mis en place une fonction de corrélation permettant de déterminer la densité et la configuration du réseau pluviométrique optimales dans le but d'avoir une estimation précise de la pluie journalière de bassin. Pour cela, ils déterminent les champs de corrélation de la pluie de chaque station pluviométrique en fonction de la distance et la direction inter-bassins ainsi que de la pluie minimale journalière. Ils ont également montré que la

précision de l'estimation de la pluie était augmentée lorsque le nombre de stations augmentait aussi, jusqu'à atteindre un pallier à partir de 5 stations, au-delà duquel le gain de précision était plus limité.

Les études ci-dessus ont ainsi montré que les densités spatiales des réseaux de mesure de pluie ont un impact sur la représentation des champs pluvieux et sur les calculs de pluie moyenne de bassin. Mais retrouve-t-on ce même impact sur la qualité des estimations des débits ?

7.2.2 Impact sur la qualité des estimations de débit

Impact de la densité du réseau pluviométrique en modélisation hydrologique

Dans un cadre de modélisation hydrologique, des auteurs ont montré que les erreurs liées à l'estimation des pluies peuvent être partiellement compensées lors du calage des paramètres du modèle ou de son utilisation [Andréassian *et al.*, 2001; Caracciolo *et al.*, 2014] ou que les bassins versants pouvaient agir comme des filtres passe-bas, permettant d'atténuer les variabilités de la pluie [Bras *et Rodriguez-Iturbe*, 1976; Oudin *et al.*, 2004]. Ainsi, une pluie « mal » représentée (mais de façon prévisible) pourrait permettre d'obtenir dans certains cas de bonnes modélisations de débit.

Les études suivantes se sont penchées sur l'impact de la densité du réseau pluviométrique sur les performances des modèles hydrologiques :

- Bras *et Rodriguez-Iturbe* [1976] ont proposé une méthode visant à quantifier la qualité des réseaux d'échantillonnage de la pluie pour leur utilisation par un modèle pluie-débit. Les résultats indiquent une forte décroissance de l'erreur quadratique moyenne quand toutes les stations disponibles sont utilisées.
- L'étude de Duncan *et al.* [1993] avait pour objectif d'évaluer, en utilisant des cartes de pluie construites à partir de données radar, l'impact de l'augmentation de la densité des stations sur la prédictibilité des hydrogrammes de sortie en utilisant un modèle hydrologique simple avec des données distribuées. Ils ont montré que la densité des stations a un très fort effet sur la précision de l'estimation des paramètres mais également sur la simulation des pics de crue (décroissance des erreurs selon une loi puissance). Il apparaît également qu'une densité plus forte est nécessaire pour estimer précisément le pic de crue, que pour estimer le débit total et la pluie totale. Ce point est important dans une perspective de prévision de crues notamment.
- Michaud *et Sorooshian* [1994] ont évalué l'effet des erreurs d'échantillonnage de la pluie sur les simulations hydrologiques distribuées, conduites lors d'événements orageux locaux. Lors de cette étude, le réseau de pluviomètres le moins dense produit les plus fortes erreurs dans la simulation des pics de crue.
- Obléd *et al.* [1994] ont également montré que l'utilisation d'un réseau plus dense était préférable pour la simulation d'événements de crue sur le bassin expérimental français du Réal Collobrier, à l'aide de la version distribuée du modèle TOPMODEL.
- A travers une analyse à l'échelle de l'événement et d'un petit bassin versant de 5 ha, Faures *et al.* [1995] ont montré que l'utilisation d'une pluie uniforme sur l'ensemble du bassin estimée à l'aide d'une seule station conduisait à de fortes incertitudes concernant l'estimation du débit.

- L'étude d'*Andréassian et al.* [2001] a eu pour but d'analyser la sensibilité de trois modèles hydrologiques (GR3J, TOPMO et IHACRES) à la connaissance imparfaite des données de pluie utilisées en entrée des modèles. Les résultats ont indiqué que l'efficacité des modèles augmente lorsqu'une meilleure description de la pluie en entrée est utilisée. Toutefois, les modèles sont tout de même capables de s'adapter à des erreurs systématiques et d'obtenir de bons résultats lorsqu'une estimation imparfaite de la pluie est utilisée, en adaptant leurs paramètres lors du calage.
- L'étude de *St-Hilaire et al.* [2003] a révélé que l'impact de la densité du réseau est également important pour la simulation du débit journalier. Les simulations ont été généralement améliorées quand la moyenne était calculée à partir du réseau le plus dense. De la même façon, les pics de crue étaient généralement mieux simulés quand le réseau le plus dense était utilisé pour calculer les précipitations moyennes journalières.
- *Segond et al.* [2007] ont étudié la relation entre la pluie spatiale et la prévision des débits de pointe. Pour cela, ils ont testé trois niveaux différents de densité (1 station pour tout le bassin, puis 7 stations (1 par sous-bassin) et enfin des données radar). Les résultats indiquent que le réseau le moins dense fournit les moins bonnes performances en prévision.
- *Meselhe et al.* [2009] ont étudié l'impact de l'échantillonnage temporel et spatial de la pluie sur les prévisions de débit. Pour cela, ils ont testé différentes densités de réseaux pluviométriques (de 1 à 30 stations pour le bassin étudié). Ces résultats ont révélé que les performances du modèle se détériorent quand la densité des stations de pluie diminue. Toutefois, pour cette taille de bassin (21,4 km²), l'augmentation de la densité de stations de pluie de 1 à 2 donne l'amélioration la plus significative pour les deux modèles hydrologiques utilisés (Figure 63).

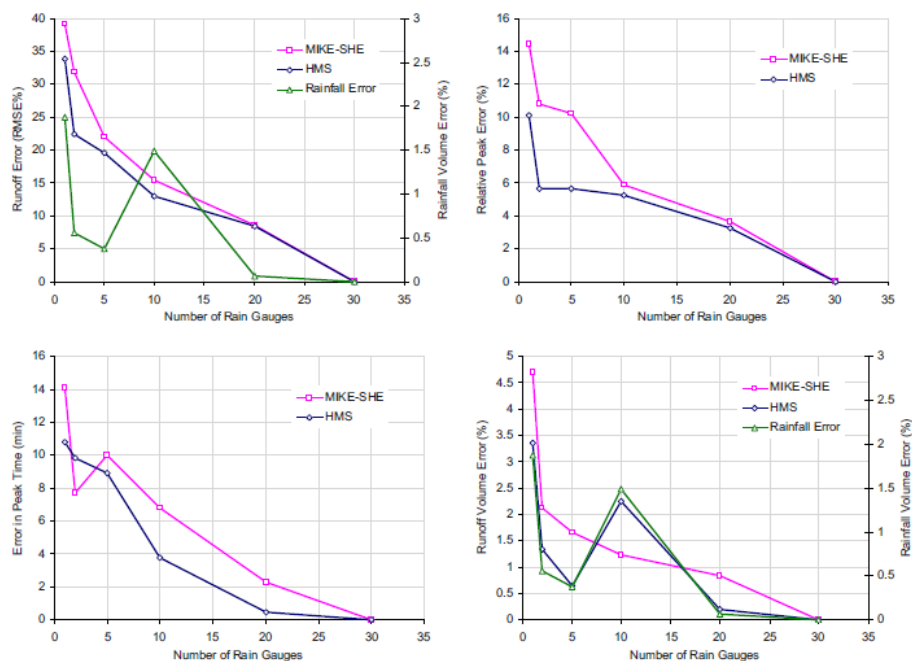


Figure 63. Erreurs calculées sur la pluie et le débit, mais aussi erreurs relatives sur le pic de crue en fonction du nombre de pluviomètres utilisées pour calculer la pluie de bassin (d'après *Meselhe et al.* [2009]).

Ces études révèlent que plus la densité de stations pluviométriques est forte, meilleures sont les simulations, et que les performances sont plus ou moins croissantes avec la densité du réseau. D'autres auteurs ont cependant montré que cette augmentation des performances des modèles hydrologiques avec l'augmentation de la densité des réseaux atteint généralement un plateau : à partir d'un certain nombre de postes, les performances ne sont plus améliorées. Ceci indique que l'ajout de stations n'apporte pas d'information supplémentaire, ni de gain dans les performances du modèle :

- *Jacoby et Dawdy* [1973] ont entrepris une étude de simulation pluie-débit, afin d'évaluer le gain de performance en simulation pour un bassin de montagne lorsque davantage de données de pluie sont utilisées pour le calage du modèle. D'un point de vue spatial, les résultats ont indiqué que l'ajout d'une troisième station pluviométrique n'apportait pas de gain en qualité de simulation.
- *Anctil et al.* [2006] ont montré que certaines combinaisons de stations donnent de meilleures prévisions de débit que lorsque toutes les stations disponibles sont utilisées pour estimer la pluie moyenne du bassin (12 sur 23 stations disponibles). Toutefois, les auteurs précisent que ces résultats ne doivent pas être interprétés comme une incitation à retirer les autres stations de pluie du bassin. Cela veut seulement dire que, du point de vue de l'optimisation, et à travers le filtre du modèle global pluie-débit, un faible nombre de stations de pluie spécifique peut être bénéfique en prenant en considération leur utilité pour la prévision des débits.
- *Bardossy et Das* [2008] ont montré que les performances du modèle HBV (modèle conceptuel pluie-débit semi-distribué) étaient radicalement plus mauvaises lorsque le nombre de stations pluviométriques était réduit de façon excessive, mais que l'ensemble des performances n'était pas significativement amélioré à partir d'un nombre de stations donné (Figure 64). Ils indiquent également que lorsqu'un réseau peu dense est utilisé pour le calage du modèle, de bonnes performances sont retrouvées en validation avec un réseau plus dense en validation. Dans le cas inverse (données denses en calage et dispersées en validation), les performances du modèle sont très nettement dégradées. D'un point de vue opérationnel, ceci souligne l'intérêt de caler le modèle avec une information similaire à celle qui sera utilisée ensuite comme entrée du modèle : on peut citer à ce sujet l'étude très complète de *Oudin et al.* [2006a] qui ont décrit l'impact de différents types d'erreurs (aléatoires et systématiques) dans les données de pluie en entrée d'un modèle.

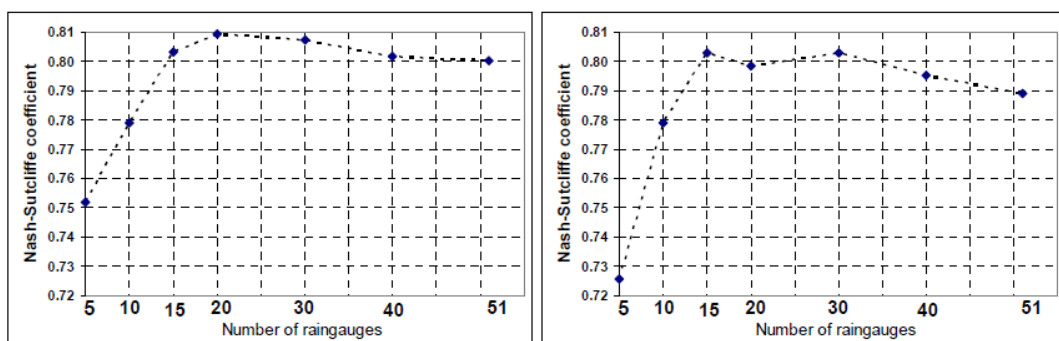


Figure 64. Performances du modèle pour différentes densités de réseau pluviométrique en période de calage (à gauche) et en période de validation (à droite) (d'après *Bardossy et Das* [2008]).

- L'étude de *Xu et al.* [2013] a eu pour but d'étudier l'influence de différentes densités de pluviomètres, la distribution du réseau et la localisation des stations pluviométriques sur les performances d'un modèle hydrologique obtenues sur un très grand bassin chinois (94 660 km²). Des scénarios de densités différentes ont été construits : 5%, 10%, 20%, 30%, 50% et 70% des 181 pluviomètres totaux. Pour chaque scénario, 100 configurations différentes de réseau ont été réalisées. Les résultats ont montré que les erreurs du modèle diminuaient graduellement avec l'augmentation du nombre de stations pluviométriques utilisées jusqu'à atteindre un palier à partir duquel les performances du modèle ne montrent pas d'amélioration évidente (à partir de 93 postes) (voir Figure 65). Ils précisent également que de meilleures performances peuvent être obtenues avec peu de stations lorsqu'une configuration spatiale optimale du réseau est utilisée.

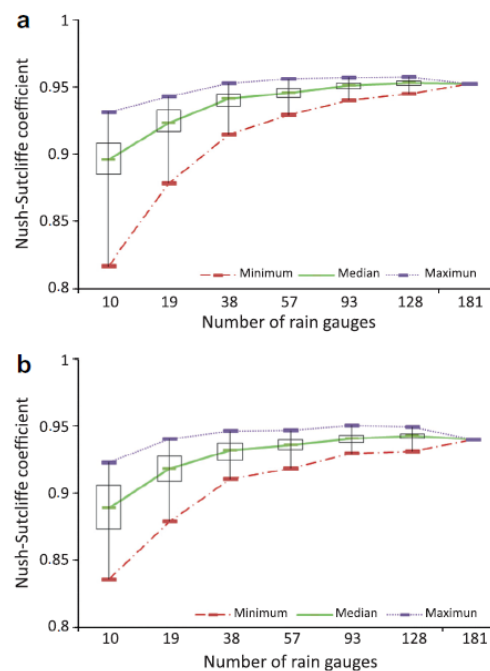


Figure 65. Performances du modèle hydrologique en fonction du nombre de pluviomètres utilisés pour estimer la pluie de bassin. a- performances en calage et b- performances en validation (d'après *Xu et al.* [2013]).

Impact de la localisation des stations de mesure de la pluie en hydrologie

D'autres études ont également mis en place des stratégies d'échantillonnage des réseaux pluviométriques afin de limiter les erreurs dans l'estimation des débits.

Sugawara [1993] propose une méthode permettant de déterminer les poids accordés à chaque station de pluie non en fonction de leurs positions géométriques mais en fonction des conditions météorologiques du bassin. Le choix des stations devraient ainsi être fonction de leur utilité pour l'estimation des débits. Etant donné que la pluie réelle de bassin ne peut être mesurée, il propose de déterminer le poids des stations de pluie à partir des débits observés, en comparant les débits observés et les débits dérivés des pluies de chaque station.

Volkman et al. [2010] ont quant à eux développé et évalué une stratégie multicritères pour identifier les meilleures localisations pour l'installation de stations de pluie, avec un intérêt particulier pour les régions semi-arides montagneuses. Pour cela, 1000 configurations de réseau de pluie ont été sélectionnées au hasard pour chaque classe de densité (de 1 à 4 stations), afin de simuler les effets des observations obtenues en attribuant aléatoirement les localisations des stations dans le réseau. Les résultats ont révélé une augmentation de la précision des simulations lorsque le réseau était plus dense mais que cette amélioration commençait à se stabiliser à partir de plus de trois stations dans le réseau. Dans cette étude, des lames d'eau de pluie sont estimées à l'aide de radar. Les données interpolées avec la technique du krigeage avec dérive externe sont supposées être les vraies pluies observées (on parle alors d'approche synthétique). Chaque centre de maille est ensuite utilisé pour définir les positions potentielles des stations. Toutefois, dans ce type d'approche le choix de la pluie de référence (estimée comme étant la pluie parfaite) conditionne le choix de la solution à explorer.

L'étude de *Dong et al.* [2005] avait également pour objectif de déterminer à partir d'analyses statistiques, le nombre et la localisation des pluviomètres appropriés au sein d'un bassin versant afin de modéliser le débit. Pour le bassin étudié, ces auteurs ont montré que les performances du modèle augmentent de manière hyperbolique, jusqu'à une valeur asymptotique à partir de cinq pluviomètres (qui est donc considérée comme étant le nombre approprié de pluviomètres pour ce bassin). De plus, les localisations géographiques des pluviomètres en fonction des performances plus ou moins bonnes du modèle ont été identifiées, montrant ainsi une forte dépendance par rapport aux caractéristiques géographiques et climatiques locales.

Arsenault et Brissette [2013] ont aussi proposé une méthode exploratoire sur un bassin Canadien de 8000 km², afin de déterminer la densité optimale du réseau de stations météo dans le cadre d'une modélisation hydrologique de long-terme. Les stations météorologiques utilisées dans cette étude sont des stations virtuelles issues du modèle climatique régional canadien (CRCM15). Ils ont montré que si les stations météorologiques sont positionnées de façon optimale, un faible nombre de stations peut suffire pour produire des performances aussi bonnes, voire meilleures, qu'avec l'utilisation du réseau complet. Toutefois, il est important de noter que le risque de produire un réseau de faible qualité hydrologique est plus élevé lorsque peu de stations sont utilisées. Ainsi, ils concluent qu'un réseau dense est préférable si les stations sont sélectionnées aléatoirement.

Enfin, *Caracciolo et al.* [2014] ont montré que la configuration optimale du réseau dans les estimations de la pluie et du débit était surtout influencée par la distribution spatio-temporelle des événements de pluie, de la distribution du type de sol et de la variable de sortie étudiée. Ces conclusions ont également été retrouvées dans l'étude de *Lopes* [1996] qui montre que la densité et l'arrangement spatial du réseau pluviométrique ainsi que les caractéristiques spatio-temporelles des événements de pluie, influencent fortement les prévisions de débit.

7.3. Impact de la longueur des enregistrements des données de pluie en hydrologie

L'impact de la longueur des chroniques de pluie ne peut être testé dans le cadre d'une modélisation hydrologique pluie-débit, puisque les modèles utilisant les données de pluie comme entrée, ont besoin de la pluie pour pouvoir calculer le débit. En l'absence de précipitation, aucun débit ne peut

être simulé par le modèle. On ne peut alors traiter de manière indépendante les deux réseaux pour la dimension temporelle. L'impact de la longueur des chroniques de pluie a été indirectement traité dans les études visant à évaluer l'impact de la longueur des chroniques de débit (voir partie I).

Ne pouvant évaluer l'impact de la longueur des chroniques de pluie pour des applications hydrologiques, nous nous sommes tout de même intéressés à son effet sur des calculs à caractérisation pluviale, comme par exemple le calcul de la pluie moyenne de bassin. Une première étude, celle de *Rodriguez-Iturbe et Mejia* [1974], a présenté l'impact de la composition du réseau sur l'estimation de la pluie moyenne de bassin à long terme. Les résultats ont amené à la conclusion que la précision de cette estimation était fonction du nombre de stations, de la surface de la zone étudiée, et de la longueur temporelle des enregistrements de précipitation. Les auteurs précisent également que pour l'estimation de la pluie moyenne à long terme, le facteur dominant est le temps depuis lequel le réseau est en fonctionnement.

Par la suite, *Bras et Colon* [1978] ont développé une méthode destinée à établir un compromis entre le nombre et la localisation des observations et la longueur des enregistrements de pluie afin d'estimer la précision de l'estimation de la pluie moyenne temporelle d'un bassin. Leurs résultats indiquent que l'estimation de l'erreur quadratique moyenne décroît quand le nombre de stations et la durée d'échantillonnage augmentent.

La longueur des chroniques de pluie a alors un impact sur l'estimation de la pluie moyenne de bassin à long terme. Il semblerait que cette estimation soit également impactée par le nombre et la localisation des stations pluviométriques.

7.4. Synthèse

Nous avons pu analyser un grand nombre d'articles traitant de l'étude de l'impact de la densité spatiale du réseau pluviométrique sur les modélisations hydrologiques. Mais malgré cette apparente abondance de travaux, il faut remarquer que la plus grande part de ceux-ci est fortement biaisée par l'approche synthétique qui est utilisée : une pluie (distribuée) est donnée comme "vraie", et une simulation de référence obtenue avec toutes les données disponibles et un jeu de paramètres du modèle sert de débit de référence, en lieu et place du débit mesuré. Dans ces conditions, il est mathématiquement impossible de faire mieux que la pluie distribuée, aussi on ne peut s'empêcher de penser que le raisonnement des modélisateurs est circulaire. Andréassian et al. (2001) mentionnaient déjà ce problème : "the use of simulation-derived streamflow series may lead to results that only confirm the very hypotheses on which the models are built."

Malgré ces réserves, on peut résumer les conclusions présentes dans la littérature hydrologique. D'une part, un réseau dense de pluviomètres permet d'obtenir de meilleures performances dans l'estimation des pluies moyennes de bassin et en modélisation et prévision des débits. Cependant, le réseau le plus dense ne fournira pas automatiquement les meilleurs résultats, et un seuil de performances peut être atteint à partir d'un certain nombre de stations présentes dans le réseau. D'autre part, il semble que la localisation géographique des stations au sein du réseau influence également les performances des modèles hydrologiques. Mais il faut rester vigilant, car ce second point est essentiellement issu de simulations synthétiques.

Chapitre 8.

Présentation de la base de données de la partie II

8.1. Introduction

Dans ce chapitre, nous présenterons les données hydro-climatiques utilisées, ainsi que le nouvel échantillon de 30 bassins versants retenu pour les travaux de la deuxième partie de cette thèse portant sur l'évaluation de l'impact de la densité du réseau pluviométrique sur les calculs hydrologiques.

Nous ferons également un rappel des différents calculs hydrologiques considérés dans cette partie.

8.2. Les données hydro-climatiques

8.2.1 Origine des données pluviométriques

Afin d'évaluer l'impact de la densité du réseau pluviométrique en hydrologie, nous avons utilisé des données journalières de pluie issues de chroniques de pluie mesurées au niveau de pluviomètres du réseau Météo-France. Dans la base de données à laquelle nous avons eu accès, des chroniques de 10 ans de données journalières de pluie sont disponibles pour 5238 stations de mesure entre 1997 et 2006 (Figure 66).

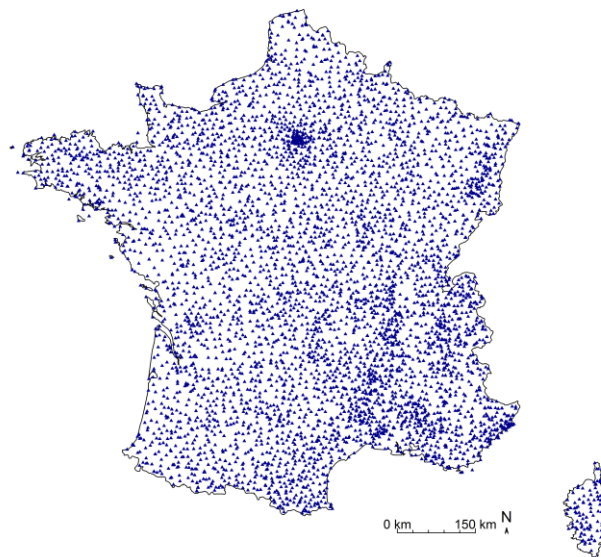


Figure 66. Localisation des 5238 stations pluviométriques journalières du réseau Météo-France utilisées dans la deuxième partie de la thèse.

8.2.2 Interpolation de la pluie de bassin

Au cours de cette thèse, nous avons fait le choix d'effectuer l'interpolation de la pluie de bassin selon la méthode de pondération par l'inverse de la distance (à la puissance 2). Nous sommes conscients que d'autres méthodes d'interpolation plus élaborées, telles que le krigeage, existent. Toutefois, nous avons dans un premier temps décidé d'effectuer l'interpolation de la pluie de bassin par une méthode plus simple, et plus rapide à mettre en œuvre. De plus, certaines études ont montré que les méthodes d'interpolation dites déterministes comme celle de pondération par l'inverse de la distance, donnaient des résultats similaires ou meilleurs que les méthodes géostatistiques en termes de modélisation hydrologique [Ly *et al.*, 2011; 2013; Ruelland *et al.*, 2008]. Otieno *et al.* [2014]

concluent également que malgré son plus haut niveau de complexité, l'utilisation du krigeage ne montre pas une plus grande capacité à estimer la pluie moyenne mensuelle que la méthode de pondération par l'inverse de la distance. *Dirks et al.* [1998] considèrent aussi que le choix du krigeage, méthode plus complexe et chronophage, ne semble pas être justifié pour l'estimation de pluie journalière, mensuelle ou annuelle lorsque le réseau pluviométrique est dense.

Afin d'appliquer la méthode de pondération par l'inverse de la distance comme méthode d'interpolation de la pluie de bassin, celui-ci est découpé selon une grille de résolution spatiale 1km x 1km. Une estimation de la pluie est alors obtenue pour chaque maille du bassin versant. La pluie de bassin est ensuite estimée en sommant l'ensemble des pluies calculées sur chacune des mailles du bassin à chaque pas de temps (ici journalier).

De plus, afin de mieux contraindre notre interpolation, nous avons ajouté une deuxième pondération, la pluie mensuelle issue de la réanalyse COMEPHORE de Météo-France qui combine données pluviométriques horaires et journalières et données radar (on utilise ainsi la connaissance de la variabilité spatiale des pluies issue du radar à notre interpolation).

Les équations suivantes décrivent les différentes étapes de l'interpolation de la pluie de bassin de cette étude :

$$P_{bassin} = \sum_{j=1}^M P_{maille}^j$$

$$\text{avec } P_{maille}^j = \frac{\sum_{i=1}^N P_{poste}^i \times \frac{1}{d^2} \times \frac{Pm_{maille}^j}{Pm_{poste}^i}}{\sum_{i=1}^N \frac{1}{d^2} \times \frac{Pm_{maille}^j}{Pm_{poste}^i}}$$

avec P_{maille} : pluie calculée pour chaque maille de la grille (résolution spatiale de 1km x 1km) du bassin versant au pas de temps journalier ; P_{poste}^i : pluie du poste i au pas de temps journalier ; d : distance séparant la maille du poste i ; Pm_{maille}^j : pluie moyenne mensuelle issue de la réanalyse radar pour la maille j ; Pm_{poste}^i : pluie moyenne mensuelle du poste.

8.2.3 Les données de débit et d'évapotranspiration potentielle

Les données journalières d'évapotranspiration potentielle et de débit nécessaires pour la modélisation de débit avec le modèle GR4J sont les mêmes que celles utilisées en première partie de cette thèse (voir chapitre 3).

8.3. La sélection d'un échantillon de bassins versants

Pour évaluer l'impact de la densité du réseau pluviométrique sur les performances des calculs hydrologiques, une nouvelle base de données a été mise en place. En effet, pour effectuer ces tests, il nous a fallu créer une nouvelle pluie de bassin à partir de pluviomètres, suffisamment nombreux pour que nous puissions faire varier leur densité.

8.3.1 Critères de sélection

Un nouvel échantillon de bassins versants a été sélectionné parmi la première liste des 609 bassins de la partie I sur la base de plusieurs critères.

Dans un premier temps, afin d'éviter de se fermer certains axes de recherche, notamment concernant l'utilisation ou non de l'information radar, nous souhaitions avoir des bassins qui étaient vu dans leur intégralité par les radars de Météo-France. Nous avons donc écarté les bassins montagneux, et plus particulièrement les bassins dont la fraction solide des précipitations était supérieure à 10% (calculée à partir de données issues de la base de données SAFRAN).

D'autre part, dans le but de pouvoir effectuer des tests concernant l'impact de la densité du réseau pluviométrique en hydrologie, la présence de bassins versants avec un réseau de pluviomètres relativement dense était nécessaire. Pour cela, nous avons sélectionné des bassins versants disposant au minimum de 7 pluviomètres utiles au calcul des polygones de Thiessen [Thiessen, 1911]. De plus, ces pluviomètres devaient présenter moins de 10% de lacunes par mois dans leur chronique de pluie.

8.3.2 Sélection de l'échantillon

A l'issue de ce tri, un nouvel échantillon de 30 bassins a ainsi été sélectionné pour cette deuxième partie d'étude. La Figure 67 indique la localisation des bassins versants sélectionnés ainsi que leurs pluviomètres associés et considérés pour le calcul de la pluie de bassin.

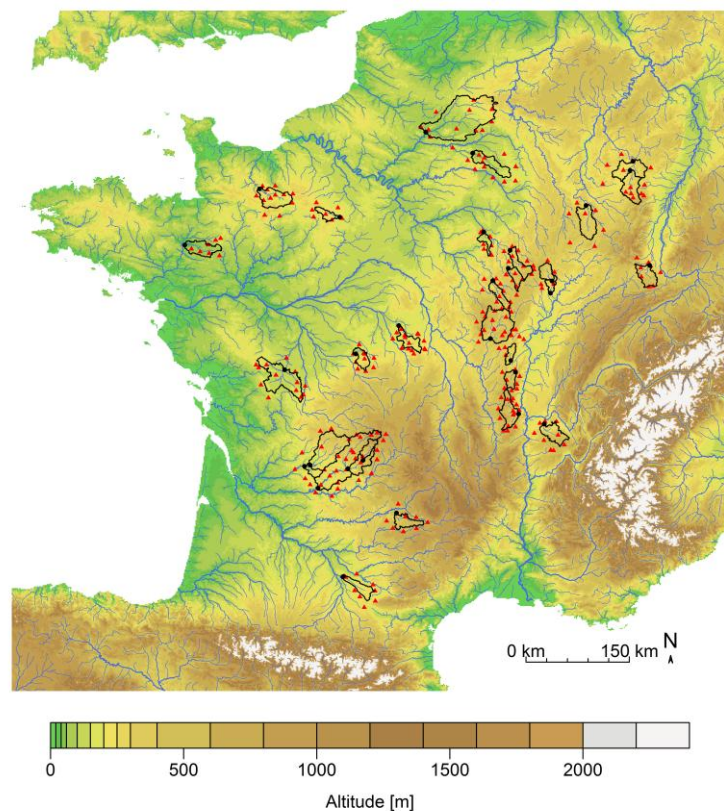


Figure 67. Localisation des 30 bassins versants sélectionnés pour la seconde partie des travaux de cette thèse. En noir, le contour et l'exutoire des bassins versants et en triangle rouge la localisation des pluviomètres utilisés.

La Figure 68 et le Tableau 10 présentent les principales caractéristiques hydro-climatiques et morphologiques de l'échantillon de bassins versants. La liste des bassins versants utilisés pour cette partie, ainsi que leurs caractéristiques hydro-climatiques et morphologiques sont disponibles en annexe C. On peut également trouver en annexe D un zoom du réseau pluviométrique associé à chaque bassin versant.

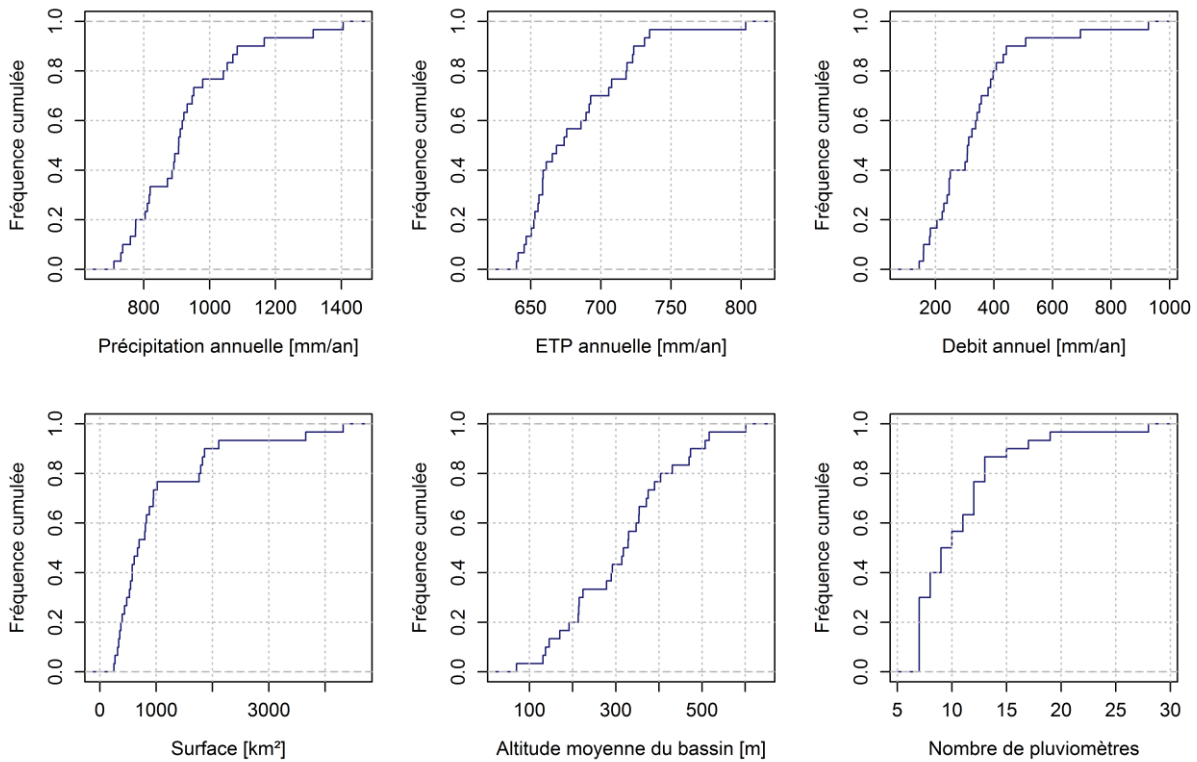


Figure 68. Principales caractéristiques de l'échantillon de bassins versants français sélectionnés dans le cadre de cette deuxième partie de thèse.

Tableau 10. Principales caractéristiques de l'échantillon de bassins versants français sélectionnés dans le cadre de cette deuxième partie de thèse.

Quantiles	0.05	0.25	0.5	0.75	0.95
Surface (km ²)	291	448	682	1003	2961
Altitude moyenne du bassin (m)	134	214	323	386	512
Nombre de pluviomètres	7	7	10	12	18
Surface pluviométrique (km ²)	137	208	295	409	800
Précipitation moyenne annuelle (mm/an)	733	812	906	972	1248
ETP moyenne annuelle (mm/an)	643	655	671	707	733
Débit moyen annuel (mm/an)	159	232	312	386	611

Cet échantillon présente une diversité entre bassins plus limitée que l'échantillon de bassins de la première partie. Nous retrouvons dans cette base de données des bassins plutôt de basse et moyenne altitudes, puisque nous avons écarté de notre échantillon les bassins neigeux. La surface des bassins versants est également moins diversifiée et est comprise entre 249 km² et 4320 km².

Les bassins versants de ce second échantillon disposent de 7 à 28 pluviomètres situés sur ou à proximité du bassin. Il s'agit des pluviomètres dont le polygone de Thiessen associé et obtenu à partir du réseau de pluviomètres disposant de moins de 10% de lacunes par mois, intersecte le bassin versant (voir annexe D).

La surface pluviométrique associée à chaque bassin varie entre 130 km² et 7500 km². Cette surface pluviométrique (S_{PLU}) correspond à la moyenne des surfaces des polygones de Thiessen [Thiessen, 1911] intersectant le bassin versant, pondérée par la surface d'occupation de chaque polygone sur le bassin [Lobligeois, 2014] :

$$S_{PLU} = \sum_{k=1}^N \rho_k \times S_k^{Th}$$

avec N le nombre de pluviomètres, ρ_k la proportion de la surface du polygone de Thiessen du pluviomètre k comprise dans le bassin versant et S_k^{Th} la surface du polygone de Thiessen du pluviomètre k .

Ainsi, plus la surface pluviométrique est grande, moins le réseau pluviométrique est dense.

Par ailleurs, la Figure 69 indique la position de chaque bassin versant dans le graphique représentant le nombre de postes pluviométriques pour 100 km² en fonction de la superficie des bassins versant. Nous observons sur ce graphique que tous les bassins versants de notre étude se situent au-dessus de la limite de densité recommandée par Schaake et al. [2000].

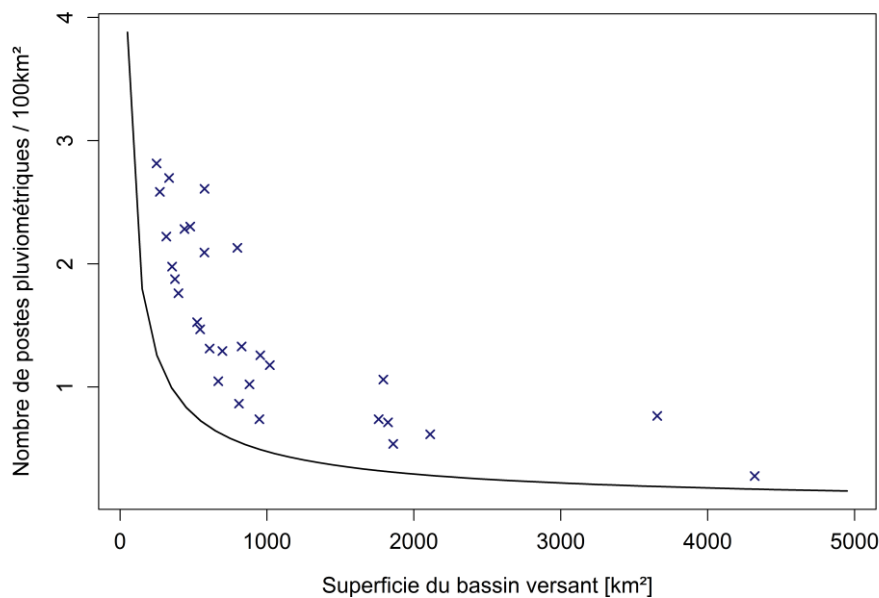


Figure 69. Nombre de postes pluviométriques pour 100 km² en fonction de la superficie des bassins versants. La courbe noire représente la limite conseillée par Schaake et al. (2000). Les croix bleues représentent les bassins de l'étude.

8.4. Les calculs hydrologiques, le calage du modèle et leur évaluation

8.4.1 Les calculs hydrologiques considérés

Les calculs hydrologiques considérés dans cette partie sont les mêmes que ceux considérés dans la partie II (à l'exception du QMNA5, qui, nous l'avons vu dans la première partie, présente des résultats très similaires au quantile de débit 0.05). Toutefois, comme nous ne disposons que de 10 ans de chronique de pluie pour cette partie, les périodes de calage des paramètres du modèle GR4J vont évoluer.

En première partie, la robustesse du modèle était évaluée selon sa capacité à être performant en régionalisation. Ainsi, nous pouvions utiliser la même période d'évaluation en calage et en validation.

Dans cette seconde partie, nous ne nous intéressons plus à la régionalisation. Nous évaluerons les performances des simulations, selon une procédure de *split-sample test* [Klemes, 1986]. On a ainsi deux périodes de calage/validation, 1998-2001 [P1] et 2003-2006 [P2] avec un an de mise en route du modèle. Cette procédure permet de mieux évaluer la robustesse du modèle face à un éventuel changement de conditions climatologiques entre la période de calage et la période de validation des paramètres.

Pour l'estimation du module, des débits de crue et des caractéristiques de débit d'étiage, le calage des paramètres du modèle a été effectué sur 9 ans (1998-2006 + 1 an de mise en route du modèle). Les différents calculs ont ensuite été effectués sur ces 9 ans de débit simulé.

8.4.2 Le calage des paramètres et évaluation du modèle

La procédure de calage des paramètres du modèle GR4J est la même que pour la partie I (voir chapitre 3). Les critères de calage et d'évaluation des calculs hydrologiques sont également les mêmes que dans la partie I (voir chapitre 5).

Nous présenterons dans cette partie les résultats sous forme de distributions relatives à tous les bassins étudiés :

- distribution des valeurs du critère de Nash pour la simulation de la chronique journalière ;
- distribution de l'erreur absolue et du biais pour les autres calculs de débit. Le biais et l'erreur sont calculés à partir d'une valeur de référence obtenue à partir de toute la chronique de débit observé sur chaque bassin versant.

8.5. Synthèse

Dans cette partie, nous avons présenté l'échantillon de 30 bassins versants sélectionné pour tester l'impact de la densité du réseau pluviométrique sur divers calculs hydrologiques. Cet échantillon présente des caractéristiques morphologiques et hydro-climatiques variées. Ces bassins versants disposent également d'un nombre varié de pluviomètres à leur proximité : de 7 à 28 pluviomètres par bassin.

Les calculs hydrologiques de cette partie sont les mêmes que ceux de la première partie mais sont calculés sur la période de 10 ans 1997-2006 (années sur lesquelles on dispose des données de

pluviométrie sur des stations Météo-France). Malheureusement, cette courte période de calcul pourra rendre difficile l'estimation des calculs nécessitant de longues chroniques de données, telles que le module ainsi que les débits de crue et d'étiage. Toutefois, étant donné que le principal but de cette étude est d'estimer comment les calculs hydrologiques sont dégradés ou améliorés en fonction de la densité spatiale du réseau pluviométrique, on gardera la même référence observée afin de calculer l'erreur associée pour chacune des densités testées.

Chapitre 9.

Analyse de la sensibilité des calculs hydrologiques à la densité spatiale des réseaux pluviométriques

9.1. Introduction

Les précipitations sont des données essentielles aux études hydrologiques. Toutefois, la pluie est variable spatialement et temporellement. Dans ce contexte, la représentativité des informations de pluie est capitale afin de limiter les incertitudes des calculs hydrologiques, liées à l'échantillonnage de cette variable dans l'espace et le temps. La densité et la configuration du réseau de mesure de la pluie peuvent alors avoir un impact majeur sur la précision de ces calculs. Le bon sens hydrologique nous dit que moins nous avons de connaissance à propos de la pluie de bassin, plus les calculs hydrologiques seront incertains. Quantifier cette tendance, c'est-à-dire la sensibilité de nos calculs à la configuration des réseaux de stations pluviométriques, est essentiel dans un contexte de demande croissante sur la connaissance des variables hydrologiques et de financement en baisse.

Dans ce chapitre, nous souhaitons tester l'impact de la densité spatiale du réseau pluviométrique sur divers calculs hydrologiques présentés dans le chapitre précédent. Dans un premier temps, la méthodologie adaptée pour cette partie sera présentée puis suivront les résultats concernant l'impact de la densité spatiale du réseau pluviométrique sur les calculs hydrologiques. Enfin, une analyse des résultats sera effectuée.

9.2. Quelques éléments de méthodologie

9.2.1 Mise en place des différentes classes de densité spatiale du réseau pluviométrique

Pour ce chapitre, nous avons évalué l'impact de différentes densités de réseau pluviométrique sur les performances de divers calculs hydrologiques. Pour cela, il a fallu constituer différentes classes de densité de pluviomètres.

Au cours de cette thèse, nous avons sous-échantillonné le réseau pluviométrique. Chaque bassin versant possède un nombre maximal n de pluviomètres pour décrire ses pluies. N classes de densité constituées de 1 à n pluviomètres ont ainsi été générées. Afin d'obtenir des résultats robustes, pour chaque bassin versant, environ 1000 configurations de réseau de pluviomètres sont envisagées lorsqu'un nombre suffisant n de pluviomètres est disponible sur le bassin (Tableau 11), en essayant de répartir au mieux ces combinaisons par classe de densité. En dessous de 10 pluviomètres présents sur le bassin, toutes les configurations possibles de réseau par classe de densité sont conservées et testées (Tableau 11). Nous rappelons que l'estimation de la pluie de bassin est effectuée à l'aide de la méthode de pondération par l'inverse de la distance au carré, présentée dans le chapitre 8.

Tableau 11. Nombre total de configurations de réseau envisagées selon le nombre de pluviomètres disponibles sur le bassin versant.

Nombre de pluviomètres	7	8	9	10	11	12	13	15	17	19	28
Nombre total de configurations	127	255	511	1023	883	934	975	1015	1016	1015	1007

Enfin, par la suite, afin de pouvoir comparer la baisse de densité entre les bassins, les densités du réseau pluviométrique ont été représentées par pourcentage. Nous présenterons dans les résultats de cette partie, les densités suivantes : 100%, 80%, 60%, 40%, 30%, 20% et 10% du réseau pluviométrique complet de chaque bassin.

Le Tableau 12 présente le nombre total de configurations de réseau selon le nombre de pluviomètres disponibles sur le bassin versant et selon la densité du réseau pluviométrique considérée.

Tableau 12. Nombre total de configurations de réseau selon le nombre n de pluviomètres disponibles sur le bassin versant et selon la densité du réseau pluviométrique (100% étant le réseau complet et 10% le réseau constitué de 10% des n pluviomètres disponibles sur le bassin).

Nombre de pluviomètres	100%	80%	60%	40%	30%	20%	10%
7	1	7	35	35	21	7	7
8	1	28	56	56	28	28	8
9	1	36	126	126	84	36	9
10	1	45	210	210	120	45	10
11	1	55	125	125	125	55	11
12	1	66	111	111	111	66	12
13	1	99	99	99	99	99	13
15	1	82	82	82	82	82	82
17	1	70	70	70	70	70	70
19	1	61	61	61	61	61	61
28	1	38	38	38	38	38	38

9.2.2 Présentation des bassins-tests

Pour cette partie, dans un premier temps, nous allons nous intéresser plus particulièrement à deux bassins versants afin d'essayer de mieux décrire et comprendre l'impact de la densité spatiale du réseau pluviométrique sur les divers calculs hydrologiques considérés. Les résultats obtenus pour les autres bassins versants sont disponibles en annexe E.

Les bassins sélectionnés pour une analyse plus approfondie sont les bassins de la Vézère à Campagne avec ses 28 pluviomètres (superficie de 3657 km²) et celui du Madon à Pulligny avec ses 7 pluviomètres (superficie de 948 km²). Ces deux bassins ont été tout simplement sélectionnés puisqu'ils comportent respectivement le nombre maximal et le nombre minimal de pluviomètres parmi les 30 bassins de notre étude.

Les Figure 70 et Figure 71 indiquent la localisation de ces deux bassins versants ainsi que celle de leurs pluviomètres. La Figure 72 présente la répartition de la pluie moyenne interannuelle calculée sur 10 ans sur ces deux bassins, à l'aide des données de la lame d'eau Comephore, combinant information radar et information pluviométrique [Tabary *et al.*, 2012]. Nous pouvons constater que sur le Madon à Pulligny, la pluie moyenne interannuelle est répartie de façon plutôt homogène alors que sur la Vézère à Campagne la répartition de la pluie est davantage contrastée entre les parties amont et aval du bassin versant. Ces observations sont à mettre en lien avec la carte de la Figure 70 représentant le relief de ce bassin. La limite amont/aval observée sur la carte des pluies moyennes interannuelles de la Figure 72 coïncident avec la limite de relief du bassin de la Figure 70. Alors que pour le Madon à Pulligny, le relief du bassin versant semble plutôt homogène sur toute la zone d'étude (Figure 71).

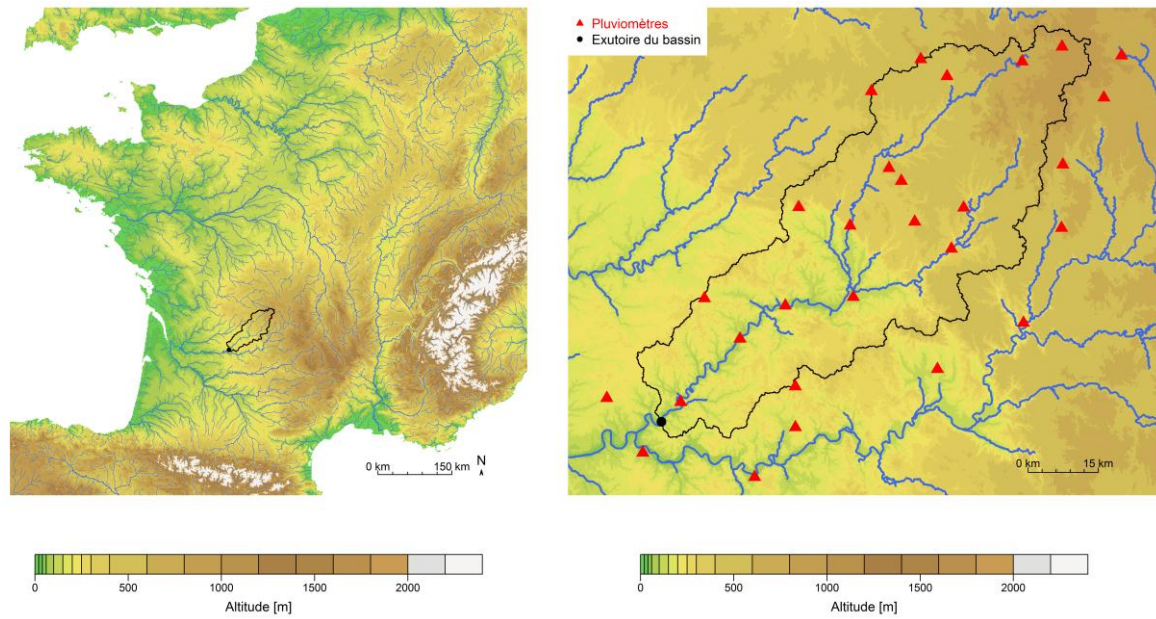


Figure 70. Localisation du bassin de La Vézère à Campagne (P4271010) ainsi que celle des 28 pluviomètres utilisés pour l'estimation de sa pluie de bassin (superficie du bassin : 3657 km²).

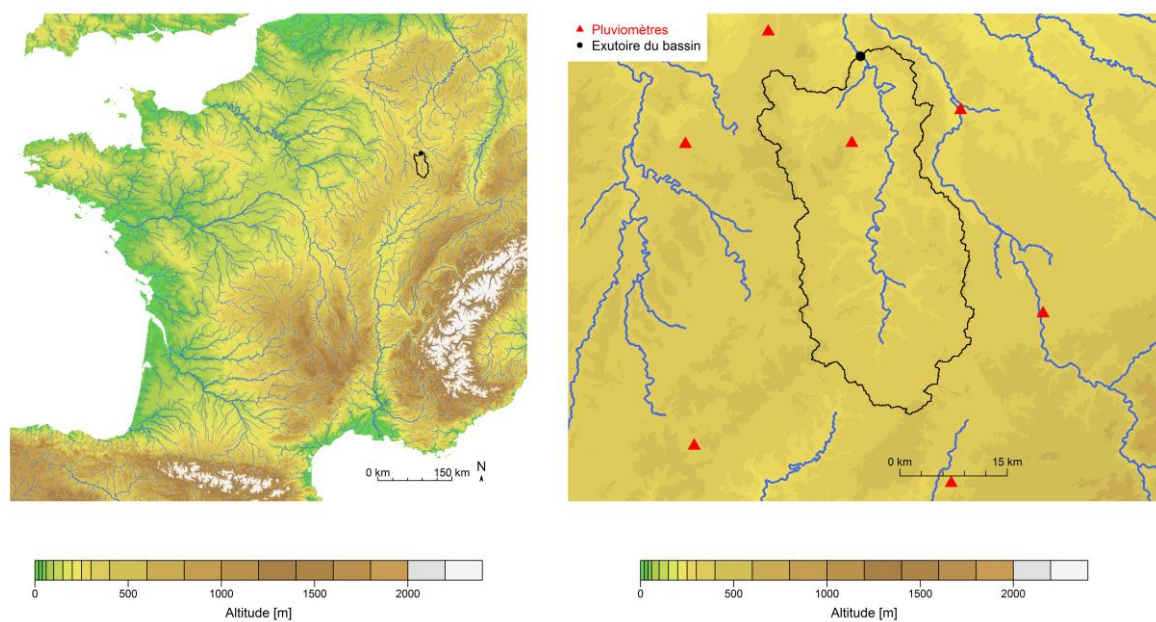


Figure 71. Localisation du bassin du Madon à Pulligny (A5431010) ainsi que celle des 7 pluviomètres utilisés pour l'estimation de sa pluie de bassin (superficie du bassin : 948 km²).

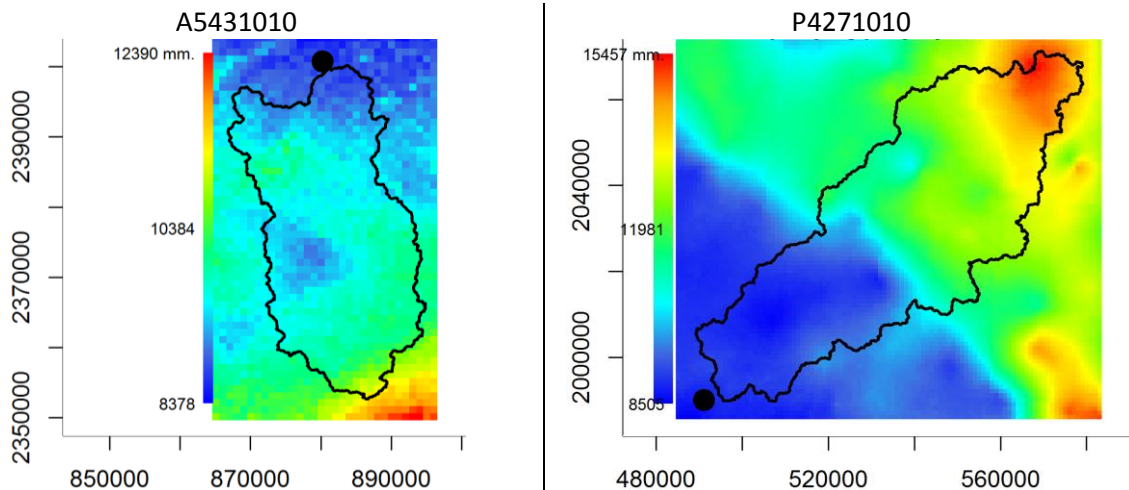


Figure 72. Pluie moyenne interannuelle (en $1/10^e$ de mm) calculée sur 10 ans à l'aide de la réanalyse Comephore de Météo-France (combinant données pluviométriques et radar) pour les bassins du Madon à Pulligny (A5431010) et de La Vézère à Campagne (P4271010).

9.3. Impact de la densité spatiale du réseau pluviométrique sur divers calculs hydrologiques

9.3.1 Sur les deux bassins-tests

Les Figure 73 et Figure 74 présentent les performances des calculs hydrologiques présentés dans le chapitre 8 lorsque la densité des pluviomètres disponibles pour calculer la pluie de bassin diminue, pour le bassin de la Vézère à Campagne et celui du Madon à Pulligny respectivement.

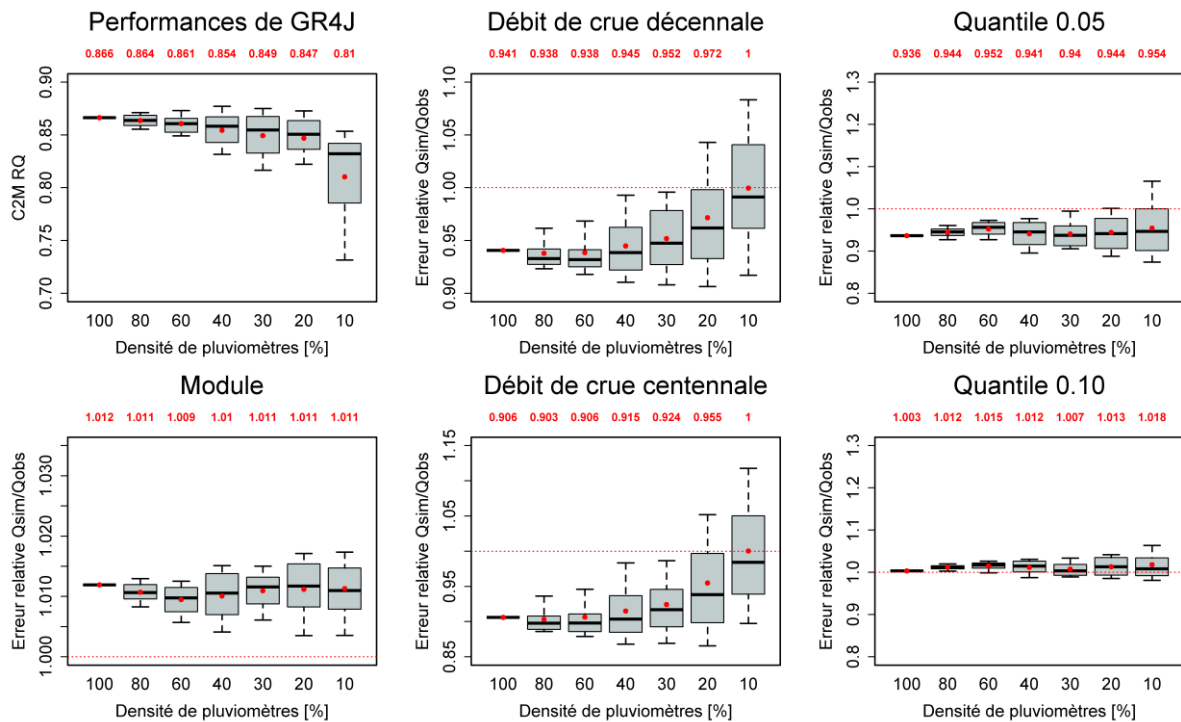


Figure 73. Performances des divers calculs hydrologiques considérés en fonction de la densité du réseau pluviométrique pour le bassin de la Vézère à Campagne.

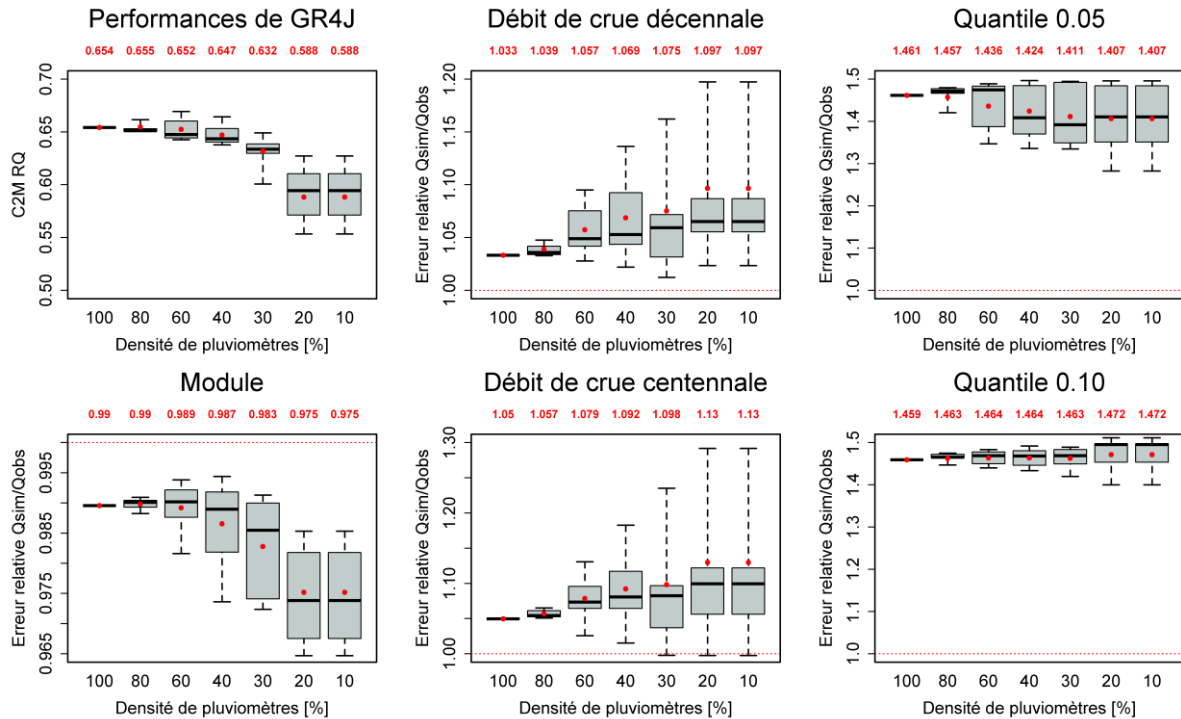


Figure 74. Performances des divers calculs hydrologiques considérés en fonction de la densité du réseau pluviométrique pour le bassin du Madon à Pulligny.

Performances du modèle GR4J

Concernant la simulation de chronique de débit journalier avec le modèle GR4J, nous observons pour les deux bassins que les performances du modèle diminuent progressivement à mesure que la densité du réseau pluviométrique diminue. Cette perte de performances lorsque l'on passe du réseau complet à un réseau réduit de 90% est en moyenne de 0.05 points de C2M pour le bassin de la Vézère à Campagne et de 0.07 points de C2M pour le bassin du Madon à Pulligny.

De plus, nous pouvons également constater pour les deux bassins que la variabilité des résultats augmente lorsque le réseau pluviométrique devient moins dense. Les calculs deviennent ainsi plus incertains à mesure que la densité du réseau baisse. Ces résultats corroborent ceux obtenus dans les études de *Andréassian et al.* [2001] et de *Oudin et al.* [2006a].

Les hydrogrammes des bassins versants de la Vézère à Campagne et du Madon à Pulligny sont respectivement disponibles en Figure 75 et Figure 76. Sur le graphique de la Vézère à Campagne, nous pouvons constater tout d'abord, que les hydrogrammes simulés avec un réseau constitué de 10% de stations présentent une très large dispersion du débit simulé, alors qu'avec l'utilisation du réseau complet, le débit est plutôt bien modélisé. Pour le bassin du Madon à Pulligny, la dispersion de débits simulés à l'aide du réseau réduit est beaucoup moins importante que pour le premier bassin mais ceci était déjà remarquable sur les Figure 73 et Figure 74. Cela est vraisemblablement lié au fait qu'il y ait moins de combinaisons de pluviomètres par classe de densité pour le bassin du Madon à Pulligny (Tableau 12).

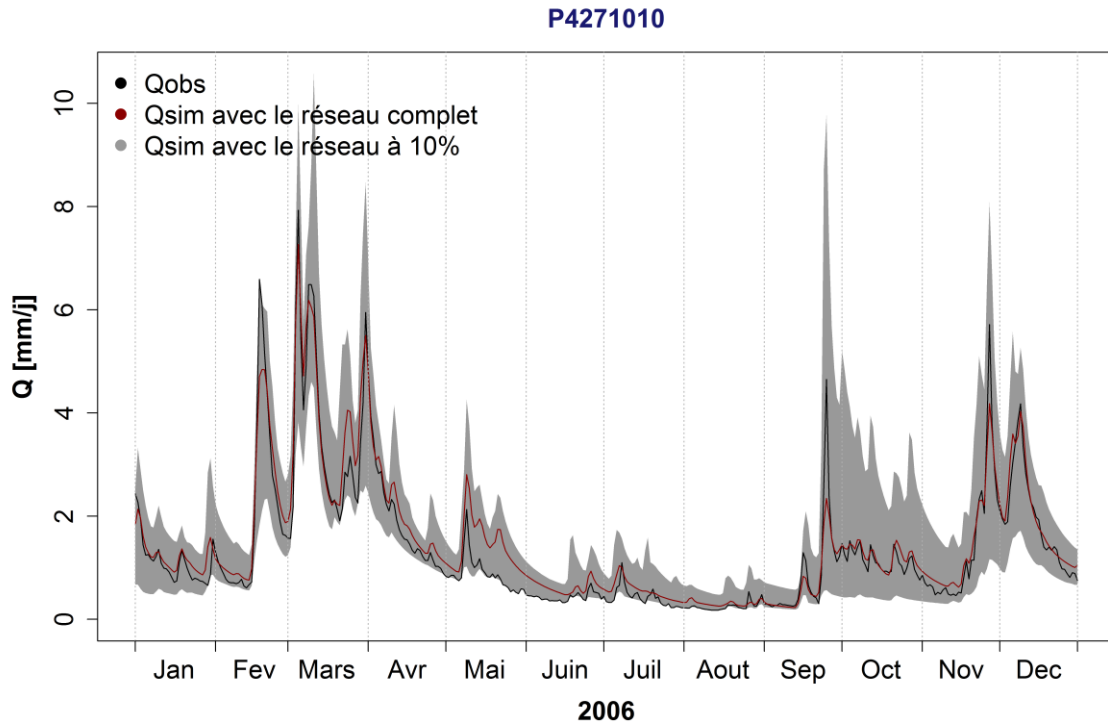


Figure 75. Hydrogrammes observé (en noir) et simulés en validation (avec le réseau pluviométrique complet en rouge et avec les réseaux constitués de 10% de stations pluviométriques en gris (borné par les débits simulés minimum et maximum de chaque jour)) sur le bassin de la Vézère à Campagne pour l'année 2006.

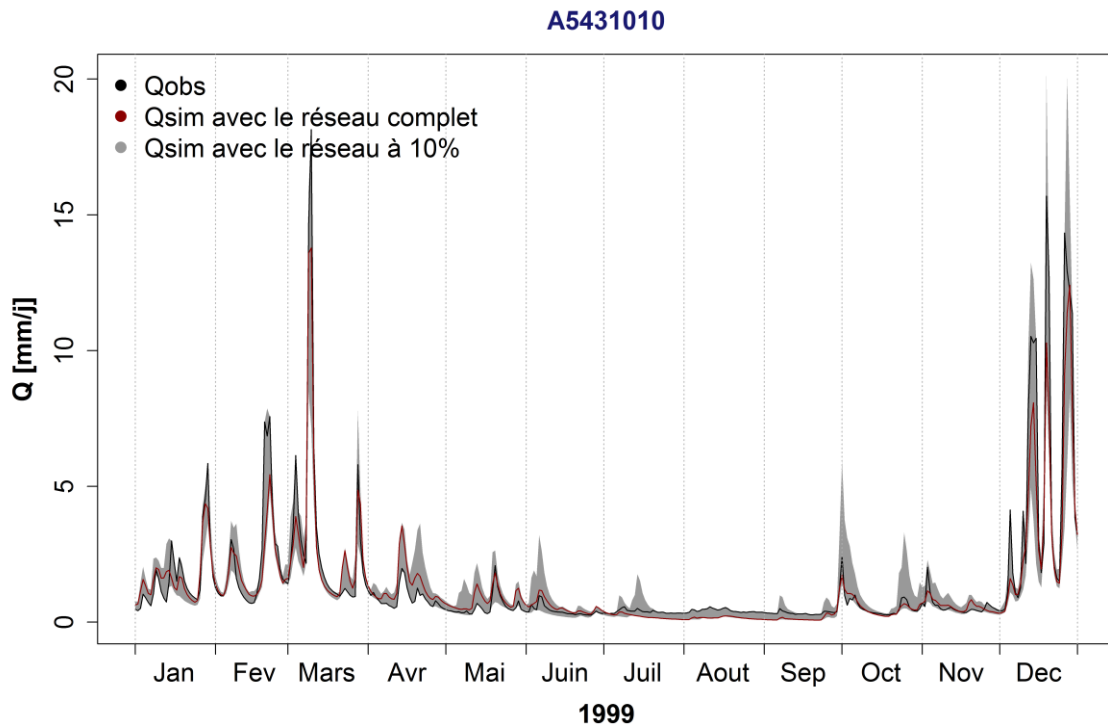


Figure 76. Hydrogrammes observé (en noir) et simulés en validation (avec le réseau pluviométrique complet en rouge et avec les réseaux constitués de 10% de stations pluviométriques en gris (borné par les débits simulés minimum et maximum de chaque jour)) sur le bassin du Madon à Pulligny pour l'année 1999.

Module

Pour le calcul du module, nous constatons comme pour la simulation de chronique de débit journalier, que les calculs sont de plus en plus incertains avec la baisse de la densité du réseau pluviométrique. Cependant, les deux bassins ne se comportent pas tout à fait de la même façon : pour le bassin de la Vézère à Campagne, l'erreur relative reste en moyenne autour de la valeur 1.01 lorsque le réseau devient moins dense, alors que pour le bassin du Madon à Pulligny, la valeur du module a tendance à être de plus en plus sous-estimée quand la densité du réseau baisse. Pour ce bassin, certaines configurations de réseaux peuvent facilement exclure la partie amont ou aval des pluviomètres, avec potentiellement des conséquences fortes sur les pluies de bassin si le champ pluvieux est très hétérogène. Alors que pour le bassin de la Vézère à Campagne, la présence d'un réseau assez dense et plutôt assez bien réparti sur le bassin, permet d'éviter des situations extrêmes avec une partie du bassin totalement dépourvu de pluviomètres.

Débits de crues décennale et centennale

De la même manière que pour le calcul du module, nous observons des comportements différents entre les deux bassins pour le calcul des débits de crues.

Pour le bassin de la Vézère à Campagne, nous observons le même comportement pour les deux calculs de débits de crues. Il y a deux façons d'interpréter les résultats (nous y reviendrons par la suite) : on peut avoir l'impression (et c'est relativement surprenant) que l'erreur relative est en moyenne améliorée lorsque le réseau devient moins dense ; on peut aussi analyser cette tendance en estimant que les crues simulées ont tendance à augmenter (quel que soit leur niveau initial) quand le réseau est moins dense. Nous constatons toutefois que la gamme des erreurs relatives devient de plus en plus large à mesure que la densité du réseau diminue.

Pour le bassin du Madon à Pulligny, les débits simulés de crues décennale et centennale augmentent lorsque le réseau devient moins dense, cette fois-ci, cela va dans le sens de la surestimation. De la même manière que précédemment, la gamme d'incertitudes des calculs augmentent également.

Dans les deux cas, lorsque les pluies de bassin sont calculées sur la base d'un faible nombre de pluviomètres, le modèle a tendance à produire des crues plus importantes. Cela peut s'expliquer par le fait que le signal de pluie obtenu à partir d'un petit nombre de pluviomètres est moins lissé par l'effet de moyenne. Dans le cas de la Vézère, où la pluie moyenne de bassin complète avait tendance à sous-estimer les crues, on observe donc un effet de compensation. Mais pour le Madon, où l'on surestimait déjà légèrement les crues, cela a pour effet d'exagérer cette surestimation.

Quantiles de bas débit

Pour les calculs des quantiles de bas débit, on n'observe pas d'évolution notable des quantiles moyens ou médians. Par contre, on observe bien une augmentation de l'incertitude comme dans les cas précédents.

9.3.2 Analyse pour l'ensemble des 30 bassins versants de l'échantillon

L'analyse des deux bassins exemples nous a montré que différents comportements étaient possibles. Nous allons maintenant essayer de généraliser nos observations sur l'ensemble de l'échantillon de bassins versants.

Performances du modèle GR4J

La Figure 77 présente les performances médianes du modèle GR4J lors de la simulation d'une chronique de débit journalier lorsque la densité du réseau pluviométrique (ici surface/nombre de pluviomètres) diminue.

Nous pouvons constater que les performances médianes ont toujours tendance à diminuer lorsque le réseau pluviométrique devient moins dense. Toutefois, nous remarquons également que cette chute de performances peut s'effectuer de manière plus brutale pour certains bassins par rapport à d'autres où la pente de décroissance observée est plus douce. Par la suite, nous allons chercher à identifier des facteurs pouvant influencer cette vitesse de décroissance des performances lorsque le réseau des stations pluviométriques diminue.

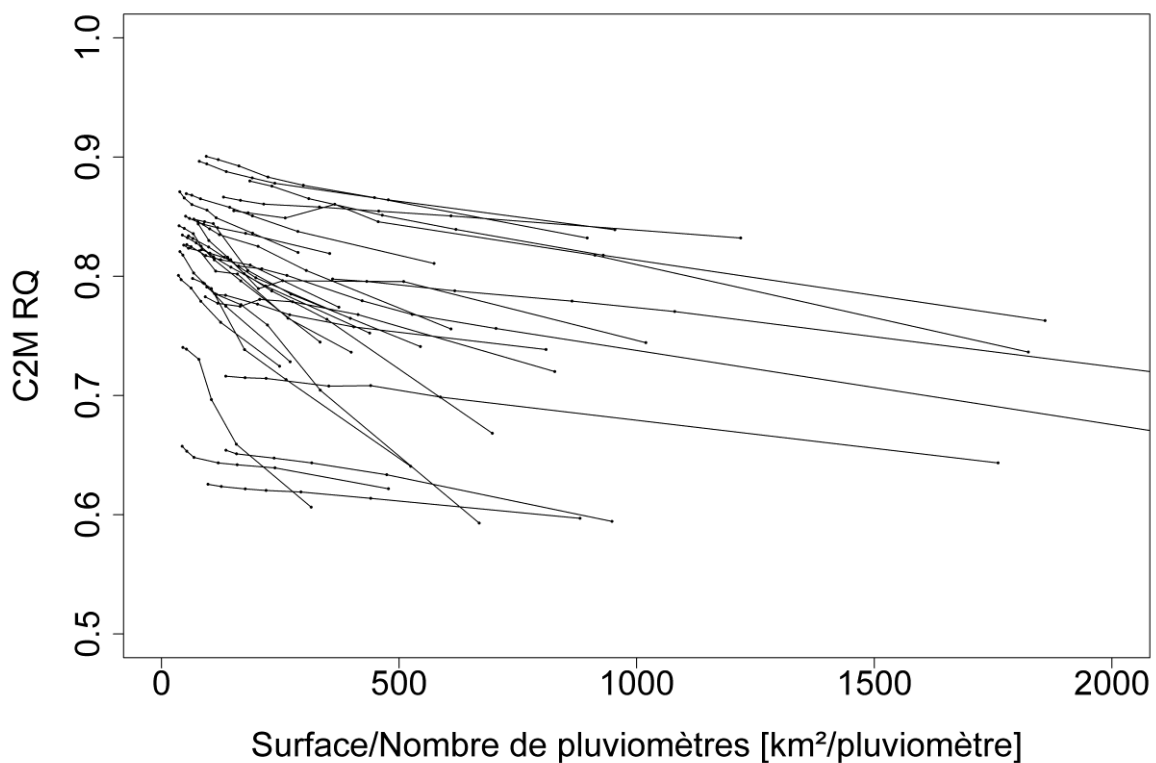


Figure 77. Valeur médiane des performances du modèle GR4J en validation en fonction de la surface par pluviomètre pour les 30 bassins versants. Chaque courbe représente les performances du modèle GR4J pour chaque bassin versant lorsque la densité du réseau pluviométrique diminue.

Le Tableau 13 présente un récapitulatif des tendances d'évolution des erreurs relatives médianes (Q_{sim}/Q_{obs}) lorsque le réseau pluviométrique passe d'un réseau complet à un réseau avec une densité de stations de 10% pour les calculs de module, de débit de crue et débit d'étiage. Il est important de souligner qu'ici il ne s'agit pas de quantifier une augmentation ou une diminution de l'erreur relative

par rapport à la valeur idéale de 1 (ces tendances sont difficilement interprétables car elles varient beaucoup d'un bassin à un autre), mais bien de quantifier une tendance d'évolution du rapport Q_{sim}/Q_{obs} entre le réseau complet et le réseau réduit à 10% de stations. Ces résultats sont basés sur une interprétation visuelle des graphiques disponibles en annexe E. Ces graphiques peuvent, dans un premier temps, nous confirmer ce qui a été observé pour les deux bassins exemple : pour l'ensemble des bassins versants, la gamme d'incertitude de ces calculs augmente lorsque le réseau pluviométrique devient moins dense.

Module

Dans le Tableau 13, nous pouvons constater pour le module qu'aucune tendance claire d'évolution ne semble se dessiner. La pente de la courbe reliant les performances médianes du réseau complet à celles du réseau constitué de 10% de pluviomètres, est dans l'ensemble nulle. Nous remarquons toutefois pour un tiers des bassins, une tendance à la décroissance. D'autre part, nous pouvons noter également à l'aide des résultats disponibles en annexe E, que seulement quatre bassins ont une erreur relative médiane qui se rapproche de la valeur optimale de 1 lorsque le réseau devient moins dense pour le calcul du module.

Débits de crues décennale et centennale

Pour les débits de crues décennale et centennale, une même tendance à la croissance de l'erreur relative est observée. Nous avons donc un débit simulé qui est de plus en plus grand en comparaison du débit observé. En enlevant progressivement des stations, nous pouvons supposer que les stations pluviométriques restantes vont avoir plus de poids sur l'ensemble du bassin. Ainsi, une pluie ponctuelle forte qui serait enregistrée sur une partie du bassin par un ou deux pluviomètres, sera interpolée à l'ensemble du bassin telle quelle, sans être lissée par la pluie d'autres stations. En allant vers des données de plus en plus ponctuelles, la pluie est alors moins lissée et les crues peuvent devenir plus fortes. Ce phénomène était également observé sur les pointes de crues illustrées sur les hydrogrammes des Figure 75 et Figure 76.

Quantiles de bas débit

Pour les quantiles de bas débit, nous observons également une tendance à l'augmentation du rapport Q_{sim}/Q_{obs} quand le réseau complet est réduit de 90%.

Tableau 13. Récapitulatif des tendances d'évolution des erreurs relatives médianes (Q_{sim}/Q_{obs}) lorsque le réseau pluviométrique passe d'un réseau pluviométrique complet à un réseau avec une densité de 10%. En rouge, le plus grand nombre de bassins recensés pour chaque calcul.

	Module	Crue décennale	Crue centennale	Quantile 0.05	Quantile 0.10
Nombre de bassins avec une pente décroissante (-)	11	8	6	10	8
Nombre de bassins avec une pente croissante (+)	0	20	19	16	20
Nombre de bassins avec une pente nulle (=)	19	2	5	4	2

9.4. Tentatives d'interprétation des résultats

9.4.1 Relation de la courbe de décroissance des performances lorsque la densité du réseau pluviométrique diminue, avec la surface pluviométrique du bassin versant

Dans un premier temps, nous avons essayé de déterminer si un lien existait entre la densité initiale de pluviomètres et la vitesse de décroissance des performances lorsque le réseau pluviométrique devient moins dense. En effet, on pourrait s'imaginer que dans le cas d'un réseau peu dense, faire varier le réseau de un ou deux pluviomètres pourrait engendrer des conséquences plus importantes sur la simulation de débit, que dans le cas d'un réseau initialement dense.

La Figure 78 présente les performances médianes du modèle GR4J lors de la simulation d'une chronique de débit journalier lorsque la densité du réseau pluviométrique (ici surface/nombre de pluviomètres) diminue, présentées en Figure 77. La couleur associée à chacune des courbes représente la valeur de la surface pluviométrique S_{plu} de chaque bassin versant (voir chapitre 8). Comme nous l'avons vu dans le chapitre 8, la surface pluviométrique du bassin versant nous donne un indice de la densité pluviométrique du bassin versant. Plus la surface pluviométrique est grande, moins le réseau pluviométrique est dense.

Nous pouvons observer sur cette figure que la vitesse de décroissance des performances semble être relativement moins brutale lorsque la surface pluviométrique est importante, soit quand la densité du réseau est plus faible. Ceci signifierait que plus un réseau pluviométrique est initialement dense sur un bassin versant (S_{plu} faible), plus la vitesse de décroissance des performances du modèle GR4J aura tendance à être forte lorsque la densité du réseau diminuera. Les résultats sont donc contraires à ce que nous attendions. Toutefois, des exceptions sont à noter.

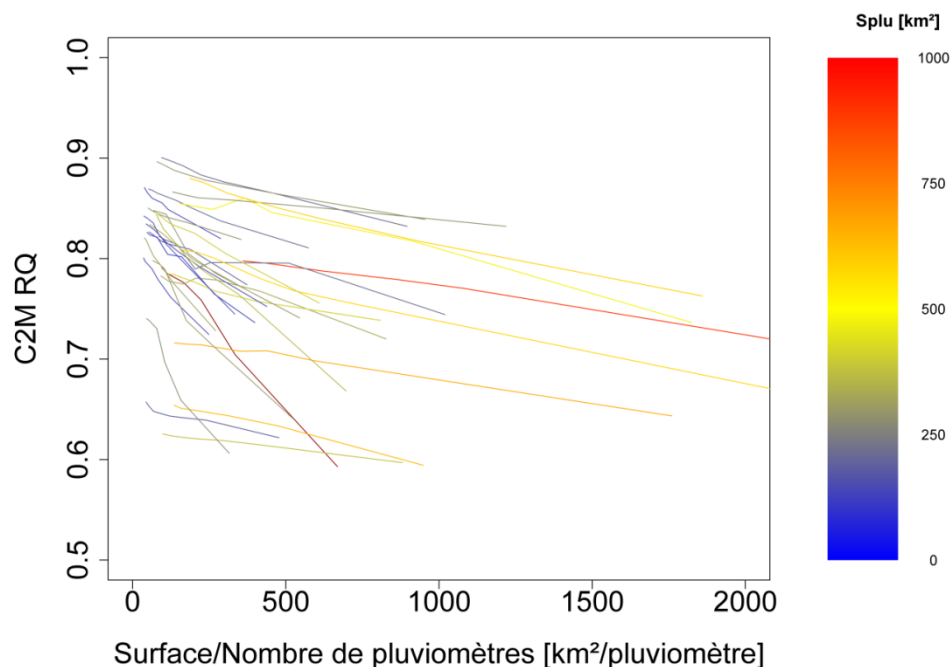


Figure 78. Valeur médiane des performances du modèle GR4J en fonction de la surface par pluviomètre pour les 30 bassins versants mais également en fonction de la valeur de la surface pluviométrique associée au bassin versant. En rouge foncée, S_{plu} est supérieure à 5000 km².

9.4.2 Relation de la décroissance des performances lorsque la densité du réseau pluviométrique diminue, avec la variabilité spatiale de la pluie

Dans un second temps, nous nous proposons d'étudier si une relation existe entre la décroissance des performances de GR4J avec la diminution de la densité du réseau pluviométrique, et la variabilité spatiale de la pluie du bassin versant. Pour cela, la géostatistique peut être un outil utile pour nous renseigner sur la variabilité spatiale d'une variable quantitative, définie en tout point de l'espace, comme ici la pluie.

9.4.2.1 Quelques rappels de géostatistique

Historiquement, la géostatistique a été développée pour le domaine minier [Krige, 1951; Matheron, 1963]. Dans les années 50, le but originel était d'estimer la variabilité de la teneur du minerai. Actuellement, le domaine d'application de la géostatistique est beaucoup plus large. La géostatistique peut s'appliquer à toute variable régionalisée, dans des domaines aussi variés que la prospection pétrolière, la météorologie, la géologie, les sciences du sol, l'environnement, etc.

Un phénomène régionalisé est un phénomène naturel dont les valeurs varient dans l'espace. Une variable régionalisée est une fonction mathématique dont les valeurs sont déterminées sur une partie de l'espace, le champ. Toutefois, nous ne connaissons la variable qu'en seulement quelques points. De plus, cette variable régionalisée peut être irrégulière et nous ne pouvons donc la modéliser par une fonction mathématique explicite. Cependant, elle présente une structure spatiale qui permet une corrélation entre les valeurs. Les outils de géostatistique vont donc permettre d'analyser, comprendre et quantifier la corrélation spatiale des valeurs de la variable régionalisée. L'idée implicite se dégageant est que deux observations en des points proches devraient se ressembler davantage que deux observations éloignées.

L'outil le plus largement utilisé en géostatistique est le variogramme. Le variogramme est une fonction décrivant la continuité spatiale d'une variable. Il permet d'obtenir en tout point de l'espace une estimation de la variable régionalisée, avec une estimation de la variance associée.

La variable régionalisée est interprétée comme une réalisation d'une fonction aléatoire $Z(x)$. Concernant les hypothèses de stationnarité liées au variogramme, celui-ci est une fonction aléatoire stationnaire d'ordre 2. Cela suppose que les deux premiers moments (moyenne et variance) existent et sont invariants par translation.

Le variogramme empirique ou expérimental s'écrit :

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i + h) - Z(x_i)]^2 \quad \text{Eq. 26}$$

où $Z(x_i)$ est la valeur de la variable au point x_i et $N(h)$ est le nombre de couples de points distants de h .

Le variogramme empirique (Figure 79) dépend de plusieurs caractéristiques dont:

- La portée: il s'agit de la distance à partir de laquelle deux observations ne sont plus corrélées. A cette distance, la valeur du variogramme correspond à la valeur de la variance

moyenne de la variable aléatoire. Au-delà, il n'y a ainsi plus de corrélation spatiale entre les valeurs.

- Le palier : il s'agit de la variance de la variable aléatoire. C'est la valeur obtenue lorsqu'on atteint la portée.
- L'effet de pépité : il s'agit d'un saut à l'origine qui signale une grande irrégularité de la variable régionalisée et l'absence de corrélation entre les valeurs prises en deux points très proches. Celui-ci n'est pas toujours observé sur un variogramme et peut témoigner d'une erreur de mesures, par exemple.

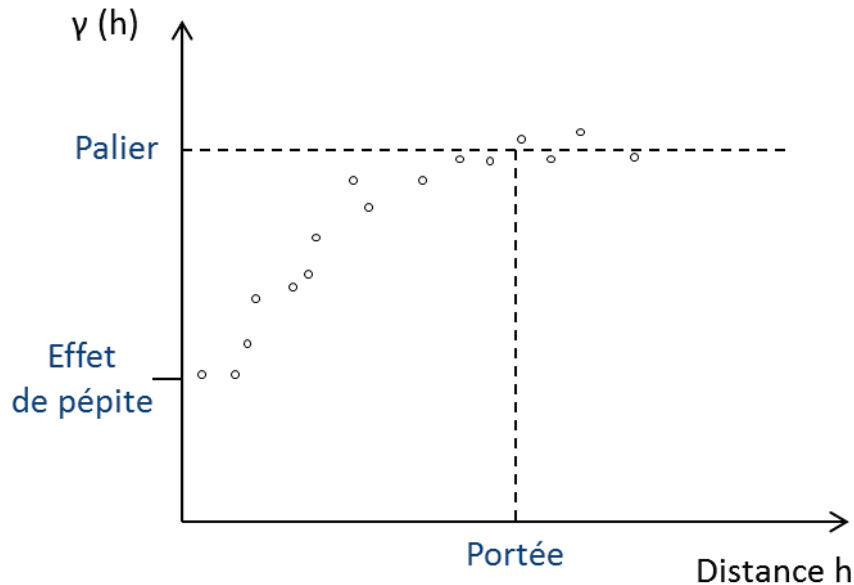


Figure 79. Schéma explicatif d'un variogramme expérimental.

Dans certains cas, la portée ne peut être déterminée car le variogramme ne montre pas de palier. On parle alors de variogramme non borné. Ceci apparaît lorsque les distances séparant les points sont trop petites par rapport à la structure spatiale étudiée (problème d'échelle) ou lorsque les propriétés de stationnarité ne sont plus respectées (la moyenne et/ou la variance ne sont pas constantes).

A proximité de l'origine du variogramme, différents comportements peuvent également être observés :

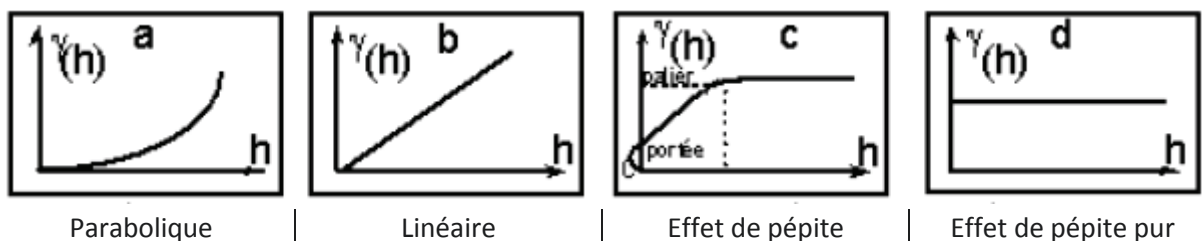


Figure 80. Différents comportements à l'origine du variogramme. (a) Allure parabolique, (b) Forme linéaire, (c) Variogramme à effet de pépité, (d) Effet de pépité pur.

Ces différents comportements à l'origine traduisent la vitesse de décroissance de la variabilité spatiale des pluies sur des distances inférieures à la portée. Par exemple, pour un cas parabolique, la variabilité va décroître très lentement pour de faible distance, alors que pour un cas linéaire la décroissance de la variabilité sera beaucoup plus rapide pour de faibles distances. Par ailleurs, un variogramme à effet de pépite pur signifie qu'aucune corrélation spatiale n'existe entre les observations.

9.4.2.2 Le variogramme climatologique empirique

Le variogramme climatologique empirique a été construit et tracé à l'aide de toutes les pluies journalières disponibles au niveau des postes pluviométriques sélectionnés pour chacun des bassins versants sur la durée de 10 ans 1997-2006. Les variogrammes climatologiques empiriques des 30 bassins versants sont disponibles en annexe F.

Ces variogrammes montrent des comportements très différents entre les bassins versants. De manière générale, deux comportements sont observés :

- Variance relativement homogène des pluies selon la distance (exemple du bassin du Madon à Pulligny (Figure 81)) :

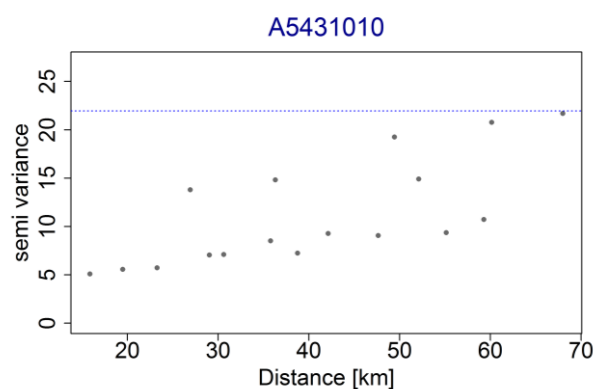


Figure 81. Variogramme climatologique empirique du bassin du Madon à Pulligny (code hydro : A5431010). La ligne en pointillé bleu représente la variance moyenne de la variable aléatoire.

- Variance qui augmente plus nettement avec la distance (exemple du bassin de la Vézère à Campagne (Figure 82)) :

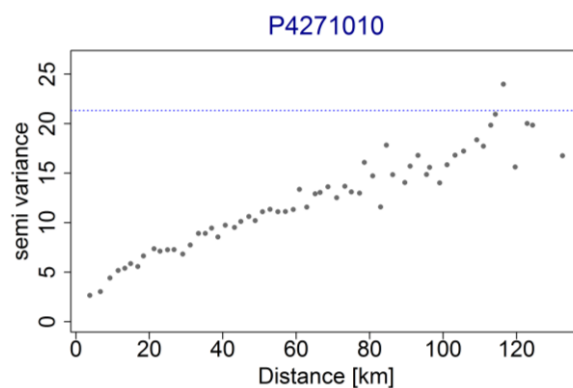


Figure 82. Variogramme climatologique empirique du bassin de la Vézère à Campagne (code hydro : P4271010). La ligne en pointillé bleu représente la variance moyenne de la variable aléatoire.

Ces deux types de comportement peuvent être mis en relation avec les cartes de pluie moyenne interannuelle de la Figure 72. Sur le bassin du Madon à Pulligny, cette pluie était plutôt homogène sur l'ensemble du bassin versant, la variabilité spatiale de la pluie était ainsi plutôt limitée, alors que sur le bassin de la Vézère à Campagne cette pluie était davantage contrastée entre les parties amont et aval du bassin versant (contraste probablement dû à l'effet d'altitude). Ainsi, sur ce bassin, plus la distance entre deux points augmente, plus la pluie varie d'un point à l'autre.

Sur certains variogrammes climatologiques empiriques, deux portées peuvent être identifiées (on parle alors de structure gigogne) : une première portée se situant entre 30 et 40 km résultant d'événements de faible distance comme par exemple les événements convectifs et une seconde portée au-dessus de 60 km, témoignant d'une plus grande structure de plus grande échelle (voir l'exemple du bassin de l'Azergues à Lozanne de la Figure 83).

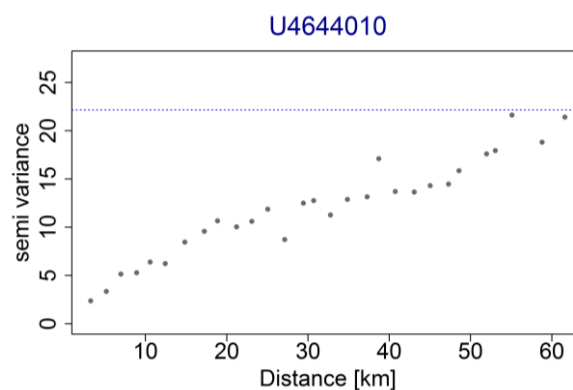


Figure 83. Variogramme climatologique empirique du bassin de l'Azergues à Lozanne (code hydro : U4644010). La ligne en pointillé bleu représente la variance moyenne de la variable aléatoire.

9.4.2.3 Courbe de décroissance des performances en fonction de la pente du variogramme climatologique empirique des pluies

Au cours de cette thèse, nous nous sommes en particulier intéressés à la pente de la régression linéaire passant par les points du variogramme climatologique empirique. Nous faisons l'hypothèse suivante : plus la pente de la régression sera forte, plus vite dans l'espace la corrélation des pluies chutera. Nous pourrions alors nous attendre à des performances des calculs hydrologiques moins bonnes à mesure de la réduction de la densité du réseau pluviométrique. *Baccour et al.* [2012] utilisent également la pente du variogramme d'évapotranspiration potentielle (rapportée à la variance) afin de définir la vitesse de variation de la structure au sein du domaine étudié dans le cas de variogramme linéaire.

La Figure 84 présente les performances médianes du modèle GR4J lors de la simulation d'une chronique de débit journalier lorsque la densité du réseau pluviométrique (ici surface/nombre de pluviomètres) diminue, présentées en Figure 77. Cette fois-ci, la palette de couleurs associée représente la pente du variogramme climatologique. Nous pouvons observer sur ce graphique que les décroissances de performances du modèle GR4J les plus douces semblent correspondre à des pentes du variogramme climatologiques faibles. Alors que les fortes valeurs de pente du variogramme semblent coïncider avec des vitesses de décroissances des performances plus élevées. Toutefois, quelques exceptions sont à noter. Nous pouvons retrouver cette tendance sur la Figure 85 qui présente la valeur de la pente de décroissance des performances du modèle GR4J lorsque la

densité du réseau pluviométrique diminue en fonction de la pente du variogramme climatologique associée à chaque bassin. Nous avons également tracé la régression linéaire obtenue à partir de ces points. Cette régression montre que la tendance générale confirme ce qui a été trouvé précédemment. Une nouvelle fois quelques exceptions peuvent survenir, d'où un coefficient de corrélation relativement peu élevé.

Les résultats obtenus semblent tout à fait logiques : plus la variabilité spatiale de la pluie sur un bassin est importante, plus vite les performances du modèle GR4J se dégraderont lorsque la densité du réseau pluviométrique diminuera.

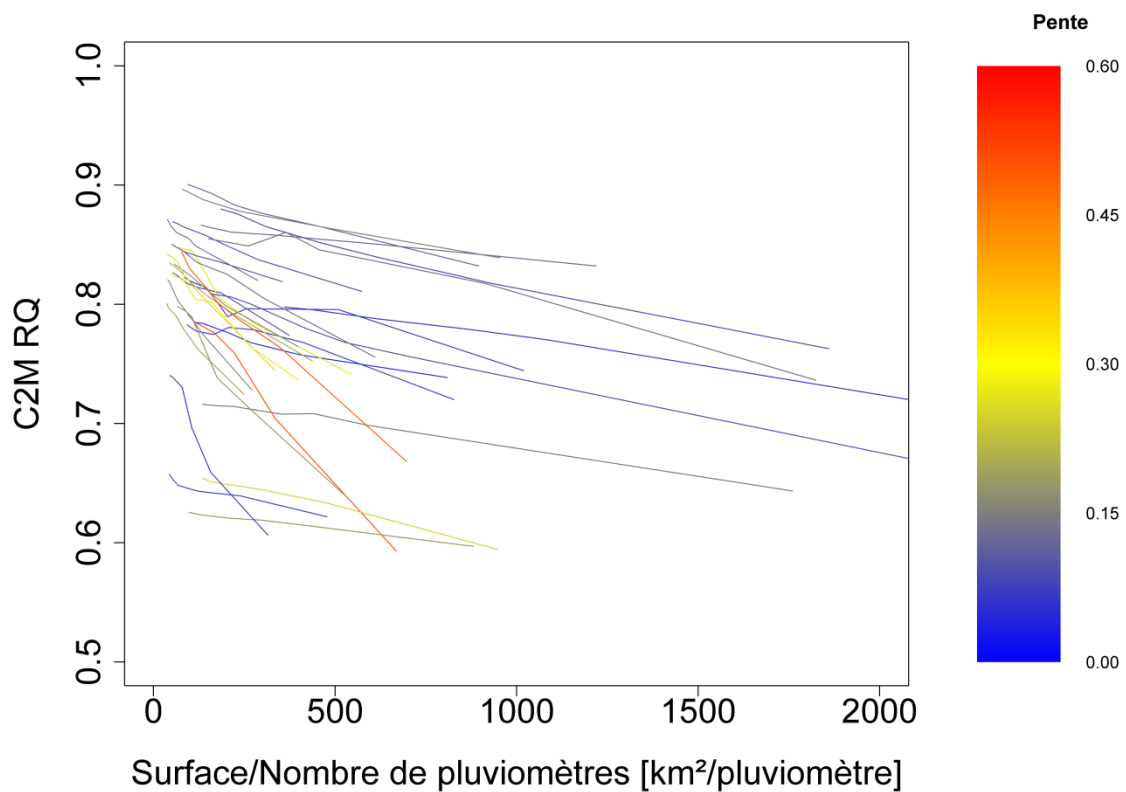


Figure 84. Valeur médiane des performances du modèle GR4J en fonction de la surface par pluviomètre pour les 30 bassins versants mais également en fonction de la valeur de la pente du variogramme climatologique.

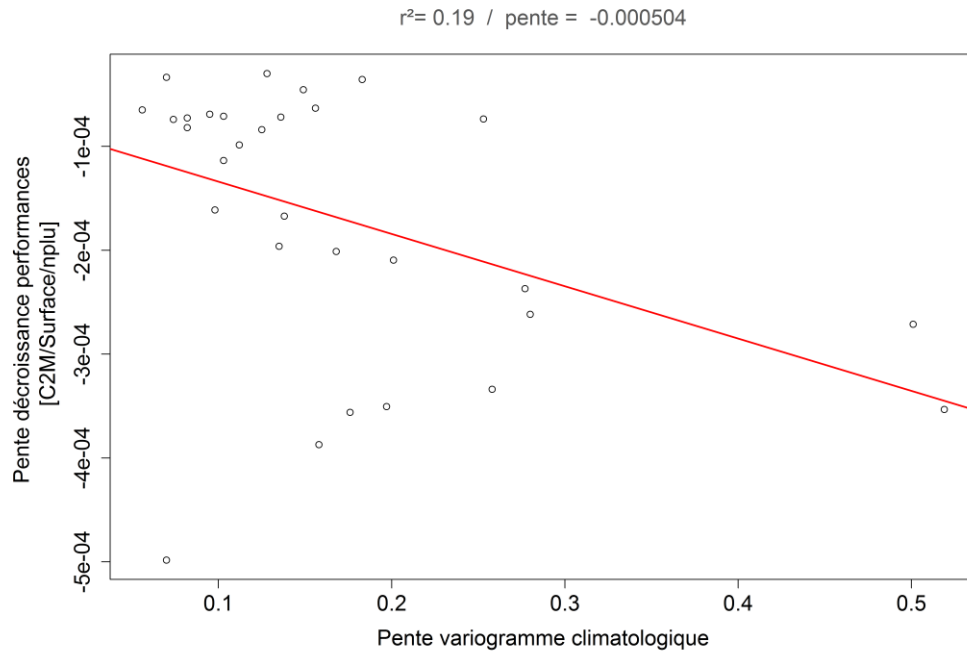


Figure 85. Graphique représentant la valeur de la pente de décroissance des performances du modèle GR4J lorsque la densité du réseau pluviométrique diminue en fonction de la pente du variogramme climatologique associée à chaque bassin. En rouge, la régression linéaire calée sur les points du graphique. En haut du graphique, r^2 correspond au coefficient de corrélation associé à la régression linéaire et la pente correspond à la pente de cette régression linéaire.

9.4.3 Interprétation de la courbe de décroissance des performances en fonction d'autres caractéristiques

Nous avons également essayé de trouver des corrélations entre la décroissance des performances en fonction de la densité pluviométrique et d'autres caractéristiques de bassin versant.

Les Figure 86 et Figure 87 présentent ainsi les performances médianes du modèle GR4J lors de la simulation d'une chronique de débit journalier lorsque la densité du réseau pluviométrique (ici surface/nombre de pluviomètres) diminue, présentées en Figure 77. La colorimétrie associée à chacune des figures correspond aux valeurs de surface de bassin versant pour la Figure 86 et l'indice de débit de base pour la Figure 87.

Pour ces deux caractéristiques, aucune corrélation n'est clairement identifiable entre la vitesse de décroissance des performances du modèle GR4J et la surface du bassin dans le premier cas, et l'indice de débit de base dans le second cas.

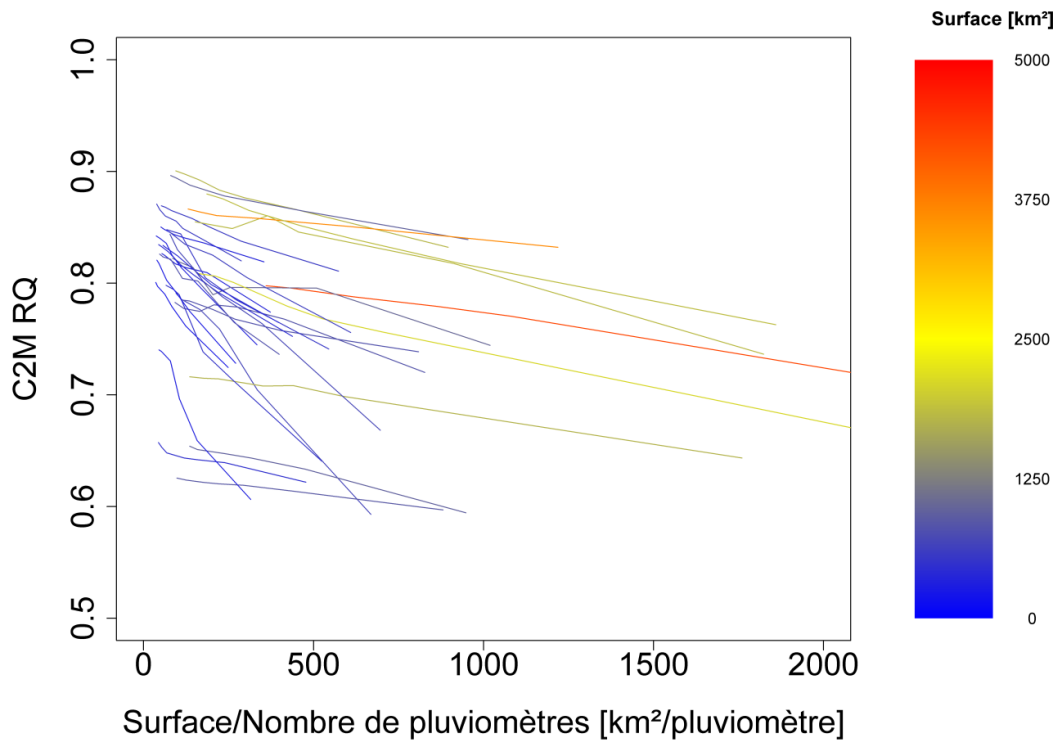


Figure 86. Valeur médiane des performances du modèle GR4J en fonction de la surface par pluviomètre pour les 30 bassins versants mais également en fonction de la valeur de la surface du bassin versant.

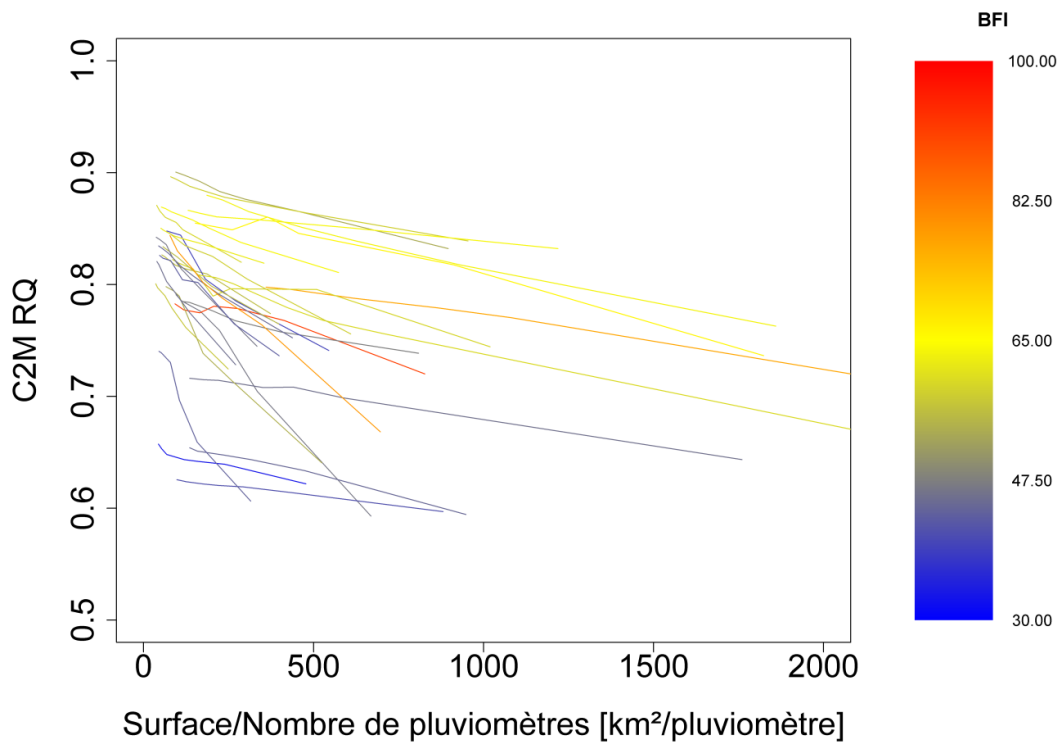


Figure 87. Valeur médiane des performances du modèle GR4J en fonction de la surface par pluviomètre pour les 30 bassins versants mais également en fonction de l'indice du débit de base du bassin versant.

9.4.4 Sensibilité des performances par rapport à la diminution de la densité du réseau pluviométrique

A partir de la Figure 77, nous nous intéressons maintenant à traduire en terme de réduction de densité, une tolérance de performances du modèle GR4J de 10%, c'est-à-dire essayer de répondre, à partir de notre échantillon de bassins, à la question : quelle serait la conséquence d'une baisse de 10% de performances du modèle en termes de densité pluviométrique ?

La Figure 88 compare les densités pluviométriques correspondant aux performances de modélisation optimum et à celles correspondant aux performances de modélisation optimum réduites de 10%. De manière générale, nous observons sur ce graphique que la densité pluviométrique diminue lorsqu'on autorise une baisse de performances de 10%. En traçant la droite de régression sur cet échantillon de points, nous constatons que sa pente est environ de 5. Cela signifie que réduire de 10% les performances du modèle équivaut à augmenter le nombre de km² par pluviomètre d'environ un facteur 5, soit de diminuer la densité pluviométrique d'environ un facteur 5.

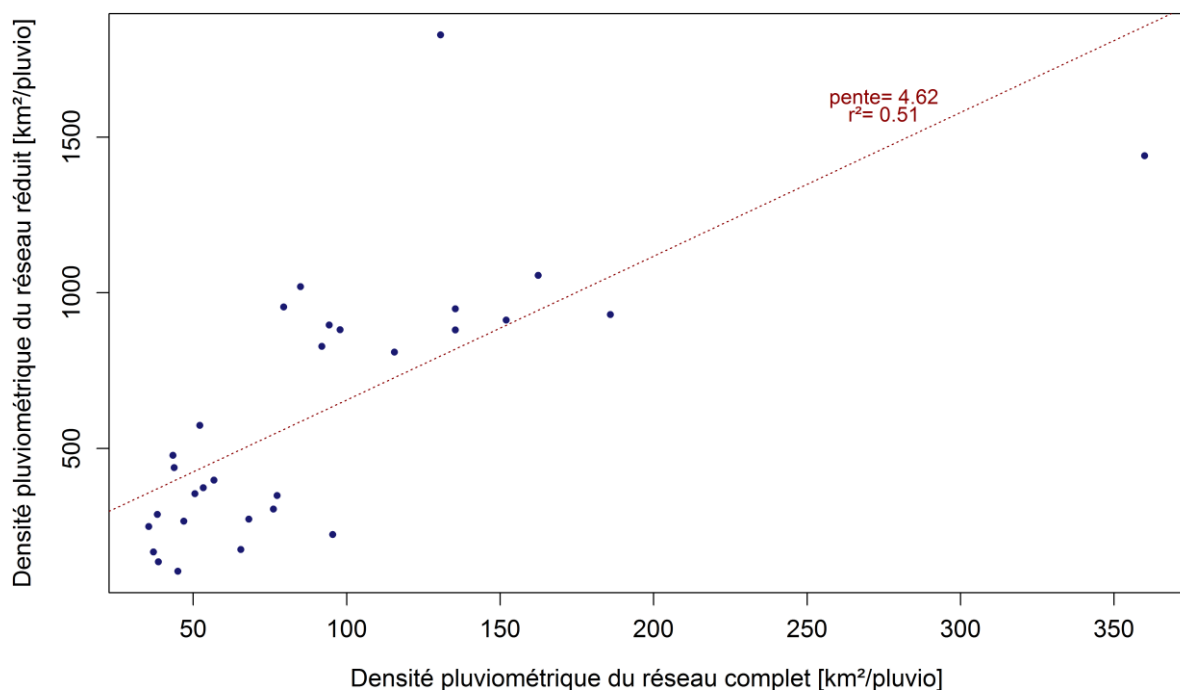


Figure 88. Comparaison des densités pluviométriques correspondant aux résultats de modélisation optimum (réseau complet) et celles correspondant aux performances de modélisation optimum réduites de 10% (réseau réduit).

9.5. Synthèse

Le but de ce chapitre était d'évaluer l'impact de la densité spatiale du réseau pluviométrique sur divers calculs hydrologiques : la simulation d'une chronique de débit journalier avec le modèle GR4J, le calcul du module, l'estimation des débits de crues décennale et centennale, et l'estimation de caractéristiques de débit d'étiage. Nous avons testé et présenté différentes densités de réseau pluviométrique pour chaque bassin versant.

Dans un premier temps, nous avons présenté et détaillé les résultats de deux bassins exemple : la Vézère à Campagne et le Madon à Pulligny. Ces résultats nous ont montré que les bassins présentaient des comportements différents.

Dans un second temps, nous avons essayé de généraliser les résultats sur l'ensemble de l'échantillon des 30 bassins versants pour chacun des calculs. Nous avons pu constater que les performances médianes du modèle GR4J diminuent lorsque le réseau pluviométrique devient moins dense pour tous les bassins, mais avec des vitesses de décroissance variables d'un bassin à l'autre. Concernant le calcul du module, pour la majorité des bassins, l'erreur relative médiane (Q_{sim}/Q_{obs}) s'écarte de la valeur idéale 1 lorsque la densité du réseau diminue. Pour les débits de crues décennale et centennale, une même tendance à la croissance de l'erreur relative est observée. Nous avons donc un débit simulé qui est de plus en plus grand en comparaison au débit observé. Ceci pourrait s'expliquer par le fait que quand les données du réseau pluviométrique deviennent de plus en plus ponctuelles, la pluie est moins lissée par la pluie d'autres stations. Les pluies ponctuelles auront alors un poids plus fort sur le calcul de la pluie de bassin et cela se traduit par une tendance à un excès en termes de débits de pointe. Pour les calculs des quantiles de bas débits, nous remarquons également une tendance à la surestimation.

Enfin, nous avons cherché d'autres explications concernant les vitesses de décroissances des performances du modèle GR4J. Parmi les caractéristiques de bassin testées, il semblerait que la variabilité spatiale de la pluie ait une influence un peu plus marquée sur la vitesse de décroissance des performances. En effet, les résultats ont dans l'ensemble montré que plus la variabilité spatiale de la pluie sur un bassin (représentée ici par la pente du variogramme climatologique du bassin) était importante, plus vite les performances du modèle GR4J semblaient se dégrader lorsque la densité du réseau pluviométrique diminuait.

Conclusion générale

Les calculs hydrologiques utiles à la société sont nombreux et variés mais nécessitent des données hydrométéorologiques en qualité et en quantité suffisantes pour assurer l'exactitude de ces calculs. Cette notion d'exactitude est d'autant plus importante lorsque ces outils hydrologiques sont destinés à être utilisés en opérationnel et à servir pour la prise de décisions importantes pour la sécurité publique. De ce fait, une bonne représentativité des informations hydrométéorologiques (en qualité et en quantité) est capitale afin de limiter les incertitudes des calculs hydrologiques, liées à l'échantillonnage de ces variables dans le temps et dans l'espace.

Toutefois, le coût des réseaux hydrométéorologiques peut être lourd pour leurs gestionnaires. Le bon sens exige que les coûts des réseaux soient comparés aux bénéfices associés aux calculs hydrologiques. Dans le premier chapitre, nous avons vu que les ratios coût-bénéfice associés aux réseaux hydrométriques pouvaient atteindre la valeur non négligeable d'1:19 au Canada [Azar *et al.*, 2003]. Toutefois, cette thèse n'avait pas de vocation économique. Il est en effet très compliqué de déterminer le bénéfice lié à la diminution de l'incertitude des calculs hydrologiques. Nos travaux se donc essentiellement portés sur l'évaluation des incertitudes de ces calculs selon la densité des réseaux hydro-pluviométriques.

Cette thèse présentait un double objectif : étudier l'impact de la densité des réseaux hydrométriques (en diminuant la densité spatiale du réseau et en testant l'échantillonnage temporel des séries disponibles de débit), mais également celui de la densité des réseaux pluviométriques (en diminuant la densité spatiale du réseau) sur les performances de divers calculs hydrologiques (simulation de débit au pas de temps journalier, estimation de débit de crue extrême et estimation de caractéristiques d'étiage). Cette thèse s'est donc articulée en deux parties : la première concernant le réseau hydrométrique et la seconde, le réseau pluviométrique.

Les synthèses bibliographiques nous ont permis de passer en revue des résultats plutôt dispersés et réalisés bien souvent dans le cadre de monographies, offrant peu de possibilités de généralisation. C'est donc dans le but de pouvoir généraliser nos résultats que nous nous sommes intéressés à un échantillon aussi large que possible de bassins versants (609 et 30 respectivement dans les deux parties de la thèse). Ceci a permis de couvrir des situations hydro-pluviométriques très variées et réparties sur le territoire métropolitain et d'essayer de produire des conclusions générales.

Nous avons aussi eu l'ambition de couvrir une certaine diversité en termes de calculs : les besoins en matière de conception de réseaux pour l'estimation des crues n'étant pas obligatoirement les mêmes que ceux destinés à l'estimation des caractéristiques d'étiage. Par ailleurs, au cours de cette thèse, il semblait difficile d'utiliser et de comparer toutes les méthodes de calcul existantes. Nous nous sommes intéressés en particulier, aux méthodes à base de modélisation pluie-débit ainsi qu'à une formule appartenant à la famille des « Budyko-curve », qui suscite un regain d'intérêt en hydrologie ces 20 dernières années.

Nous proposons d'articuler cette conclusion générale autour de deux parties : une première partie concernant le réseau hydrométrique et une seconde partie portant sur le réseau pluviométrique.

Réseau hydrométrique

L'objectif principal des réseaux hydrométriques est d'obtenir suffisamment d'information sur la disponibilité des ressources en eau, leur distribution géographique et leur variabilité dans le temps, afin de répondre à certaines problématiques opérationnelles telles que la prévision des crues et des étiages, le dimensionnement d'ouvrages hydrotechniques, le maintien de la capacité à produire l'eau potable, etc.

En hydrologie, le cas idéal est d'avoir une station hydrométrique installée depuis de nombreuses années, avec une chronique de données sans lacunes et de bonne qualité. Il va de soi que rien ne peut être meilleur que les observations. Il est ainsi nécessaire de bien réfléchir avant de fermer une station ancienne. L'utilisation de données mesurées sur quelques années peut également être utile pour permettre une reconstitution de longues séries de débits à l'aide de modèles hydrologiques. Toutefois, ces chroniques de débit ne sont pas toujours disponibles au point d'intérêt du cours d'eau et les incertitudes liées aux méthodes alternatives permettant l'estimation de ces débits en des points non jaugés sont importantes et dépendantes de la disponibilité spatio-temporelle des données hydrométriques.

Nous présentons quelques recommandations concernant les pratiques à mettre en œuvre afin de limiter les incertitudes des calculs en ces points où aucune station pérenne n'a été préalablement installée, et de fournir des estimations de leurs incertitudes associées.

Cas d'un bassin totalement non jaugé

Dans le cas où aucune donnée de débit n'est disponible au point stratégique du cours d'eau, diverses méthodes de régionalisation peuvent être appliquées afin d'estimer le débit en ces points non jaugés. La méthode adoptée au cours de cette thèse est celle basée sur la proximité spatiale des bassins versants. Cette méthode consiste à transférer au bassin cible, les paramètres des modèles hydrologiques des bassins les plus proches.

Au cours de cette thèse, nous avons mis en place une méthode permettant d'évaluer l'impact de la densité spatiale du réseau hydrométrique sur les performances de cette méthode de régionalisation et sur les divers calculs hydrologiques présentés précédemment. Pour cela, nous avons appliqué la méthode appelée « désert hydrométrique » qui consiste à écarter progressivement les bassins donneurs les plus proches : les paramètres sont transférés de voisins qui sont de plus en plus éloignés du bassin versant cible non-jaugé, en fixant une limite inférieure en dessous de laquelle les voisins sont écartés. Cette étude a montré que pour tous ces calculs, la performance du processus de régionalisation diminue lorsque le réseau de bassins voisins devient moins dense (lorsque l'on supprime les bassins voisins les plus proches). Toutefois, cette chute d'efficacité liée à l'augmentation de la distance des plus proches voisins est moindre en comparaison de celle attribuée à la méthode de régionalisation elle-même. Ceci confirme que rien ne vaut les observations de débit en un point donné. De plus, la quantification de la perte de performances est variable entre les bassins et les différentes variables hydrologiques et aucune explication spatiale (proximité du plus proche voisin, zone géographique de la zone d'étude) ne semble se démarquer.

Par ailleurs, pour une application opérationnelle sur un bassin versant non-jaugé donné, en connaissant la distance à laquelle se trouve les plus proches voisins, les courbes obtenues au chapitre

5 (Figure 39 à Figure 45) donnent une estimation des gammes de performances auxquelles on peut s'attendre sur un bassin versant non jaugé et pour divers calculs hydrologiques. Ces « abaques » d'ambition généraliste permettent d'obtenir une indication de la valeur attendue de l'efficacité des calculs régionalisés en fonction de la distance à laquelle se situent les bassins voisins, avec un intervalle d'incertitude sur la valeur attendue.

Cas d'un bassin partiellement jaugé

Si aucune station n'a été installée au point d'intérêt du cours d'eau, quelques mesures ponctuelles peuvent venir aider à l'utilisation d'une méthode de régionalisation. Ce point a été étudié au cours de cette thèse dans le chapitre 6, où nous avons notamment repris les travaux de thèse de *Rojas-Serna* [2005], ce qui a permis de confirmer l'intérêt de mettre en place une méthode combinant une information partielle fournie par les quelques mesures de débit disponibles sur le bassin cible, avec une information régionale, fournie par les paramètres des bassins voisins. Cette étude montre que la méthode proposée peut efficacement exploiter les deux sources d'information, ce qui est particulièrement intéressant quand peu de mesures sont disponibles sur un bassin versant. En effet, même lorsqu'une seule mesure de débit est disponible, des améliorations de performances peuvent être obtenues en comparaison de la méthode de régionalisation utilisée sous sa forme la plus simple.

Au cours de ce chapitre, nous avons également montré que certaines gammes de débit étaient plus informatives selon l'objectif hydrologique visé (estimation des bas, moyens ou hauts débits), lors de l'application de cette méthode. Par ailleurs, nous avons également montré qu'utiliser un plus petit nombre de débits mesurés aléatoirement permettait d'obtenir de meilleures performances que l'utilisation d'un grand nombre de débits équivalents.

Réseau pluviométrique

Dans la deuxième partie de la thèse, nous avons évalué l'impact de la densité spatiale du réseau pluviométrique sur ces calculs hydrologiques. Nous avons testé et présenté différentes densités de réseau pluviométrique pour chaque bassin versant. Nous avons essayé de faire émerger des comportements « génériques » avec difficulté, que ce soit pour l'ensemble des bassins versants ou l'ensemble des calculs envisagés. L'étude a montré que selon l'objectif hydrologique considéré, les résultats variaient :

- les performances médianes du **modèle GR4J** diminuent lorsque le réseau pluviométrique devient moins dense pour tous les bassins, mais avec des vitesses de décroissance des performances variables d'un bassin à l'autre. Cette différence de vitesses de décroissances de performances peut être expliquée par la variabilité spatiale de la pluie du bassin. Les résultats ont dans l'ensemble montré que plus la variabilité spatiale de la pluie sur un bassin (caractérisée dans cette thèse par la pente du variogramme climatologique du bassin) était importante, plus les performances du modèle GR4J se dégradaient vite lorsque la densité du réseau pluviométrique diminuait. Ces résultats sont en accord avec ceux trouvés par *Lobligeois* [2014] au cours de sa thèse, puisque dans son étude, l'utilisation de l'information radar (ajout d'une information pluviométrique) améliore les simulations pluie-débit pour les bassins avec une forte variabilité spatiale de la pluie. Au contraire, pour les

bassins présentant une faible hétérogénéité spatiale de la pluie, le bénéfice apporté par l'observation radar est plus contrasté.

- Concernant le calcul du **module**, pour l'ensemble des bassins, l'erreur relative médiane (Q_{sim}/Q_{obs}) s'écarte de la valeur idéale 1 lorsque la densité du réseau diminue.

- Pour les **débits de crues décennale et centennale**, une même tendance à la croissance de l'erreur relative est observée. Nous avons donc un débit simulé qui est de plus en plus grand en comparaison du débit observé. Ceci pourrait s'expliquer par le fait que quand le réseau pluviométrique devient de moins en moins dense, la pluie est moins lissée par la pluie d'autres stations. Les pluies ponctuelles auront alors un poids plus fort sur le calcul de la pluie de bassin et cela se traduit par une tendance à surestimer les débits de pointe.

- Pour les calculs des **quantiles de bas débits**, nous remarquons également une tendance à la surestimation des débits quand la densité du réseau pluviométrique diminue. Malheureusement, nous n'avons pu identifier les raisons de ces résultats.

Limites de notre approche et perspectives

L'objectif de cette thèse était de quantifier les conséquences hydrologiques de la diminution des densités des réseaux d'observation. Cette thèse ne visait ainsi pas à l'optimisation ou la détermination de configuration idéale des réseaux hydrométéorologiques en France. L'idéal serait pour chaque calcul hydrologique considéré, d'essayer de déterminer la densité minimale des réseaux hydro-pluviométriques à respecter pour garder un certain niveau de performances à budget global donné.

D'autre part, le but de cette thèse n'était pas non plus de développer une nouvelle méthode de régionalisation ou de déterminer quelle méthode était la plus performante avec le réseau hydrométrique actuel, mais seulement d'évaluer à partir d'une méthode (ici celle basée sur la proximité spatiale des bassins), quel était l'impact de la baisse de la densité de ce réseau sur les performances des calculs hydrologiques. Il pourrait être intéressant d'évaluer cet impact en utilisant une autre méthode de régionalisation, comme par exemple celle basée sur la similarité physique. Ceci consisterait à écarter progressivement les bassins les plus proches physiquement du bassin non jaugé et d'en évaluer les conséquences sur les calculs hydrologiques. On pourrait ainsi déterminer quelle méthode de régionalisation est la plus sensible à la densité du réseau hydrométrique.

D'un point de vue davantage opérationnel, la méthode présentée dans le chapitre 6 utilisant les jaugeages ponctuels afin d'améliorer les simulations de débit sur des bassins non jaugés, pourrait également être utilisée pour la prévision des crues. Ce chapitre pourrait, en effet, amorcer une réflexion sur la mise au point d'une stratégie opérationnelle itérative de jaugeages. Ceci consisterait, à partir de paramètres estimés en non jaugés, à déterminer à l'aide de prévision, le moment le plus opportun d'aller effectuer un premier jaugeage relativement à l'objectif hydrologique prédéfini. Ensuite, à l'aide de ce premier jaugeage, un ajustement du jeu de paramètres pourrait être effectué à l'aide de la méthode présentée au chapitre 6, et le processus pourrait être répété autant de fois que nécessaire.

Par ailleurs, un autre aspect intéressant qui mériterait d'être étudié est l'amplification des réseaux (en évaluant des stations hydrométéorologiques « virtuelles »). Pour le réseau pluviométrique, nous

avons pensé à utiliser la réanalyse Comephore, qui combine données pluviométriques et données radar [Tabary *et al.*, 2012]. La méthode aurait consisté à sélectionner des pixels de cette réanalyse dans le but d'être utilisés comme pseudo-pluviomètres additionnels. Ces pluviomètres additionnels auraient reflété la lame d'eau prise en compte par les radars et auraient ainsi apporté une information de pluie supplémentaire. A l'aide de cette méthode, une approche permettant de configurer de manière optimale le réseau pluviométrique afin de limiter les incertitudes des calculs hydrologiques tout en respectant un certain budget, aurait pu alors être mise en place. Toutefois, cette étude a été écartée de cette thèse en raison de la trop grande dépendance de la lame d'eau Comephore avec la lame d'eau issue des pluviomètres seuls : au cours de sa thèse, *Lobligeois* [2014] a montré que l'utilisation de la réanalyse dans la simulation d'événements de crue par un modèle pluie-débit, n'apportait pas forcément de meilleures performances au modèle hydrologique par rapport au krigeage des pluviomètres seuls (52% des événements de crue sont mieux simulés avec l'utilisation des pluviomètres contre 48% avec la réanalyse). Il précise également que les bassins versants présentant de fortes hétérogénéités spatiales des pluies présentaient aussi une densité de pluviomètres élevée, et que les événements qui contenaient le plus d'information radar étaient aussi très bien mesurés par un réseau de mesure ponctuelle très dense. Ceci pouvait ainsi expliquer le contraste des résultats obtenus avec les deux sources de pluie utilisées. Par contre, l'utilisation seule de l'information radar pourrait faciliter la sélection de postes pluviométriques virtuels, ayant peu d'influence avec le réseau déjà présent, et pourrait ainsi permettre de déterminer une configuration optimale du réseau pluviométrique en fonction de tel ou tel type d'événements pluvieux et en fonction du budget.

Enfin, la mise en œuvre d'une méthode permettant une gestion coordonnée des deux réseaux de mesure mériterait également d'être étudiée plus en détail. Ce n'est certes pas une idée neuve. Par exemple, *Linsley et Crawford* [1965] proposaient d'installer des stations hydrométriques sur de courtes durées, le temps qu'une relation entre la pluie issue des stations pluviométriques et le débit de la station hydrométrique soit établie sans trop d'incertitudes à l'aide de modèle ; la station serait par la suite déplacée à un autre endroit et ainsi de suite. Il serait tout à fait envisageable de reprendre (mais avec nos outils actuels) ce type d'étude dans le but d'étendre la couverture hydrométrique avec un budget raisonnable, à l'aide de stations volantes pour des objectifs de connaissance de la ressource en eau, par exemple.

Bibliographie

- Adhikary, S. K., A. G. Yilmaz, et N. Muttill (2014), Optimal design of rain gauge network in the Middle Yarra River catchment, Australia, *Hydrol. Process.*, n/a-n/a, doi: 10.1002/hyp.10389.
- Anaman, K. A., D. J. Thampapillai, A. Henderson-Sellers, P. F. Noar, et P. J. Sullivan (1995), Methods for assessing the benefits of meteorological services in Australia, *Meteorological Applications*, 2(1), 17-29, doi: 10.1002/met.5060020104.
- Anaman, K. A., S. C. Lellyett, L. Drake, R. J. Leigh, A. Henderson-Sellers, P. F. Noar, P. J. Sullivan, et D. J. Thampapillai (1998), Benefits of meteorological services: evidence from recent research in Australia, *Meteorological Applications*, 5(2), 103-115, doi: 10.1017/S1350482798000668.
- Anctil, F., C. Perrin, et V. Andreassian (2004), Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models, *Environ. Modell. Softw.*, 19(4), 357-368, doi: 10.1016/s1364-8152(03)00135-x.
- Anctil, F., N. Lauzon, V. Andreassian, L. Oudin, et C. Perrin (2006), Improvement of rainfall-runoff forecasts through mean areal rainfall optimization, *Journal of Hydrology*, 328(3-4), 717-725, doi: 10.1016/j.jhydrol.2006.01.016.
- Andréassian, V., et C. Perrin (2012), On the ambiguous interpretation of the Turc-Budyko nondimensional graph, *Water Resources Research*, 48, doi: 10.1029/2012wr012532.
- Andréassian, V., C. Perrin, E. Parent, et A. Bárdossy (2010a), The Court of Miracles of Hydrology: can failure stories contribute to hydrological science?, *Hydrological Sciences Journal*, 55(6), 849-856.
- Andréassian, V., C. Perrin, E. Parent, et A. Bardossy (2010b), The Court of Miracles of Hydrology: can failure stories contribute to hydrological science?, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 55(6), 849-856, doi: 10.1080/02626667.2010.506050.
- Andréassian, V., J. Lerat, N. Le Moine, et C. Perrin (2012a), Neighbors: nature's own hydrological models, *Journal of Hydrology* 414/415, 49-58, doi: 10.1016/j.jhydrol.2011.10.007.
- Andréassian, V., C. Perrin, C. Michel, I. Usart-Sanchez, et J. Lavabre (2001), Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models, *Journal of Hydrology*, 250(1-4), 206-223, doi: 10.1016/s0022-1694(01)00437-1.
- Andréassian, V., N. Le Moine, C. Perrin, M. H. Ramos, L. Oudin, T. Mathevet, J. Lerat, et L. Berthet (2012b), All that glitters is not gold: the case of calibrating hydrological models, *Hydrol. Process.*, 26, 2206-2210, doi: 10.1002/hyp.9264.
- Andréassian, V., C. Perrin, L. Berthet, N. Le Moine, J. Lerat, C. Loumagne, L. Oudin, T. Mathevet, M. H. Ramos, et A. Valery (2009), HESS Opinions 'Crash tests for a standardized evaluation of hydrological models', *Hydrol. Earth Syst. Sci.*, 13(10), 1757-1764.
- Arnaud, P., J. Lavabre, C. Fouchier, S. Diss, et P. Javelle (2011), Sensitivity of hydrological models to uncertainty in rainfall input, *Hydrological Sciences Journal*, 56(3), 397-410, doi: 10.1080/02626667.2011.563742.
- Arora, V. K. (2002), The use of the aridity index to assess climate change effect on annual runoff, *Journal of Hydrology*, 265, 164-177.

- Arsenault, R., et F. Brissette (2013), Determining the Optimal Spatial Distribution of Weather Station Networks for Hydrological Modeling Purposes Using RCM Datasets: An Experimental Approach, *Journal of Hydrometeorology*, 15(1), 517-526, doi: 10.1175/JHM-D-13-088.1.
- Arsenault, R., et F. P. Brissette (2014), Continuous streamflow prediction in ungauged basins: The effects of equifinality and parameter set selection on uncertainty in regionalization approaches, *Water Resources Research*, 50(7), 6135-6153, doi: 10.1002/2013WR014898.
- Aschwanden, H., et C. Kan (1999), Le débit d'étiage Q347 - Etat de la question., 134 pp, Service hydrologique et géologique national (Suisse).
- Azar, J., D. Sellars, et D. Schroeter (2003), Water Quantity Monitoring in British Columbia: A Business Review of the BC Hydrometric Programs, BC Ministry of Sustainable Resources Management.
- Baccour, H., M. Slimani, et C. Cudennec (2012), Structures spatiales de l'évapotranspiration de référence et des variables climatiques corrélées en Tunisie, *Hydrological Sciences Journal*, 57(4), 818-829, doi: 10.1080/02626667.2012.672986.
- Bagrov, N. (1953), On long-term average of evapotranspiration from land surface (O srednem mnogoletnem isparenii s poverkhnosti souchi), *Meteorologia i Gidrologia*, 10, 20-25.
- Bardossy, A., et T. Das (2008), Influence of rainfall observation network on model calibration and application, *Hydrol. Earth Syst. Sci.*, 12(1), 77-89.
- Bell, V. A., et R. J. Moore (2000), The sensitivity of catchment runoff models to rainfall data at different spatial scales, *Hydrol. Earth Syst. Sci.*, 4(4), 653-667.
- Berne, A., G. Delrieu, et B. Boudevillain (2009), Variability of the spatial structure of intense Mediterranean precipitation, *Adv. Water Resour.*, 32(7), 1031-1042, doi: <http://dx.doi.org/10.1016/j.advwatres.2008.11.008>.
- Berne, A., G. Delrieu, J.-D. Creutin, et C. Obled (2004), Temporal and spatial resolution of rainfall measurements required for urban hydrology, *Journal of Hydrology*, 299(3-4), 166-179, doi: <http://dx.doi.org/10.1016/j.jhydrol.2004.08.002>.
- Berthet, L., V. Andreassian, C. Perrin, et C. Loumagne (2010), How significant are quadratic criteria? Part 1. How many years are necessary to ensure the data-independence of a quadratic criterion value?, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 55(6), 1051-1062, doi: 10.1080/02626667.2010.505890.
- Beven, K., et I. Westerberg (2011), On red herrings and real herrings: disinformation and information in hydrological inference, *Hydrol. Process.*, 25(10), 1676-1680, doi: 10.1002/hyp.7963.
- Bhatt, V. K., et A. K. Tiwari (2008), Estimation of peak streamflows through channel geometry, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 53(2), 401-408, doi: 10.1623/hysj.53.2.401.
- Boldetti, G. (2012), Estimation of the parameters of hydrological models on ungauged basins: a comparison of direct and indirect approaches, PhD thesis, 207 pp, AgroParisTech - Irstea (Antony).
- Boldetti, G., M. Riffard, V. Andréassian, et L. Oudin (2010), Data-set cleansing practices and hydrological regionalization: is there any valuable information among outliers?, *Hydrological*

- Sciences Journal-Journal Des Sciences Hydrologiques*, 55(6), 941-951, doi: 10.1080/02626667.2010.505171.
- Bras, R. L., et I. Rodriguez-Iturbe (1976), Rainfall network design for runoff prediction, *Water Resources Research*, 12(6), 1197-1208.
- Bras, R. L., et R. Colon (1978), Time-averaged areal mean of precipitation - estimation and network design, *Water Resources Research*, 14(5), 878-888, doi: 10.1029/WR014i005p00878.
- Brath, A., A. Montanari, et E. Toth (2004), Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model, *Journal of Hydrology*, 291(3-4), 232-253, doi: 10.1016/j.jhydrol.2003.12.044.
- Brigode, P., V. Andréassian, F. Bourgin, et F. Lobligelois (2014), Reconstitution de 55 années de débits sur plus de 10 000 points non-jaugés en France par régionalisation d'un modèle pluie-débit, *Rapport d'étude*, Irstea & ONEMA.
- Budyko, M. I. (1974), *Climate and Life*, edited, p. 508, Academic Press, New York.
- Caracciolo, D., E. Arnone, et L. Noto (2014), Influence of Spatial Precipitation Sampling on Hydrological Response at the Catchment Scale, *J. Hydrol. Eng.*, 19(3), 544-553, doi: doi:10.1061/(ASCE)HE.1943-5584.0000829.
- Catalogne, C. (2012), Amélioration des méthodes de prédétermination des débits de référence d'étiage en sites peu ou pas jaugés, PhD thesis, 285 pp, Université de Grenoble - Irstea (Lyon).
- Catalogne, C., et E. Sauquet (2011), Interpolation des débits de référence d'étiage, *Rapport d'étude*, 35 pp, Cemagref & ONEMA.
- Catalogne, C., E. Sauquet, et M. Lang (2014), Valorisation des données de jaugeages épisodiques pour l'estimation du débit de référence d'étiage QMNA5, *La Houille Blanche*(4), 78-87.
- Choudhury, B. J. (1999), Evaluation of an empirical equation for annual evaporation using field observations and results from a biophysical model, *Journal of Hydrology*, 216(1-2), 99-110, doi: 10.1016/S0022-1694(98)00293-5.
- Cordery, I., et P. S. Cloke (1994), Benefits of flow data for flood-protection design, *Journal of the Institution of Water and Environmental Management*, 8(1), 33-38.
- Cosandey, C., et M. Robinson (2000), *Hydrologie continentale*, Armand Colin ed., 368 pp.
- Dirks, K. N., J. E. Hay, C. D. Stow, et D. Harris (1998), High-resolution studies of rainfall on Norfolk Island Part II: Interpolation of rainfall data, *Journal of Hydrology*, 208(3-4), 187-193, doi: 10.1016/S0022-1694(98)00155-3.
- Dong, X. H., C. M. Dohmen-Janssen, et M. J. Booij (2005), Appropriate spatial sampling of rainfall for flow simulation, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 50(2), 279-298, doi: 10.1623/hysj.50.2.279.61801.
- Donohue, R., M. Roderick, et T. McVicar (2011), Assessing the differences in sensitivities of runoff to changes in climatic conditions across a large basin, *J. Hydrol.*, 406(3-4), 234-244, doi: 10.1016/j.jhydrol.2011.07.003.

- Donohue, R. J., M. L. Roderick, et T. R. McVicar (2012), Roots, storms and soil pores: Incorporating key ecohydrological processes into Budyko's hydrological model, *Journal of Hydrology*, 436, 35-50, doi: 10.1016/j.jhydrol.2012.02.033.
- Dooge, J. C. I. (1992), Sensitivity of runoff to climate change - A Hortonian approach, *Bull. Amer. Meteorol. Soc.*, 73(12), 2013-2024, doi: 10.1175/1520-0477(1992)073<2013:sortcc>2.0.co;2.
- Dooge, J. C. I., M. Bruen, et B. Parmentier (1999), A simple model for estimating the sensitivity of runoff to long-term changes in precipitation without a change in vegetation, *Adv. Water Resour.*, 23(2), 153-163, doi: 10.1016/s0309-1708(99)00019-6.
- Drogue, G. P., et J. Plasse (2014), How can a few streamflow measurements help to predict daily hydrographs at almost ungauged sites?, *Hydrological Sciences Journal*, 59(12), 2126-2142, doi: 10.1080/02626667.2013.865031.
- Dubreuil, P. (1968), Etude de l'extension rationnelle du réseau hydrométrique du ministère de l'agriculture, 96 pp, Ministère de l'agriculture, Service de l'hydraulique; Office de la recherche scientifique et technique Outre-Mer, Service hydrologique, Paris.
- Duncan, M. R., B. Austin, F. Fabry, et G. L. Austin (1993), The effect of gauge sampling density on the accuracy of streamflow prediction for rural catchments, *Journal of Hydrology*, 142(1-4), 445-476, doi: 10.1016/0022-1694(93)90023-3.
- Dyck, S. (1976), Applied Hydrology, part I. Calculation and regulation of the discharge of streams. , VEB Verlag für Bauwesen, Berlin.
- Edijatno, N. D. Nascimento, X. L. Yang, Z. Makhlof, et C. Michel (1999), GR3J: a daily watershed model with three free parameters, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 44(2), 263-277, doi: 10.1080/02626669909492221.
- Emmanuel, I., H. Andrieu, E. Leblois, et B. Flahaut (2012), Temporal and spatial variability of rainfall at the urban hydrological scale, *Journal of Hydrology*, 430-431(0), 162-172, doi: <http://dx.doi.org/10.1016/j.jhydrol.2012.02.013>.
- Faures, J. M., D. C. Goodrich, D. A. Woolhiser, et S. Sorooshian (1995), Impact of small-scale spatial rainfall variability on runoff modeling, *Journal of Hydrology*, 173(1-4), 309-326, doi: 10.1016/0022-1694(95)02704-s.
- Freebairn, J. W., et J. W. Zillman (2002), Economic benefits of meteorological services, *Meteorological Applications*, 9(1), 33-44, doi: 10.1017/S1350482702001044.
- Frei, T. (2010), Economic and social benefits of meteorology and climatology in Switzerland, *Meteorological Applications*, 17(1), 39-44, doi: 10.1002/met.156.
- Fu, B. (1981), On the calculation of the evaporation from land surface, *Atmospherica Sinica*, 5, 23-31.
- Gan, T. Y., E. M. Dlamini, et G. F. Biftu (1997), Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling, *Journal of Hydrology*, 192(1-4), 81-103.
- Gerrits, A. M. J., H. H. G. Savenije, E. J. M. Veling, et L. Pfister (2009), Analytical derivation of the Budyko curve based on rainfall characteristics and a simple evaporation model, *Water Resources Research*, 45, doi: 10.1029/2008wr007308.

- Ghosh, B. (1951), Random Distances Within A Rectangle and Between Two Rectangles, *Bulletin of Cal. Math. Soc.*, 43, 17-24.
- Gottschalk, L. (1993a), Correlation and covariance of runoff, *Stoch. Hydrol. Hydraul.*, 7(2), 85-101, doi: 10.1007/bf01581418.
- Gottschalk, L. (1993b), Interpolation of runoff applying objective methods, *Stoch. Hydrol. Hydraul.*, 7(4), 269-281, doi: 10.1007/bf01581615.
- Gottschalk, L., E. Leblois, et J. O. Skøien (2011), Distance measures for hydrological data having a support, *Journal of Hydrology*, 402(3-4), 415-421.
- Gottschalk, L., J. L. Jensen, D. Lundquist, R. Solantie, et A. Tollan (1979), Hydrologic regions in the Nordic countries, *Nordic Hydrology*, 10(5), 273-286.
- Gupta, H. V., H. Kling, K. K. Yilmaz, et G. F. Martinez (2009), Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *Journal of Hydrology*, 377(1-2), 80-91.
- Helloco, F. (2010), Etude nationale de rationalisation du réseau de mesure pluviométrique pour l'hydrologie de la Prévision des Crues par les services de l'Etat, SCHAPI.
- Hendrick, R. L., et G. H. Comer (1970), Space variations of precipitation and implications for raingage network design, *Journal of Hydrology*, 10(2), 151-163, doi: [http://dx.doi.org/10.1016/0022-1694\(70\)90185-X](http://dx.doi.org/10.1016/0022-1694(70)90185-X).
- Herbaud, J. (1969), Essai sur les problèmes de rationalisation de réseaux - Mise en oeuvre sur un réseau pluviométrique, Cahiers O.R.S.T.O.M, Série Hydrologique.
- Hickel, K., et L. Zhang (2006), Estimating the impact of rainfall seasonality on mean annual water balance using a top-down approach, *Journal of Hydrology*, 331(3-4), 409-424, doi: 10.1016/j.jhydrol.2006.05.028.
- Hrachowitz, M., et M. Weiler (2011), Uncertainty of Precipitation Estimates Caused by Sparse Gauging Networks in a Small, Mountainous Watershed, *J. Hydrol. Eng.*, 16(5), 460-471, doi: 10.1061/(asce)he.1943-5584.0000331.
- Hrachowitz, M., et al. (2013), A decade of Predictions in Ungauged Basins (PUB)a review, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 58(6), 1198-1255, doi: 10.1080/02626667.2013.803183.
- Huff, F. A. (1970), Sampling errors in measurement of mean precipitation, *Journal of Applied Meteorology*, 9(1), 35-44.
- IPCC (2001), Special report on the regional impacts of climate change: An assessment of vulnerability, 517 pp, Cambridge Univ. Press, UK.
- Jacoby, S., et D. R. Dawdy (1973), The relation of rainfall network density to accuracy of runoff prediction in a mountainous basin, *WMO Publication 326*, 1.
- James, L. D. (1972), Hydrologic modeling, parameter estimation, and watershed characteristics, *Journal of Hydrology*, 17(4), 283-307.

- Juston, J., J. Seibert, et P. O. Johansson (2009), Temporal sampling strategies and uncertainty in calibrating a conceptual hydrological model for a small boreal catchment, *Hydrol. Process.*, 23(21), 3093-3109, doi: 10.1002/hyp.7421.
- Kim, U., et J. J. Kaluarachchi (2009), Hydrologic model calibration using discontinuous data: an example from the upper Blue Nile River Basin of Ethiopia, *Hydrol. Process.*, 23(26), 3705-3717, doi: 10.1002/hyp.7465.
- Klemes, V. (1986), Operational testing of hydrological simulation-models, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 31(1), 13-24, doi: 10.1080/02626668609491024.
- Konz, M., S. Uhlenbrook, L. Braun, A. Shrestha, et S. Demuth (2007), Implementation of a process-based catchment model in a poorly gauged, highly glacierized Himalayan headwater, *Hydrol. Earth Syst. Sci.*, 11(4), 1323-1339.
- Krajewski, W. F., G. J. Ciach, et E. Habib (2003), An analysis of small-scale rainfall variability in different climatic regimes, *Hydrological Sciences Journal*, 48(2), 151-162, doi: 10.1623/hysj.48.2.151.44694.
- Krige, D. (1951), A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand, *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, 52(6), 119-139, doi: citeulike-article-id:3070804-doi: 10.2307/3006914.
- Laaha, G., J. O. Skoien, et G. Blöschl (2014), Spatial prediction on river networks: comparison of top-kriging with regional regression, *Hydrol. Process.*, 28(2), 315-324, doi: 10.1002/hyp.9578.
- Lang, M., et J. Lavabre (2007), *Estimation de la crue centennale pour les plans de prévention des risques d'inondations*, Quae ed., 232 pp.
- Lazo, J., et L. Chestnut (2002), Economic value of current and improved weather forecasts in the US household sector, Report prepared for the National Oceanic and Atmospheric Administration. Stratus consulting INC. Boulder, Colorado, USA.
- Le Moine, N., V. Andréassian, C. Perrin, et C. Michel (2007), How can rainfall-runoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments, *Water Resour. Res.*, 43(6), W06428, doi: 10.1029/2006WR005608.
- Lebel, T., G. Bastin, C. Obled, et J. D. Creutin (1987), On the accuracy of areal rainfall estimation: A case study, *Water Resources Research*, 23(11), 2123-2134, doi: 10.1029/WR023i011p02123.
- Leviäkangas, P., R. Hautala, J. Räsänen, R. Öörni, S. Sonninen, M. Hekkanen, M. Ohlström, A. Venäläinen, et S. Saku (2007), Benefits of meteorological services in Croatia, in *VTT Technical Research Centre of Finland - Research Notes 2420*. Finnish Meteorological Institute, Helsinki., edited, pp. 3-71.
- Linsley, R. K., et N. H. Crawford (1965), Coordination of precipitation and streamflow networks, paper presented at Symposium on the design of hydrological networks, IAHS, Québec.
- Liu, J., et D. W. Han (2010), Indices for calibration data selection of the rainfall-runoff model, *Water Resources Research*, 46, doi: W0451210.1029/2009wr008668.

- Lobligeois, F. (2014), Mieux connaître la distribution spatiale des pluies améliore-t-il la modélisation des crues ? Diagnostic sur 181 bassins versants français, PhD thesis, 312 pp, AgroParisTech - Irstea (Antony).
- Lobligeois, F., V. Andreassian, C. Perrin, P. Tabary, et C. Loumagne (2014), When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events, *Hydrol. Earth Syst. Sci.*, 18(2), 575-594, doi: 10.5194/hess-18-575-2014.
- Lopes, V. L. (1996), On the effect of uncertainty in spatial distribution of rainfall on catchment modelling, *CATENA*, 28(1-2), 107-119, doi: [http://dx.doi.org/10.1016/S0341-8162\(96\)00030-6](http://dx.doi.org/10.1016/S0341-8162(96)00030-6).
- Ly, S., C. Charles, et A. Degre (2011), Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium, *Hydrol. Earth Syst. Sci.*, 15(7), 2259-2274, doi: 10.5194/hess-15-2259-2011.
- Ly, S., C. Charles, et A. Degre (2013), Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale. A review, *Biotechnol. Agron. Soc.*, 17(2), 392-406.
- Magette, W. L., V. O. Shanholtz, et J. C. Carr (1976), Estimating selected parameters for Kentucky watershed model from watershed characteristics, *Water Resources Research*, 12(3), 472-476, doi: 10.1029/WR012i003p00472.
- Mason, B. J. (1966), The role of meteorology in the national economy, *Weather*, 21(11), 382-393, doi: 10.1002/j.1477-8696.1966.tb02787.x.
- Matheron, G. (1963), Principles of geostatistics, *Economic Geology*, 58(8), 1246-1266, doi: 10.2113/gsecongeo.58.8.1246.
- Mathevet, T. (2005), Quels modèles pluie-débit globaux au pas de temps horaire ? Développements empiriques et comparaison de modèles sur un large échantillon de bassins versants, PhD thesis, 463 pp, ENGREF, Paris.
- Mathevet, T., C. Michel, V. Andréassian, et C. Perrin (2006), A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins, *IAHS-AISH Publication*(307), 211-219.
- McIntyre, N., H. Lee, H. Wheeler, A. Young, et T. Wagener (2005), Ensemble predictions of runoff in ungauged catchments, *Water Resources Research*, 41(12), doi: 10.1029/2005wr004289.
- Merz, R., et G. Blöschl (2004), Regionalisation of catchment model parameters, *Journal of Hydrology*, 287(1-4), 95-123, doi: 10.1016/j.jhydrol.2003.09.028.
- Meselhe, E. A., E. H. Habib, O. C. Oche, et S. Gautam (2009), Sensitivity of Conceptual and Physically Based Hydrologic Models to Temporal and Spatial Rainfall Sampling, *J. Hydrol. Eng.*, 14(7), 711-720, doi: 10.1061/(asce)1084-0699(2009)14:7(711).
- Mezentsev, V. (1955), More on the computation of total evaporation (Yechio raz o rastchetie srednevo summarnovo ispareniiia), *Meteorologia i Gidrologia*, 5, 24-26.
- Michaud, J. D., et S. Sorooshian (1994), Effect of rainfall-sampling errors on simulations of desert flash floods, *Water Resources Research*, 30(10), 2765-2775, doi: 10.1029/94wr01273.

- Milly, P. C. D. (1994), Climate, soil water storage, and the average annual water balance, *Water Resources Research*, 30(7), 2143-2156, doi: 10.1029/94wr00586.
- Ministère des transports, de l'équipement, du tourisme et de la mer, et Ministère de l'écologie et du développement durable (2006), Circulaire relative à l'organisation de l'hydrométrie dans les DIREN et les SPC (services de prévision des crues).
- Mishra, A. K., et P. Coulibaly (2009), Developments in hydrometric network design: a review, *Rev. Geophys.*, 47, doi: Rg200110.1029/2007rg000243.
- Mishra, A. K., et P. Coulibaly (2010), Hydrometric network evaluation for Canadian watersheds, *Journal of Hydrology*, 380(3-4), 420-437, doi: 10.1016/j.jhydrol.2009.11.015.
- Mouelhi, S. (2003), Vers une chaîne cohérente de modèles pluie-débit conceptuels globaux aux pas de temps pluriannuel, annuel, mensuel et journalier, PhD thesis, 323 pp, ENGREF, Paris.
- Mouelhi, S., C. Michel, C. Perrin, et V. Andréassian (2006), Linking stream flow to rainfall at the annual time step: the Manabe bucket model revisited, *Journal of Hydrology*, 328, 283-296.
- Nash, J. E., et J. V. Sutcliffe (1970), River flow forecasting through conceptual models part I - A discussion of principles, *Journal of Hydrology*, 10(3), 282-290.
- Nicolle, P., V. Andréassian, et E. Sauquet (2013), Blending neighbor-based and climate-based information to obtain robust low-flow estimates from short time series, *Water Resources Research*, 49(12), 8017-8025, doi: 10.1002/2012WR012940.
- Obled, C., J. Wendling, et K. Beven (1994), The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data, *Journal of Hydrology*, 159(1-4), 305-333, doi: [http://dx.doi.org/10.1016/0022-1694\(94\)90263-1](http://dx.doi.org/10.1016/0022-1694(94)90263-1).
- Ol'Dekop, E. M. (1911), *On evaporation from the surface of river basins.*, University of Tartu, Tartu, Estonia.
- OMM (1972), *Casebook on Hydrological Network Design Practice*, Genève.
- Otieno, H., J. X. Yang, W. W. Liu, et D. W. Han (2014), Influence of Rain Gauge Density on Interpolation Method Selection, *J. Hydrol. Eng.*, 19(11), doi: 10.1061/(asce)he.1943-5584.0000964.
- Oudin, L., V. Andréassian, C. Perrin, et F. Anctil (2004), Locating the sources of low-pass behavior within rainfall-runoff models, *Water Resources Research*, 40(11), W11101, doi: 10.1029/2004WR003291.
- Oudin, L., V. Andréassian, J. Lerat, et C. Michel (2008a), Has land cover a significant impact on mean annual streamflow? An international assessment using 1508 catchments, *Journal of Hydrology*, 357(3-4), 303-316, doi: 10.1016/j.jhydrol.2008.05.021.
- Oudin, L., A. Kay, V. Andréassian, et C. Perrin (2010), Are seemingly physically similar catchments truly hydrologically similar?, *Water Resources Research*, 46(11), W11558, doi: 10.1029/2009WR008887.

- Oudin, L., C. Perrin, T. Mathevet, V. Andreassian, et C. Michel (2006a), Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models, *Journal of Hydrology*, 320(1-2), 62-83, doi: 10.1016/j.jhydrol.2005.07.016.
- Oudin, L., V. Andréassian, T. Mathevet, C. Perrin, et C. Michel (2006b), Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations, *Water Resources Research*, 42(7), W07410, doi: 10.1029/2005WR004636.
- Oudin, L., V. Andreassian, C. Perrin, C. Michel, et N. Le Moine (2008b), Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments, *Water Resources Research*, 44(3), doi: W03413.10.1029/2007wr006240.
- Oudin, L., F. Hervieu, C. Michel, C. Perrin, V. Andréassian, F. Anctil, et C. Loumagne (2005), Which potential evapotranspiration input for a lumped rainfall-runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling, *Journal of Hydrology*, 303(1-4), 290-306, doi: <http://dx.doi.org/10.1016/j.jhydrol.2004.08.026>.
- Palmer, T. N. (2002), The economic value of ensemble forecasts as a tool for risk assessment: From days to decades, *Quarterly Journal of the Royal Meteorological Society*, 128(581), 747-774, doi: 10.1256/0035900021643593.
- Parajka, J., R. Merz, et G. Blöschl (2005), A comparison of regionalisation methods for catchment model parameters, *Hydrol. Earth Syst. Sci.*, 9(3), 157-171.
- Parajka, J., A. Viglione, M. Rogger, J. L. Salinas, M. Sivapalan, et G. Blöschl (2013), Comparative assessment of predictions in ungauged basins – Part 1: Runoff-hydrograph studies, *Hydrol. Earth Syst. Sci.*, 17(5), 1783-1795, doi: 10.5194/hess-17-1783-2013.
- Pearson, C. P. (1998), Changes to New Zealand's national hydrometric network in the 1990s, *Journal of Hydrology New Zealand*, 37(1), 1-17.
- Perrin, C., C. Michel, et V. Andréassian (2003), Improvement of a parsimonious model for streamflow simulation, *Journal of Hydrology*, 279(1-4), 275-289, doi: 10.1016/s0022-1694(03)00225-7.
- Perrin, C., L. Oudin, V. Andreassian, C. Rojas-Serna, C. Michel, et T. Mathevet (2007), Impact of limited streamflow data on the efficiency and the parameters of rainfall-runoff models, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 52(1), 131-151, doi: 10.1623/hysj.52.1.131.
- Porporato, A., E. Daly, et I. Rodriguez-Iturbe (2004), Soil water balance and ecosystem response to climate change, *Am. Nat.*, 164(5), 625-632, doi: 10.1086/424970.
- Potter, N. J., et L. Zhang (2009), Interannual variability of catchment water balance in Australia, *Journal of Hydrology*, 369, 120-129.
- Potter, N. J., L. Zhang, P. C. D. Milly, T. A. McMahon, et A. J. Jakeman (2005), Effects of rainfall seasonality and soil moisture capacity on mean annual water balance for Australian catchments, *Water Resources Research*, 41(6), doi: 10.1029/2004wr003697.
- Pushpalatha, R., C. Perrin, N. L. Moine, et V. Andréassian (2012), A review of efficiency criteria suitable for evaluating low-flow simulations, *Journal of Hydrology*, 420-421, 171-182, doi: <http://dx.doi.org/10.1016/j.jhydrol.2011.11.055>.

- Quintana-Seguí, P., P. Le Moigne, Y. Durand, E. Martin, F. Habets, M. Baillon, C. Canellas, L. Franchisteguy, et S. Morel (2008), Analysis of Near-Surface Atmospheric Variables: Validation of the SAFRAN Analysis over France, *J. Appl. Meteorol. Climatol.*, 47(1), 92-107, doi: 10.1175/2007JAMC1636.1.
- Randrianasolo, A. (2012), Généralisation de l'approche d'ensemble à la prévision hydrologique dans les bassins versants non jaugés, PhD thesis, 314 pp, AgroParisTech - Irstea (Antony).
- Roche, M. (1963), Point de vue sur le problème de l'abattement des pluies, *IAHS-AISH Publication*, 65, 266-278.
- Roche, M. (1967), Composition et rationalisation des réseaux hydrométéorologiques, 47-60 pp, Cahiers O.R.S.T.O.M., Série Hydrologie.
- Roderick, M. L., et G. D. Farquhar (2011), A simple framework for relating variations in runoff to variations in climatic conditions and catchment properties, *Water Resour. Res.*, 47, doi: 10.1029/2010wr009826.
- Rodriguez-Iturbe, I., et J. M. Mejia (1974), The design of rainfall networks in time and space, *Water Resources Research*, 10(4), 713-728.
- Rojas-Serna, C. (2005), Quelle connaissance hydrométrique minimale pour définir les paramètres d'un modèle pluie-débit ? , ENGREF (Paris), France.
- Rojas-Serna, C., C. Michel, C. Perrin, et V. Andreassian (2006), Ungauged catchments: how to make the most of a few streamflow measurements? , *Large Sample Basin Experiments of Hydrological Model Parameterization: Results of the Models Parameter Estimation Experiment*, 203-236.
- Ruelland, D., S. Ardoin-Bardin, G. Billen, et E. Servat (2008), Sensitivity of a lumped and semi-distributed hydrological model to several methods of rainfall interpolation on a large basin in West Africa, *Journal of Hydrology*, 361(1-2), 96-117, doi: 10.1016/j.jhydrol.2008.07.049.
- Samuel, J., P. Coulibaly, et R. A. Metcalfe (2011), Estimation of Continuous Streamflow in Ontario Ungauged Basins: Comparison of Regionalization Methods, *J. Hydrol. Eng.*, 16(5), 447-459, doi: 10.1061/(asce)he.1943-5584.0000338.
- Sankarasubramanian, A., R. M. Vogel, et J. F. Limbrunner (2001), Climate elasticity of streamflow in the United States, *Water Resources Research*, 37(6), 1771-1781, doi: 10.1029/2000wr900330.
- Sauquet, E., et C. Catalogne (2010), Interpolation des modules: quelles évolutions depuis la note technique de 1987 ? , 29 pp, Cemagref.
- Sauquet, E., L. Gottschalk, et E. Leblois (2000), Mapping average annual runoff: a hierarchical approach applying a stochastic interpolation scheme, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 45(6), 799-815, doi: 10.1080/02626660009492385.
- Schaake, J. (1981), Summary of river forecasting raingage network density requirements, *Unpublished*, available at <http://www.nws.noaa.gov/oh/mopex/raingage%20density%20requirement.htm>.
- Schaake, J., Q. Duan, M. Smith, et V. Koren (2000), Criteria to select basins for hydrologic model development and testing, in *15th Conference on Hydrology AMS*, edited, Long Beach, CA.

- Schreiber, P. (1904), Über die Beziehungen zwischen dem Niederschlag und der Wasserführung der Flüsse in Mitteleuropa, *Zeitschrift für Meteorologie*, 21, 441-452.
- Seed, A. W., et G. L. Austin (1990), Sampling errors for rain-gage-derived mean areal daily and monthly rainfall, *Journal of Hydrology*, 118(1-4), 163-173, doi: 10.1016/0022-1694(90)90256-w.
- Segond, M. L., H. S. Wheater, et C. Onof (2007), The significance of spatial rainfall representation for flood runoff estimation: A numerical evaluation based on the Lee catchment, UK, *Journal of Hydrology*, 347(1-2), 116-131, doi: 10.1016/j.jhydrot.2007.09.040.
- Seibert, J., et K. J. Beven (2009), Gauging the ungauged basin: how many discharge measurements are needed?, *Hydrol. Earth Syst. Sci.*, 13(6), 883-892.
- Seibert, J., et J. J. McDonnell (2015), Gauging the Ungauged Basin: Relative Value of Soft and Hard Data, *J. Hydrol. Eng.*, 20(1), doi: 10.1061/(asce)he.1943-5584.0000861.
- Sene, K. J., et F. A. K. Farquharson (1998), Sampling Errors for Water Resources Design: The Need for Improved Hydrometry in Developing Countries, *Water Resources Management*, 12(2), 121-138, doi: 10.1023/a:1007952319495.
- Skoien, J. O., R. Merz, et G. Blöschl (2006), Top-kriging - geostatistics on stream networks, *Hydrol. Earth Syst. Sci.*, 10(2), 277-287.
- Sorooshian, S., V. K. Gupta, et J. L. Fulton (1983), Evaluation of maximum-likelihood parameter-estimation techniques for conceptual rainfall-runoff models - Influence of calibration data variability and length on model credibility, *Water Resources Research*, 19(1), 251-259, doi: 10.1029/WR019i001p00251.
- Spence, C., P. Saso, et J. Rausch (2007), Quantifying the impact of hydrometric network reductions on regional streamflow prediction in Northern Canada, *Canadian Water Resources Journal*, 32(1), 1-20.
- St-Hilaire, A., T. Ouarda, M. Lachance, B. Bobee, J. Gaudet, et C. Gignac (2003), Assessment of the impact of meteorological network density on the estimation of basin precipitation and runoff: a case study, *Hydrol. Process.*, 17(18), 3561-3580, doi: 10.1002/hyp.1350.
- Storm, B., K. H. Jensen, et J. C. Refsgaard (1988), Estimation of catchment rainfall uncertainty and its influence on runoff prediction, *Nordic Hydrology*, 19(2), 77-88.
- Sugawara, M. (1993), On the weights of precipitation stations, *Advances in theoretical hydrology*, 59-74.
- Tabary, P., P. Dupuy, G. L'Henaff, C. Gueguen, L. Moulin, O. Laurantin, C. Merlier, et J. M. Soubeyrou (2012), A 10-year (1997–2006) reanalysis of Quantitative Precipitation Estimation over France: methodology and first results in *Weather Radar and Hydrology, IAHS Publ.*, 351, 255-260.
- Tallaksen, L. M., et H. A. J. van Lanen (2004), *Hydrological Drought: Processes and Estimation Methods for Streamflow and Groundwater*, 579 pp., Elsevier.
- Tan, S. B. K., L. H. C. Chua, E. B. Shuy, E. Y. M. Lo, et L. W. Lim (2008), Performances of rainfall-runoff models calibrated over single and continuous storm flow events, *J. Hydrol. Eng.*, 13(7), 597-607, doi: 10.1061/(asce)1084-0699(2008)13:7(597).

- Tasker, G. D. (1986), Generating efficient gaging plans for regional information, *Integrated design of hydrological networks. Proc. Budapest symposium, 1986*, 269-281.
- Thiessen, A. H. (1911), Precipitation averages for large areas, *Monthly Weather Review*, 39(7), 1082-1089, doi: 10.1175/1520-0493(1911)39<1082b:PAFLA>2.0.CO;2.
- Tsintikidis, D., K. R. Georgakakos, J. A. Sperflage, D. E. Smith, et T. M. Carpenter (2002), Precipitation uncertainty and raingauge network design within Folsom Lake watershed, *J. Hydrol. Eng.*, 7(2), 175-184, doi: 10.1061/(asce)1084-0699(2002)7:2(175).
- Turc, L. (1954), Le bilan d'eau des sols: relation entre les précipitations, l'évaporation et l'écoulement, *Annales Agronomiques, Série A*(5), 491-595.
- Valéry, A. (2010), Modélisation précipitations – débit sous influence nivale. Élaboration d'un module neige et évaluation sur 380 bassins versants., PhD thesis, 405 pp, Cemagref-AgroParisTech, Antony & Paris.
- Vidal, J. P., E. Martin, L. Franchisteguy, M. Baillon, et J. M. Soubeyroux (2010), A 50-year high-resolution atmospheric reanalysis over France with the Safran system, *Int. J. Climatol.*, 30(11), 1627-1644, doi: 10.1002/joc.2003.
- Volkman, T. H. M., S. W. Lyon, H. V. Gupta, et P. A. Troch (2010), Multicriteria design of rain gauge networks for flash flood prediction in semiarid catchments with complex terrain, *Water Resources Research*, 46, doi: W1155410.1029/2010wr009145.
- Wagner, T., N. McIntyre, M. J. Lees, H. S. Wheater, et H. V. Gupta (2003), Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis, *Hydrol. Process.*, 17(2), 455-476, doi: 10.1002/hyp.1135.
- Walker, S. (2000), The value of hydrometric information in water resources management and flood control, *Meteorological Applications*, 7(4), 387-397, doi: 10.1017/S135048270001626.
- Williams, C. A., et al. (2012), Climate and vegetation controls on the surface water balance: Synthesis of evapotranspiration measured across a global network of flux towers, *Water Resources Research*, 48(6), doi: 10.1029/2011wr011586.
- WMO (1965), Guide to Hydrometeorological practices, 115 pp, WMO - No. 168.
- WMO (2008), Guide to Hydrological Practices, WMO - No. 168.
- Xu, H., C.-Y. Xu, H. Chen, Z. Zhang, et L. Li (2013), Assessing the influence of rain gauge density and distribution on hydrological model performance in a humid region of China, *Journal of Hydrology*, 505, 1-12, doi: <http://dx.doi.org/10.1016/j.jhydrol.2013.09.004>.
- Yang, H., D. Yang, Z. Lei, et F. Sun (2008), New analytical derivation of the mean annual water-energy balance equation, *Water Resour. Res.*, 44, W03410, doi:03410.01029/02007WR006135.
- Yapo, P. O., H. V. Gupta, et S. Sorooshian (1996), Automatic calibration of conceptual rainfall-runoff models: Sensitivity to calibration data, *Journal of Hydrology*, 181(1-4), 23-48, doi: 10.1016/0022-1694(95)02918-4.

- Yokoo, Y., M. Sivapalan, et T. Oki (2008), Investigating the roles of climate seasonality and landscape characteristics on mean annual and monthly water balances, *Journal of Hydrology*, 357(3-4), 255-269, doi: 10.1016/j.jhydrol.2008.05.010.
- Zalachori, I. (2013), Prévisions hydrologiques d'ensemble : développements pour améliorer la qualité des prévisions et estimer leur utilité, PhD thesis, 397 pp, AgroParisTech - Irstea (Antony).
- Zhang, L., N. Potter, K. Hickel, Y. Q. Zhang, et Q. X. Shao (2008), Water balance modeling over variable time scales based on the Budyko framework - Model development and testing, *Journal of Hydrology*, 360(1-4), 117-131, doi: 10.1016/j.jhydrol.2008.07.021.

Annexes

Annexe A. Liste des 609 bassins versants de la partie I

Tableau 14. Liste des 609 bassins versants utilisés pour la partie I sur l'impact de la densité du réseau hydrométrique en hydrologie. S désigne la surface, Z l'altitude moyenne des bassins versants et Q, P et ETP désignent respectivement leurs moyennes annuelles de débit, de pluie et d'évapotranspiration potentielle.

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
A1050310	L'III à Altkirch	238	434	326	957	653
A1080330	L'III à Didenheim	668	390	309	932	659
A1122010	La Largue à Dannemarie	140	442	272	971	650
A1122020	La Largue à Friesen	101	470	351	991	648
A1252010	La Doller à Reiningue	182	575	728	1390	621
A1310310	L'III à Ensisheim	1030	400	319	991	657
A1350310	L'III à Colmar [Ladhof]	2105	409	296	988	657
A2023030	La Petite Fecht à Stosswihr [Village]	44	874	1013	1595	566
A2052020	La Fecht à Ostheim	457	673	460	1249	611
A2332110	La Lièpvrette à Lièpvre	107	602	549	1365	607
A2352020	Le Giessen à Sélestat [amont]	277	501	388	1153	622
A2512010	L'Andlau à Andlau	42	656	606	1209	604
A2612010	L'Ehn à Niedernai	57	443	296	1003	645
A2732010	La Bruche à Russ [Wisches]	224	629	848	1248	591
A2842010	La Mossig à Sultz-les-Bains	167	349	260	820	659
A3301010	La Moder à Schweighouse-sur-Moder [aval]	622	267	282	833	661
A3422010	La Zorn à Saverne [Schinderthal]	184	424	387	1076	636
A3472010	La Zorn à Waltenheim-sur-Zorn	684	291	270	846	661
A3792010	La Sauer à Beinheim	502	223	234	812	673
A3832010	Le Seltzbach à Niederroedern	204	187	258	807	680
A3902010	La Lauter à Wissembourg [Weiler]	275	326	288	887	646
A4250640	La Moselle à Épinal	1218	650	999	1506	590
A4362030	La Vologne à Cheniménil [2]	354	646	890	1537	592
A5431010	Le Madon à Pulligny	948	328	357	906	641
A5730610	La Moselle à Toul	3346	442	601	1132	625
A6051020	La Meurthe à Saint-Dié	371	642	676	1544	592
A6443010	Le Vacon à Barbas	36	303	412	1047	640
A6571110	La Vezouze à Lunéville	560	315	378	945	642
A6621210	La Mortagne à Sainte-Hélène [Autrey]	98	484	520	1433	627
A6731220	La Mortagne à Gerbéviller	498	361	364	1127	639
A6761010	La Meurthe à Damelevières	2294	410	468	1169	629
A7010610	La Moselle à Custines	6835	400	526	1088	632
A7642010	La Petite Seille à Château-Salins	150	250	266	850	654
A7881010	La Seille à Metz	1274	234	243	802	657
A7930610	La Moselle à Hauconcourt	9384	357	435	1013	638
A8071010	L'Orne à Boncourt	414	232	297	860	648
A8322010	Le Woigot à Briey	77	282	508	829	635
A8431010	L'Orne à Rosselange	1241	246	338	848	643

Annexe A. Liste des 609 bassins versants de la partie I

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
A8732010	La Canner à Koenigsmacker	109	246	278	801	647
A9001050	La Sarre Blanche à Laneuveville-lès-Lorquin	65	435	653	1148	629
A9013050	La Sarre Rouge à Vasperviller	90	503	511	1213	614
A9091050	La Sarre à Keskastel	881	318	337	952	647
A9091060	La Sarre à Diedendorf	717	331	276	973	644
A9221010	La Sarre à Sarreinsming	1760	292	325	907	650
A9402110	L'Horn à Bousseviller	94	332	306	899	641
A9862010	La Nied Allemande à Varize	367	283	313	835	642
B0220010	La Meuse à Goncourt	368	372	357	908	648
B1092010	Le Mouzon à Circourt-sur-Mouzon [Villars]	401	378	354	948	633
B1150010	La Meuse à Domrémy-la-Pucelle	1035	370	410	936	638
B1282010	Le Vair à Soulosse-sous-Saint-Élophe	440	367	354	942	634
B1340010	La Meuse à Vaucouleurs [Chalaines]	1728	364	376	937	636
B2220010	La Meuse à Saint-Mihiel	2543	350	373	935	636
B3150020	La Meuse à Stenay	3915	319	400	933	637
B4031010	La Chiers à Montigny-sur-Chiers	285	337	514	882	618
B4132010	La Crusnes à Pierrepont	207	339	361	860	615
B4572010	Le Loison à Han-lès-Juvigny	349	237	352	871	645
B4601010	La Chiers à Chauvency-le-Château	1706	286	424	889	632
B4631010	La Chiers à Carignan	1978	282	412	897	631
B5172010	La Bar à Cheveuges	434	209	383	943	641
B5572010	La Sormonne à Belval	369	255	550	1022	621
B6111010	La Semoy à Haulmé	1339	385	718	1041	598
B7200010	La Meuse à Chooz [Ile Graviat]	10184	300	453	963	627
B7322010	La Houille à Landrichamps	186	363	571	1122	595
D0156520	L'Helpe Majeure à Taisnières-en-Thiérache	322	209	382	886	621
D0206010	La Solre à Ferrière-la-Grande	115	190	389	877	625
E3346010	La Marque à Bouvines	132	43	211	710	644
E3511210	La Lys à Ligny	87	146	382	1014	612
E3511220	La Lys à Delettes	162	128	403	1011	613
E3646210	La Clarence à Robecq	222	84	163	776	627
E4035710	L'Aa à Wizernes	392	135	427	974	608
E4306010	La Hem à Tournehem-sur-la-Hem [Guémy]	106	122	450	996	616
E4905710	L'Yser à Bambecque [Engelshof]	234	29	239	700	631
E5300210	La Liane à Wirwignes	103	93	562	1001	620
E5400310	La Canche à Brimeux	917	110	416	947	617
E5406510	La Ternoise à Hesdin	345	119	410	942	615
E5505720	L'Authie à Dompierre-sur-Authie	792	115	312	858	620
E6397010	L'Hallue à Bavelincourt	118	111	150	734	622
G1003010	L'Yères à Touffreville-sur-Eu	255	152	346	902	623
H0100010	La Seine à Nod-sur-Seine	373	404	389	911	653
H0100020	La Seine à Plaines-Saint-Lange	686	346	514	885	661
H0321030	L'Ource à Autricourt	574	353	350	894	658
H0503010	L'Hozain à Buchères [Courgerennes]	249	170	180	760	674
H1231010	L'Aube à Blaincourt-sur-Aube	1659	303	361	899	655

Annexe A. Liste des 609 bassins versants de la partie I

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
H1333010	La Laine à Soulaines-Dhuys	22	200	1439	848	669
H2062010	Le Beuvron à Ouagne [Champmoreau]	264	254	260	875	685
H2073110	Le Sauzay à Corvol-l'Orgueilleux	87	272	352	898	677
H2322010	Le Serein à Bierre-lès-Semur	267	430	275	913	658
H2402010	L'Armançon à Brianny	221	410	234	876	658
H2412010	L'Armançon à Quincy-le-Vicomte	478	354	223	872	661
H2473010	L'Armanche à Chessy-les-Prés	477	188	249	770	670
H2513110	Le Tholon à Champvallon	133	176	196	722	699
H2721010	L'Yonne à Courlon-sur-Yonne	10687	263	270	840	678
H3102010	L'Ouanne à Toucy	152	266	200	804	675
H3122020	L'Ouanne à Gy-les-Nonains	877	209	170	770	683
H3201010	Le Loing à Châlette-sur-Loing	2302	170	166	731	688
H3613010	Le Lunain à Paley	162	148	101	717	687
H3613020	Le Lunain à Épisy	247	131	89	713	687
H3621010	Le Loing à Épisy	3917	148	154	712	689
H4002010	L'Oeuf à Bondaroy	241	125	67	642	689
H4022020	L'Essonne à Guigneville-sur-Essonne [La Mothe]	851	117	142	647	689
H4033010	La Juine à Méréville	210	129	143	621	684
H4042010	L'Essonne à Ballancourt-sur-Essonne	1857	121	141	639	683
H4223110	La Remarde à Saint-Cyr-sous-Dourdan	152	147	135	669	671
H4252010	L'Orge à Morsang-sur-Orge	934	133	131	656	678
H5062010	Le Rognon à Doulaincourt-Saucourt	618	363	472	983	642
H5172010	La Saulx à Vitry-en-Perthois	2109	264	387	991	644
H5173110	Le Bruxenelle à Brusson	131	155	254	830	674
H5302010	Le Surmelin à Saint-Eugène	443	207	187	742	643
H5412010	Le Petit Morin à Montmirail	358	175	178	735	650
H5732010	Le Grand Morin à Pommeuse	769	164	228	736	660
H5920010	La Seine à Paris [Austerlitz après création lacs]	43825	201	222	786	671
H6162010	L'Aire à Chevières	957	251	451	991	640
H6201010	L'Aisne à Mouron	2242	208	339	916	649
H6221010	L'Aisne à Givry	2888	195	341	902	650
H6313020	La Suipe à Orainville	810	128	167	731	658
H6412010	La Vesle à Saint-Brice-Courcelles	735	135	136	715	658
H6412020	La Vesle à Châlons-sur-Vesle	828	131	160	710	658
H6423010	L'Ardres à Fismes	300	174	162	687	647
H7021010	L'Oise à Hirson	320	268	516	995	614
H7033010	Le Thon à Origny-en-Thiérache	256	216	457	945	625
H7401010	L'Oise à Sempigny	4320	137	251	804	640
H7423710	L'Aronde à Clairoux	280	93	141	675	648
H7602010	La Brèche à Nogent-sur-Oise	459	118	158	681	641
H7611012	L'Oise à Pont-Sainte-Maxence [Sarron]	13484	139	256	788	649
H7702010	Le Thérain à Bonnières	200	180	258	859	619
H7742020	Le Thérain à Maysel	1210	136	212	766	632
H7813010	La Nonette à Courteuil [Saint-Nicolas d'Acy]	335	102	151	707	665
H7853010	Le Sausseron à Nesles-la-Vallée	102	107	169	691	659

Annexe A. Liste des 609 bassins versants de la partie I

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
H7900010	La Seine à Poissy	61642	181	249	781	665
H8012010	L'Epte à Gournay-en-Bray	247	148	240	906	624
H8042010	L'Epte à Fourges	1386	130	208	788	641
H8043310	L'Aubette de Magny à Ambleville	99	123	146	711	650
H8212010	L'Andelle à Vascoeuil	377	159	331	907	624
H9021010	L'Eure à Saint-Luperce	315	224	159	731	652
H9202010	L'Avre à Acon	477	210	165	730	650
H9402030	L'Iton à Normanville	1029	173	114	676	650
H9501010	L'Eure à Louviers	5891	162	137	658	659
H9913020	Le Cailly à Notre-Dame-de-Bondeville	199	147	417	910	623
I0011010	La Risle à Rai	147	272	300	811	630
I0113010	Le Guiel à Montreuil-l'Argillé	82	251	260	817	631
I0211019	La Risle [pour Seine rive gauche] à Pont-Authou	1813	188	272	759	642
I1203010	La Calonne aux Authieux-sur-Calonne	173	141	323	800	645
I2021010	La Dives à Beaumais	283	142	176	735	649
I3131010	L'Orne à Rabodanges	1019	213	247	817	645
I4053010	La Mue à Reviars	100	61	108	750	656
I6943110	La Divette à Octeville	103	92	470	953	639
I7913610	Le Thar à Jullouville	73	93	438	1026	660
I9221010	La Sélune à Ducey [Vézins]	744	159	451	1010	654
J0121510	Le Couesnon à Romazy	518	110	297	848	674
J0144010	La Loysance à Saint-Ouen-la-Rouërie	82	128	307	858	668
J0201510	Le Couesnon à Sougéal [Le Poirier]	863	101	281	832	673
J0323010	Le Guyoult à Epiniac	62	73	238	757	667
J0621610	La Rance à Guenroc [Rophemel]	372	115	230	782	665
J1103010	L'Arguenon à Jugon-les-Lacs	103	123	251	840	660
J1114010	La Rosette à Mégrit	113	97	211	762	667
J1513010	Le Gouët à Saint-Julien	135	213	396	1008	639
J1813010	Le Leff à Quemper-Guézenec	342	115	257	892	657
J2233010	Le Léguer à Belle-Isle-en-Terre	265	215	576	1060	632
J2514010	Le Dourduff à Garlan	44	104	386	1095	654
J2603010	Le Jarlot à Plougouven	44	161	523	1154	648
J2605410	Le Tromorgant à Plougouven	42	137	465	1126	643
J2723010	La Penze à Taulé [Penhoat]	142	149	631	1156	649
J3014310	L'Horn à Mespaul [Pont Milin]	49	85	450	954	657
J3024010	Le Guillec à Trézilidé	45	86	476	941	653
J3205710	L'Aber Wrac'h au Drennec	24	77	577	1088	660
J3213020	L'Aber-Benoît à Plabennec [Loc Maria]	27	80	598	1092	661
J3323020	L'Aber Ildut à Brélès [Keringar]	95	70	491	1045	661
J3413020	L'Élorn à Ploudiry [Kerfaven]	201	138	681	1156	648
J3514010	La Mignonne à Irvillac [Pont Mell]	67	115	685	1147	650
J3811810	L'Aulne à Châteauneuf-du-Faou [Pont Pol ty Glass]	1225	176	570	1180	643
J3834010	La Douffine à Saint-Ségal [Kerbriant]	140	163	726	1254	650
J4014010	Le Goyen à Pont-Croix [Kermaria]	89	79	493	966	683
J4211910	L'Odet à Ergué-Gabéric [Tréodet]	203	145	757	1242	658

Annexe A. Liste des 609 bassins versants de la partie I

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
J4224010	Le Jet à Ergué-Gabéric	108	119	654	1170	669
J4231911	L'Odet à Quimper [Kervir virtuelle]	328	132	719	1213	661
J4313010	Le Steir à Guengat [Ty Planche]	181	117	639	1185	666
J4514010	Le Moros à Concarneau [Pont D 22]	21	91	494	1077	685
J4614010	Le Ster Goz à Bannalec [Pont Meya]	72	127	664	1210	668
J4734010	L'Inam au Faouët [Pont Priant]	118	162	626	1141	652
J4742010	L'Éllé à Arzano [Pont Ty Nadan]	576	165	517	1134	652
J4902012	La Laïta à Quimperlé [virtuelle 2]	852	152	499	1152	655
J5102210	Le Scorff à Plouay [Pont Kerlo]	299	145	523	1130	659
J5613010	L'Evel à Guénin	316	106	336	917	672
J5704810	Le Coët-Organ à Quistinic [Kerdec]	46	132	562	1102	671
J6213010	Le Loch à Brech	182	84	441	986	682
J7024010	La Valière à Erbrée [Pont D 110]	30	136	275	880	672
J7083110	Le Chevré à la Bouëxière [Le Drugeon]	152	97	246	826	679
J7214010	La Flume à Pacé [Pont D 231]	91	96	220	751	673
J7353010	Le Meu à Montfort-sur-Meu [L'Abbaye]	477	96	206	766	674
J7483010	La Seiche à Bruz [Carcé]	809	70	183	737	693
J7633010	Le Semnon à Bain-de-Bretagne [Rochereuil]	406	81	221	765	694
J7824010	L'Aron à Grand-Fougeray [La Bernardais]	112	63	198	769	698
J7973010	Le Canut Sud à Saint-Just [La rivière Colombel]	40	61	230	764	700
J8002310	L'Oust à Saint-Martin-des-Prés [La Ville Rouault]	29	218	443	1054	634
J8433010	La Claie à Saint-Jean-Brévelay	135	112	392	962	677
J8502310	L'Oust à Saint-Gravé [Echuse Le Guélin]	2454	121	304	866	668
J8632410	L'Aff à Quelneuc [La rivière]	343	102	249	795	676
J8813010	L'Arz à Molac [Le Qinquizio]	161	91	448	972	682
J9999999	Le Couesnon à Antrain	805	102	278	827	674
K0010010	La Loire à Usclades-et-Rieutord [Rieutord]	60	1280	1350	1378	539
K0030010	La Loire à Issarlès [Pont de la Borie]	229	1235	1096	1375	539
K0100020	La Loire à Goudet	442	1178	361	1223	546
K0523010	L'Ance du Nord à Saint-Julien-d'Ance [Laprat]	347	1003	400	1089	574
K0550010	La Loire à Bas-en-Basset	3260	969	352	953	592
K0567530	La Semène à Jonzieux	58	1000	472	1003	585
K0568310	L'Ecotay à Marlhès	5	996	606	1039	568
K0614010	Le Furan à Andrézieux-Bouthéon	175	662	440	871	659
K0663310	La Coise à Larajasse [Le Nézel]	61	753	333	882	650
K0704510	La Toranche à Saint-Cyr-les-Vignes	55	578	269	847	681
K0724510	Le Chanasson à Civens [La rivière]	13	486	218	773	693
K0733220	Le Lignon de Chalmazel à Chalmazel [Chevelières]	60	1180	898	1125	545
K0744010	L'Anzon à Débats-Rivière-d'Orpra [Cotes]	180	763	454	986	632
K0763310	Le Vizezy à Essertines-en-Châtelneuf [La Guillanche]	41	1040	493	1024	591
K0773220	Le Lignon de Chalmazel à Poncins [2]	662	743	390	925	638
K0813020	L'Aix à Saint-Germain-Laval	197	750	485	990	624
K0943010	Le Rhins à Amplepuis	114	643	492	1028	665
K0974010	Le Gand à Neaux	86	570	327	907	673
K1173210	L'Arconce à Montceaux-l'Étoile	593	358	298	898	687

Annexe A. Liste des 609 bassins versants de la partie I

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
K1273110	Le Ternin à Tavernay [Pré Charmoy]	259	488	449	1147	655
K1284810	La Selle à la Celle-en-Morvan [Polroy]	135	558	702	1168	642
K1321810	L'Arroux à Étang-sur-Arroux [Pont du Tacot]	1792	431	397	979	665
K1503010	La Besbre à Châtel-Montagne	157	738	534	1137	623
K1524010	Le Barbenan au Breuil	121	585	486	1022	664
K1724210	La Dragne à Vandenesse	114	393	516	1084	666
K1753110	L'Alène à Cercy-la-Tour [Coueron]	333	329	413	962	673
K1930010	La Loire à Nevers	17630	516	323	895	661
K1954010	La Nièvre d'Arzembouy à Poiseux [Poisson]	226	305	334	929	665
K2064010	Le Langouyrou à Langogne	66	1207	729	942	543
K2123010	Le Chapeauroux à Châteauneuf-de-Randon [Hermet]	125	1281	406	963	528
K2134010	La Clamouze à Chastanier	50	1216	464	933	551
K2163110	Le Grandrieu à Grandrieu	71	1284	506	964	519
K2233020	L'Ance du Sud à Monistrol-d'Allier [Pouzas]	231	1161	402	898	553
K2283110	La Desges à Chanteuges	141	1083	418	876	567
K2365510	Le Lidenne à Couteuges [La Tuilerie Basse]	48	697	164	747	639
K2514010	L'Allanche à Joursac [Pont du Vernet]	156	1115	595	1268	548
K2523010	L'Alagnon à Joursac [Le Vialard]	322	1123	643	1232	545
K2544010	L'Allagnonette à Massiac	68	883	268	855	608
K2593010	L'Alagnon à Lempdes	994	971	385	1047	584
K2674010	La Couze Chambon à Montaigut-le-Blanc [Champeix]	159	960	521	1126	591
K2783010	La Morge à Maringues [Côte Rouge]	668	512	199	826	686
K2884010	La Faye à Olliegues [Giroux-Faye]	73	866	719	1064	599
K3153010	L'Andelot à Loriges	203	354	150	772	688
K3264010	La Saunade à Pontaugur	111	686	315	951	632
K3292020	La Sioule à Saint-Priest-des-Champs [Fades-Besserve]	1300	781	433	1005	616
K3374710	Le Boublon Lagées à Fourilles	71	339	133	773	694
K3650810	L'Allier à Cuffy [Pont du Guétin]	14349	660	328	895	641
K4094010	Le Nohain à Saint-Martin-sur-Nohain [Villiers]	478	236	222	829	679
K4383110	Le Dhuy à Saint-Cyr-en-Val [Gobson]	219	114	98	682	700
K4443010	L'Ardoux à Lailly-en-Val	165	107	101	719	702
K5200910	Le Cher à Teillet-Argenty	1611	520	312	924	661
K5653010	L'Auron à Bourges [L'Ormediot]	575	192	205	775	706
K6373020	La Petite Sauldre à Ménétréol-sur-Sauldre	332	255	307	832	677
K7312610	L'Indre à Saint-Cyran-du-Jambot	1707	201	247	777	706
K7514010	L'Échandon à Saint-Branchs	128	106	151	691	716
L0010610	La Vienne à Peyrelevade [Servières]	64	841	868	1407	596
L0010620	La Vienne à Peyrelevade [La Rigole du Diable]	77	834	816	1400	596
L0314010	La Vige à Saint-Martin-Sainte-Catherine	131	482	634	1121	677
L0563010	La Briance à Condat-sur-Vienne [Chambon Veyrinas]	605	386	428	1102	701
L0813010	La Glane à Saint-Junien [Le Dérot]	298	315	416	1036	709
L0920610	La Vienne à Étagnac [Pont de Pilas]	4109	446	543	1121	684
L2321610	Le Clain à Vivonne [Danlot]	1824	146	247	819	723
L3200610	La Vienne à Ingrandes	10087	270	396	929	707
L4033010	La Rozeille à Moutier-Rozeille [Aubusson]	190	681	437	1123	633

Annexe A. Liste des 609 bassins versants de la partie I

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
L4110710	La Creuse à Ahun	748	621	411	1119	644
L4220710	La Creuse à Fresselines	1235	537	411	1060	660
L4321710	La Petite Creuse à Genouillac	561	418	307	893	675
L4411710	La Petite Creuse à Fresselines [Puy Rageaud]	853	393	318	900	680
L4530710	La Creuse à Éguzon-Chantôme	2427	459	371	986	672
L4653010	La Bouzanne à Velles [Forges]	438	214	229	811	708
L4710710	La Creuse à Ciron [Scoury]	3339	389	343	937	683
L5101810	La Gartempe à Folles [Bessines]	568	446	457	1022	681
L5223020	Le Vincou à Bellac [2]	286	325	392	1029	703
L5301811	La Gartempe à Folles [virtuelle]	368	450	545	1003	678
L5323010	La Brame à Oradour-Saint-Genest	232	267	310	935	715
L6020710	La Creuse à Leugny	8013	299	304	897	699
L6202030	La Claise au Grand-Pressigny [Étableau 2]	886	117	160	742	726
L7000610	La Vienne à Nouâtre	19895	267	313	896	706
M0050620	La Sarthe à Saint-Céneri-le-Gérei [Moulin du Désert]	909	185	246	791	660
M0243010	L'Orne Saosnoise à Montbizot [Moulin Neuf Cidrerie]	502	103	164	744	691
M0250610	La Sarthe à Neuville-sur-Sarthe [Montreuil]	2713	160	240	784	672
M0361510	L'Huisne à Nogent-le-Rotrou [Pont de bois]	833	200	236	768	658
M0680610	La Sarthe à Saint-Denis-d'Anjou [Beffes]	7523	135	198	756	682
M1034020	L'Ozanne à Trizay-lès-Bonneval [Prémoteux]	267	187	174	707	663
M1041610	Le Loir à Saint-Maur-sur-le-Loir	1080	178	97	665	666
M1151610	Le Loir à Villavard	4529	147	103	650	679
M1213010	La Braye à Valennes [La Caboche]	269	176	202	727	677
M1354020	L'Escotais à Saint-Paterne-Racan	68	109	145	702	705
M1531610	Le Loir à Durtal	7918	130	127	674	687
M3230920	La Mayenne à Saint-Fraimbault-de-Prières	1857	198	338	903	649
M3253110	L'Aron à Moulay	185	182	330	837	664
M3313010	L'Ernée à Ernée	121	184	357	902	659
M3323010	L'Ernée à Andouillé [Les Vaugeois]	376	157	331	892	666
M3340910	La Mayenne à l'Huisserie [Bonne]	2908	178	322	881	658
M3423010	La Jouanne à Forcé	404	119	235	800	685
M3504010	Le Vicoïn à Nuillé-sur-Vicoïn	234	126	255	833	679
M3600910	La Mayenne à Château-Gontier	3935	160	291	856	665
M3630910	La Mayenne à Chambellay	4166	154	298	848	668
M3771810	L'Oudon à Châtellais [Marcillé]	732	81	174	736	696
M3774010	Le Chéran à la Boissière	77	88	210	723	694
M4101921	La Maine à Angers [2]	22199	128	184	737	684
M5300010	La Loire à Montjean-sur-Loire	110188	296	253	811	685
M6014010	Le Beuvron à Andrezé [Tuvache]	38	106	252	748	716
M6240010	La Loire [totale] à Nantes [Saint Félix]	111975	292	246	810	686
M6323010	L'Erdre à Candé [La Grée]	171	61	181	706	709
M6333020	L'Erdre à Nort-sur-Erdre [Moulin de Vault]	463	54	182	742	710
M6353020	L'Erdre [totale] à Nantes [Saint Félix]	976	44	180	760	713
M7112410	La Sèvre Nantaise à Tiffauges [La Moulinette]	872	170	333	907	710
M7453010	La Maine à Remouillé	595	80	293	857	728

Annexe A. Liste des 609 bassins versants de la partie I

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
N3024010	Le Louing à Chantonnay [St-Philbert du Pont Charrault]	121	105	298	886	728
N3222010	La Smagne à Sainte-Pexine [Les Mottes]	184	68	264	844	742
N3301610	Le Lay à Mareuil-sur-Lay-Dissais	1045	94	278	871	729
N4010610	La Sèvre Niortaise à Azay-le-Brûlé [Pont de Ricou]	244	130	526	901	728
O0015310	Le Maudan à Fos	36	1454	1205	1473	554
O0105110	La Neste de Cap de Long à Aragnouet [Les Edelweiss]	5	2706	1508	1614	369
O0362510	Le Salat à Soueix-Rogalle [Kercabanac]	385	1356	1333	1594	596
O0485110	Le Lachein à Balaguères [Baget - Las Hountas]	10	837	1530	1173	709
O0502520	Le Salat à Saint-Lizier [Saint Girons]	1159	1166	961	1473	636
O0525010	La Gouarèze à Cazavet [Aliou]	14	695	999	1107	723
O0554010	L'Arbas à Castelbiague [Pont de Prades]	92	676	604	1053	726
O0592510	Le Salat à Roquefort-sur-Garonne	1579	986	850	1349	666
O0624010	Le Volp à Montberaud [Sainte-Croix-Volvestre]	90	483	354	1023	748
O0744030	L'Arize au Mas-d'Azil	220	600	568	1122	736
O0794010	L'Arize à Rieux [Volvestre]	479	458	356	956	756
O0964030	La Louge au Fousseret	277	360	228	866	763
O0984010	La Louge à Muret	481	307	401	806	775
O1076010	Le Quioulès à Château-Verdun [Riete 2]	54	2131	1149	1561	431
O1115010	L'Artigue à Uzatz [Cibelle]	24	2187	1687	1519	481
O1484320	La Touyre à Montferrier [2]	29	1467	1422	1251	624
O1494330	La Touyre à Lérans	95	939	649	1085	691
O1874010	La Lèze à Labarthe-sur-Lèze	356	305	178	779	788
O1900010	La Garonne à Portet-sur-Garonne	10001	901	565	1132	677
O1900015	La Garonne à Portet-sur-Garonne	10001	901	563	1132	677
O1984310	Le Touch à Toulouse [Saint-Martin-du-Touch]	500	248	246	725	787
O2034010	L'Aussonnelle à Seilh	188	204	133	702	801
O2215010	La Saune à Quint-Fonsegrives	109	210	124	725	803
O2552910	La Save à Larra	1118	257	172	775	781
O2620010	La Garonne à Verdun-sur-Garonne	13806	714	437	1024	710
O3011010	Le Tarn au Pont-de-Montvert [Fontchalettes]	65	1302	1633	1784	542
O3015520	Le Rieumalet au Pont-de-Montvert	22	1341	1310	1806	517
O3064010	Le Tarnon à Florac	132	982	893	1026	623
O3084320	La Mimente à Florac	126	937	890	1599	630
O3121010	Le Tarn à Montbrun [Pont de Montbrun]	588	1020	950	1402	605
O3141010	Le Tarn à Mostuéjols [La Muse]	945	973	1004	1270	617
O3165010	La Brèze à Meyrueis	34	1083	899	1077	582
O3194010	La Jonte à Meyrueis [aval]	98	1057	679	1083	588
O3364010	La Dourbie à Nant [Pont de Gardies]	428	871	546	1206	651
O3454310	La Muze à Montjaux [Saint-Hippolyte]	112	780	399	1038	666
O3841010	Le Tarn à Marsal [Pécotte]	4493	734	603	1093	680
O4102510	L'Agout à Fraisse-sur-Agout	45	1010	1120	1604	633
O4931010	Le Tarn à Villemur-sur-Tarn	9119	570	497	1041	719
O5042510	L'Aveyron à Palmas [Pont de Manson]	300	784	351	1060	664
O5055010	Le Serre à Coussergues [Resuenhe]	108	765	304	1107	679
O5092520	L'Aveyron à Onet-le-Château [Rodez]	584	749	348	1054	676

Annexe A. Liste des 609 bassins versants de la partie I

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
O5192520	L'Aveyron à Villefranche-de-Rouergue [Recoules]	1060	684	383	1025	691
O5224010	L'Alzou à Villefranche-de-Rouergue [barrage Cabal]	208	482	396	969	730
O5292510	L'Aveyron à Laguépie [1]	1604	597	372	1000	707
O5344010	Le Vioulou à Salles-Curan [Trébons-Bas]	57	938	595	1084	636
O5424010	Le Céor à Centrès [Estrebaldie]	161	676	425	1016	688
O5685010	La Bonnette à Saint-Antonin-Noble-Val	181	331	221	937	758
O5754020	La Vère à Bruniquel [La Gauterie]	310	267	219	853	783
O5964020	Le Lemboulas à Lafrançaise [Lunel]	403	186	176	815	788
O6140010	La Garonne à Lamagistère	32427	563	387	981	727
O7015810	L'Esclancide à Pelouse [Les Salces]	33	1278	468	964	524
O7035010	Le Bramont à Saint-Bauzile [Les Fonts]	117	1083	481	1310	583
O7041510	Le Lot à Balsièges [Bramonas]	468	1113	537	1138	566
O7085010	Le Coulagnet à Marvejols	83	1010	341	1023	582
O7094010	La Colagne au Monastier-Pin-Moriès [Le Monastier]	458	1075	349	1008	575
O7101510	Le Lot à Banassac [La Mothe]	1158	1046	430	1057	583
O7202510	La Truyère à Serverette	71	1221	990	957	529
O7234010	La Rimeize à Rimeize	117	1138	498	1067	562
O7234030	La Rimeize à Fau-de-Peyre [Vareilles]	81	1181	570	1078	559
O7245010	Le Chapouillet à Rimeize [Chassignoles]	65	1080	396	1030	572
O7265010	La Limagnole à Fontans [Saint-Alban]	78	1164	411	935	548
O7272510	La Truyère au Malzieu-Ville [Le Soulier]	583	1123	432	995	561
O7354010	La Lander à Saint-Georges	309	992	368	842	578
O7404010	Le Bes à Marchastel [Gour du Gou - 2]	35	1281	791	1111	529
O7444010	Le Bès à Saint-Juéry	286	1204	868	1078	543
O7502510	La Truyère à Neuvéglise [Grandval]	1795	1070	454	955	568
O7635010	La Bromme à Brommat [EDF]	109	984	1037	1625	588
O7874010	Le Dourdou à Conques	545	507	431	1084	723
O7944020	Le Rieu-Mort à Viviez [2]	153	375	384	1093	756
O8133520	Le Célé à Orniac [Les Amis du Célé]	1246	417	467	1122	731
O8255010	Le Vert à Labastide-du-Vert [Les Campagnes]	119	269	372	937	759
O8394310	La Lémance à Cuzorn	220	228	183	905	764
O9034010	Le Tolzac à Varès	258	103	164	807	789
O9134010	L'Avance à Montpouillan	410	111	153	815	783
P0010010	La Dordogne à Saint-Sauves-d'Auvergne	89	1194	1309	1310	528
P0084010	Le Chavanon à Messeix [La Cellete]	363	799	566	1121	611
P0115010	La Burande à la Tour-d'Auvergne	21	1221	1519	1482	515
P0115020	La Burande [ou ru de Burons] à Singles	85	1038	1127	1413	564
P0364010	La Santoire à Condat [Roche-Pointue]	169	1150	854	1353	543
P0714010	La Diège à Chaveroche	236	778	652	1263	613
P0894010	La Sumène à Bassignac [Pont de Vendes]	392	825	611	1291	614
P1154010	La Luzège à Lamazière-Basse [Pont de Bouyges]	250	686	825	1371	637
P1502510	La Maronne à Pleaux [Enchanet]	455	775	997	1436	630
P1712910	La Cère à Vic-sur-Cère [Comblat-le-Pont]	88	1114	1497	1831	534
P1772910	La Cère à Sansac-de-Marmiesse	349	913	1046	1670	591
P2114010	La Sourdoire à la Chapelle-aux-Saints	63	288	485	1257	724

Annexe A. Liste des 609 bassins versants de la partie I

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
P2484010	Le Céou à Saint-Cybranet	573	272	184	934	762
P3001010	La Vézère à Saint-Merd-les-Oussines [Maisonnial]	42	867	1075	1413	595
P3021010	La Vézère à Bugeat	138	835	1010	1400	609
P3131020	La Vézère à Uzerche	599	606	795	1361	656
P3201020	La Vézère à Saint-Viance	993	520	604	1290	677
P3234010	La Loyre à Voutezac [Pont de l'Aumonerie]	104	371	530	1167	714
P3274010	La Loyre à Saint-Viance [Pont de Burg]	254	292	409	1112	729
P3502510	La Corrèze à Tulle [Pont des soldats]	354	601	928	1405	655
P3922510	La Corrèze à Brive-la-Gaillarde [Le Prieur]	954	473	696	1315	692
P4015010	La Couze à Chasteaux [Le Soulier]	58	276	479	1081	755
P4114010	Le Cern au Lardin-Saint-Lazare [Rispe]	92	229	337	940	763
P4271010	La Vézère à Campagne	3657	371	508	1166	718
P5404010	L'Eyraud à la Force [Bitarel]	74	100	214	889	788
P5715010	L'Engranne à Baigneaux	30	72	181	826	785
P6081510	L'Isle à Cognac-sur-l'Isle	448	331	438	1117	718
P6134010	La Loue à Saint-Médard-d'Excideuil	197	334	476	1125	721
P7001510	L'Isle à Bassilac [Charrieras]	1859	289	442	1071	731
P7041510	L'Isle à Périgueux	2111	279	408	1053	735
P7261510	L'Isle à Abzac	3757	214	316	982	753
P8012510	La Dronne à Saint-Pardoux-la-Rivière [Le Manet]	140	354	508	1106	712
P8215010	La Belle à Mareuil	40	178	248	946	751
P8312520	La Dronne à Bonnes	1912	176	327	953	753
P9001510	L'Isle à Coutras [virtuelle]	6564	185	300	957	756
Q0100010	L'Adour [Adour seul] à Asté [Pont d'Asté]	270	1572	914	1318	564
Q0100025	L'Adour [Adour+canal] à Asté [Pont d'Asté]	270	1572	1019	1318	564
Q0115710	L'Oussouet à Trébons	32	938	983	1379	688
Q0214010	L'Échez à Louey	78	503	580	1265	744
Q0224020	L'Échez à Bordères-sur-l'Échez	154	440	611	1219	747
Q0280030	L'Adour à Estirac	876	772	543	1172	692
Q0435010	La Laysa à Villefranque	38	262	536	958	780
Q0450010	L'Adour à Cahuzac-sur-Adour	1279	610	539	1110	719
Q0522520	L'Arros à Gourgue	183	655	624	1290	712
Q0612510	L'Arros à Juillac	568	431	389	1087	750
Q1100010	L'Adour à Aire-sur-l'Adour [2]	2921	426	412	1054	750
Q2593310	La Midouze à Campagne	2478	114	266	920	789
Q3120010	L'Adour à Saint-Vincent-de-Paul	7707	228	368	1014	777
Q3224010	Le Luy de France à Monget	199	232	456	1104	784
Q3354010	Le Luy de Béarn à Saint-Médard	257	188	502	1092	793
Q3464010	Le Luy à Saint-Pandelon	1144	133	492	1139	803
Q4401010	Le Gave de Pau à Sassis [Pescadères]	416	1983	1479	1560	443
Q4674010	Le Gave d'Azun à Arras-en-Lavedan [Arras-Nouaux]	190	1739	1280	1430	522
Q4801010	Le Gave de Pau à Saint-Pé-de-Bigorre [Pont de Rieulhes]	1123	1643	1269	1439	525
Q5501010	Le Gave de Pau à Bérenx [Pont de Bérenx]	2592	916	1040	1358	664
Q6142910	Le Gave d'Ossau à Oloron-Sainte-Marie [Oloron-Sainte-Croix]	493	1392	1267	1903	579
Q6332510	Le Gave d'Aspe à Bedous [Pont d'Escot]	371	1409	2020	2024	574

Annexe A. Liste des 609 bassins versants de la partie I

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
Q7002910	Le Gave d'Oloron à Oloron-Sainte-Marie [Oloron-SNCF]	1101	1266	1519	1897	604
Q7322510	Le Saison à Mauléon-Licharre	498	729	1534	1837	700
Q7412910	Le Gave d'Oloron à Escos	2484	807	1291	1716	687
Q8032510	La Bidouze à Aïcirits-Camou-Suhast [Saint-Palais]	246	367	673	1576	775
Q8345910	Le Mendialçu à Hasparren	17	132	922	1423	811
Q9164610	La Nive des Aldudes à Saint-Étienne-de-Baïgorry	157	672	1402	1817	716
S2224610	Le Grand Arriou à Moustey [Biganon]	113	69	276	1048	786
S2242510	L'Eyre à Salles	1678	78	323	1035	785
S4214010	Le Magescq à Magescq	77	64	436	1186	800
S5144010	La Nivelle à Saint-Pée-sur-Nivelle	142	277	1109	1518	784
U0104010	Le Coney à Xertigny	64	418	643	1148	630
U0230010	La Saône à Cendrecourt	1161	344	468	1004	642
U0415010	Le Breuchin à la Proiselière-et-Langle	123	561	1137	1572	612
U0474010	La Lanterne à Fleurey-lès-Faverney	1028	395	676	1291	640
U0610010	La Saône à Ray-sur-Saône	3761	334	498	1065	652
U0635010	La Gourgeonne à Tincey-et-Pontrebeau	146	265	370	942	665
U0724010	Le Salon à Denèvre	385	302	380	915	664
U0924010	La Vingeanne à Saint-Maurice-sur-Vingeanne	397	347	314	917	656
U0924020	La Vingeanne à Oisilly	609	314	310	892	668
U1004010	L'Ognon à Servance [Fourguenons]	71	696	1394	1828	588
U1025010	Le Rahin à Plancher-Bas	32	796	1680	2020	579
U1035410	Le Scey à Beveuge	183	365	446	1203	648
U1054010	L'Ognon à Beaumotte-Aubertans	1259	380	591	1256	652
U1084010	L'Ognon à Pesmes	2071	338	516	1177	668
U1109010	La Venelle à Selongey	56	419	339	920	652
U1204010	La Tille à Crécey-sur-Tille	230	415	374	919	652
U1224010	La Tille à Arceau [Arcelot]	845	397	267	899	659
U1224020	La Tille à Cessey-sur-Tille	882	390	245	895	661
U1235020	La Norges à Genlis	271	273	316	804	694
U1334010	L'Ouche à Trouhans	906	427	314	876	665
U1334020	L'Ouche à Crimolois	874	435	306	879	663
U1420010	La Saône à Pagny-la-Ville [Lechatelet]	11693	318	426	992	669
U2002010	Le Doubs à Mouthe	33	1229	1694	1714	533
U2012010	Le Doubs à Labergement-Sainte-Marie	170	1107	825	1687	535
U2022010	Le Doubs à la Cluse-et-Mijoux [Pontarlier amont]	382	1054	648	1664	544
U2035020	Le Drugeon à Vuillecin [2]	189	890	526	1583	573
U2102010	Le Doubs à Ville-du-Pont	641	989	528	1633	553
U2122010	Le Doubs à Goumois	1159	990	797	1588	550
U2142010	Le Doubs à Glère [Courclavon]	1306	960	817	1560	556
U2215020	Le Dessoubre à Saint-Hippolyte	590	795	733	1447	587
U2222010	Le Doubs à Mathay	2236	869	761	1489	573
U2305210	Le Saint-Nicolas à Rougemont-le-Château	9	713	1079	1279	641
U2345020	La Savoureuse à Giromagny	30	800	1598	2112	583
U2345030	La Savoureuse à Belfort	144	563	958	1817	620
U2345410	Le Rhome à Lachapelle-sous-Chaux	19	571	1210	2066	618

Annexe A. Liste des 609 bassins versants de la partie I

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
U2356610	Le Rupt à Dung	43	387	429	1104	661
U2402010	Le Doubs à Voujeaucourt	3556	714	715	1384	600
U2425260	Le Cusancin à Baume-les-Dames	541	549	449	1334	644
U2512010	Le Doubs à Besançon	4658	657	668	1347	612
U2542010	Le Doubs à Rochefort-sur-Nenon	5169	622	658	1326	620
U2604030	La Loue à Vuillafans	291	806	2296	1530	589
U2615820	Le Lison [source] à Nans-sous-Sainte-Anne	210	735	800	1539	602
U2616410	Le Verneau à Nans-sous-Sainte-Anne	15	749	1064	1509	602
U2634010	La Loue à Champagne-sur-Loue	1264	620	1307	1430	628
U2654010	La Loue [totale] à Parcey	1644	551	1134	1370	643
U2722010	Le Doubs à Neublans-Abergement	7346	580	753	1316	632
U3214010	La Grosne à Jalogny [Cluny]	334	470	380	947	676
U3225010	La Guye à Sigy-le-Châtel [Corcelles]	271	330	240	886	689
U3415020	La Brenne à Sellières	37	343	545	1232	704
U3424010	La Seille à Saint-Usuge	938	306	465	1181	703
U4014010	La Reyssouze à Montagnat	84	278	207	1015	711
U4204010	La Veyle à Lent	41	290	302	1003	715
U4216010	Le Vieux Jonc à Buellas [Corgenon]	121	272	247	999	717
U4235010	Le Renon à Neuville-les-Dames	93	272	272	1000	718
U4300010	La Saône à Mâcon [amont]	26941	373	470	1070	672
U4505010	L'Ardières à Beaujeu	55	597	475	1040	665
U4644010	L'Azergues à Lozanne	798	517	301	922	686
V0205420	Le Borne à Saint-Jean-de-Sixt	65	1560	1505	1995	466
V0222010	L'Arve à Arthaz-Pont-Notre-Dame	1645	1493	1373	1749	480
V0245610	L'Aire à Saint-Julien-en-Genevois [Thaïry]	47	691	448	1231	624
V0345210	Le Redon à Margencel	32	706	462	1191	648
V0355010	Le Foron à Sciez	51	656	585	1269	644
V1015010	La Valserine à Lélex [Niaizet]	76	1248	917	1943	529
V1015030	La Valserine à Chézery-Forens [Chézery]	110	1200	1304	1907	536
V1214010	Le Fier à Dingy-Saint-Clair	224	1244	1417	1803	536
V1225010	La Filière à Argonay	163	1035	864	1626	576
V1235210	L'Ire à Doussard	25	1273	1285	1715	576
V1235610	L'Eau Morte à Doussard	93	1093	949	1646	597
V1237410	Le Laudon à Saint-Jorioz	30	998	764	1434	605
V1257810	Les Eparis à Alby-sur-Chéran	24	762	523	1413	606
V1264010	Le Fier à Vallières	1286	951	1025	1568	595
V1315020	La Laysse à la Motte-Servolex [Pont du Tremblay]	281	742	700	1515	636
V1464310	Le Furans à Arbignieu [Pont de Peyzieu]	159	525	748	1451	661
V1504010	Le Guiers Mort à Saint-Laurent-du-Pont	94	1248	1534	1867	554
V1774010	La Bourbre à Tignieu-Jameyzieu	696	375	343	1042	719
V2035010	La Lemme à Fort-du-Plasne [Pont-de-Lemme]	95	938	300	1897	568
V2114010	L'Angillon à Champagnole	107	668	984	1629	605
V2202010	L'Ain à Marigny [Chalain]	734	825	1181	1724	584
V2206010	Le Hérisson à Doucier	51	727	1035	1765	610
V2409010	L'Orbe aux Rousses	20	1216	848	2053	521

Annexe A. Liste des 609 bassins versants de la partie I

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
V2414010	La Bienne à Saint-Claude [Chenavier]	203	1005	1476	1964	545
V2444020	La Bienne à Jeurre	593	934	1582	1895	568
V2924010	L'Albarine à Saint-Rambert-en-Bugey	210	826	937	1689	599
V3015010	L'Yzeron à Craponne	46	504	226	799	713
V3114010	Le Gier à Rive-de-Gier	331	600	246	857	680
V3124010	Le Gier à Givors	407	555	246	841	690
V3315010	La Valencize à Chavanay	37	646	310	892	676
V4034020	La Véore à Beaumont-lès-Valence [Laye]	195	379	190	1013	725
V4145210	La Glueyre à Gluiras [Tisoneche]	72	852	945	1353	641
V4214010	La Drôme à Luc-en-Diois	189	1022	408	1095	598
V4225010	Le Bez à Châtillon-en-Diois	227	1275	593	1343	545
V4275910	La résurgence des Fontaigneux à Beaufort-sur-Gervanne	104	802	272	1176	635
V4455010	Le Jabron à Souspierre	79	525	320	934	716
V5004010	L'Ardèche à Pont-de-Labeaume	293	888	1771	1732	622
V5014010	L'Ardèche à Vogüé	625	710	1336	1573	666
V5045810	Le Borne à Saint-Laurent-les-Bains [Pont de Nicoulaud]	63	1115	1362	1842	574
V5046610	L'Altier à Altier [La Goulette]	103	1152	1046	1510	561
V5064010	L'Ardèche à Saint-Martin-d'Ardèche [Sauze-Saint-Martin]	2264	598	900	1403	703
V5424010	La Cèze à Bessèges	229	657	867	1627	705
V6035010	Le Toulourenc à Malaucène [Veaux]	157	847	263	1000	654
V7124010	Le Gardon de Mialet à Générargues [Roucan]	244	563	809	1583	727
V7135010	Le Gardon de Saint-Jean à Corbès [Roc Courbe]	262	537	826	1497	750
V7216510	Le Vigueirat à Tarascon [Saint-Gabriel]	258	42	389	652	865
W1024010	L'Arc à Bramans	641	2525	388	1126	289
W1034010	L'Arc à Saint-Michel-de-Maurienne [La Saussaz]	947	2424	1024	1127	312
W2314010	La Bonne à Entraigues [Pont Battant]	144	1946	918	1289	411
W2335210	La Roizonne à la Valette [La Rochette]	70	1761	1067	1586	470
W2405010	La Jonche à la Mure	51	1187	471	1407	545
W2714010	La Romanche à Mizoën [Chambon amont]	223	2391	1044	1288	316
W2767210	La Duy à Vizille	49	923	442	1478	574
W3315010	Le Meaudret à Méaudre	74	1259	269	1451	552
W3335210	L'Adouin à Saint-Martin-en-Vercors [Tourtre]	37	1443	488	1404	546
W3344010	La Bourne à Saint-Just-de-Claix [Pont de Manne]	813	1113	807	1331	582
W3355010	La Vernaison à Pont-en-Royans	289	1243	180	1356	557
X0010010	La Durance à Val-des-Prés [Les Alberts]	206	2217	780	1184	374
X0100010	La Durance à Briançon [aval]	548	2186	807	1125	372
X0130010	La Durance à l'Argentière-la-Bessée	963	2177	874	1154	380
X0310010	La Durance à Embrun [La Clapière]	2283	2105	723	1049	396
X0434010	L'Ubaye à Barcelonnette [Abattoir]	542	2213	587	1018	373
X0454010	L'Ubaye au Lauzet-Ubaye [Roche Rousse]	943	2083	684	1013	401
X0500010	La Durance à Espinasses [Serre-Ponçon]	3580	2029	697	1036	411
X1034020	Le Buech à Serres [Les Chambons]	731	1249	612	1141	564
X1225010	Le Bes à la Javie [Esclangon-Péroure]	165	1458	539	921	533
X1424010	L'Asse à Beynes [Chabrières]	382	1133	378	970	630
X2114010	L'Issole à Saint-André-les-Alpes [Mourefrey]	138	1502	563	1081	537

Annexe A. Liste des 609 bassins versants de la partie I

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P (mm/an)	E [mm/an]
X2305010	Le Jabron à Comps-sur-Artuby [Pont de l'Evescat]	67	1070	287	1026	651
Y0474030	La Têt [partielle] à Perpignan	1302	1109	233	863	679
Y0624020	L'Agly à Saint-Paul-de-Fenouillet [Clue de la Fou]	218	730	404	975	727
Y1435410	L'Argent Double à la Redorte [Les Salices]	108	425	288	1013	765
Y2015010	L'Arre au Vigan [La Terrisse]	155	664	1077	1556	713
Y2214010	La Lergue à Lodève	181	565	811	1313	729
Y2554010	L'Orb à Vieussan	912	539	766	1279	742
Y3204010	Le Lez à Montferrier-sur-Lez [Lavalette]	116	146	589	952	841
Y3464010	Le Vidourle à Marsillargues	790	153	302	1000	841
Y3514020	Le Vistre à Bernis	280	80	240	744	873
Y3534010	Le Vistre [lit mineur] au Cailar	496	70	243	741	875
Y4306510	La Chapelette à Arles [Caparon]	124	57	216	642	862
Y4624010	Le Gapeau à Hyères [Sainte-Eulalie]	536	316	228	940	839
Y5032010	L'Argens à Châteauvert	505	403	214	788	808
Y5112010	L'Argens à Carcès [aval]	1141	370	196	844	823
Y5202010	L'Argens aux Arcs	1651	363	219	849	821
Y5312010	L'Argens à Roquebrune-sur-Argens	2514	350	227	865	823
Y5514040	La Siagne à Callian [Ajustadoux]	166	956	1121	1079	678
Y6624010	La Roya à Breil-sur-Roya	453	1461	813	1177	577

Annexe B. Article “How should a rainfall-runoff model be parametrized in an almost ungauged catchment? A methodology tested on 609 catchments”

Claudia Rojas-Serna¹, Charles Perrin², Vazken Andréassian², Laure Lebecherel², Ludovic Oudin³

¹ Universidad Autónoma Metropolitana, División de Ciencias Básicas e Ingeniería, México, Mexico

² Irstea, Hydrosystems and Bioprocesses Research Unit, Antony, France

³ UMR 7619 Sisyphe, Université Pierre et Marie Curie, Paris, France

Abstract

This paper deals with catchments that are *almost* ungauged, i.e., catchments in which only a small number of point flow measurements are available. In these catchments, hydrologists may still need to simulate continuous streamflow time series using a rainfall-runoff model, and the methodology we present allows use of the few point measurements for model parameterization. Our method combines regional information (parameter sets of neighboring gauged stations) and local information (brought by the point measurements) in a framework where the relative weight of each source of information is made dependent on the number of point measurements available. This approach is tested with two different hydrological models on a set of 609 catchments. The results show that even a single measurement can improve the simulation's efficiency, while ten measurements reduce more than 50% of the performance gap between the gauged and ungauged situations.

Keywords: Rainfall-runoff modeling; Regionalization; Point flow measurements; Ungauged catchment; Neighborhood; Parameter estimation; Model parameterization

1 Introduction

1.1 Point streamflow measurements are ubiquitous

This paper deals with the use of noncontinuous streamflow measurements in hydrology. On the one hand, there are many locations in the world where it is difficult to maintain long-term flow-gauging stations and where only short series are available [see, e.g., *Bhatt and Tiwari, 2008; Kim and Kaluarachchi, 2009; Konz et al., 2007*]. On the other hand, when a hydrological question is raised for an ungauged river, practicing hydrologists may not have the time and resources to install a perennial gauging station, but they will generally have the opportunity to take a small number of flow measurements.

In Sweden, routine point flow measurements have been taken for a number of rivers (B. Arheimer, personal communication). In France, some systematic point flow measurement campaigns have been coordinated by the Rhine-Meuse Water Authority over the last two decades to improve knowledge of low flows [*Corbonnois et al., 1999; Decloux and Sary, 1991; François and Sary, 1990,1994*]. Point flow measurements are also made over a range of water quality assessment stations. These point flow measurements may be useful for a number of issues, such as low flow estimation [*Chopart and Sauquet, 2008; Eng and Milly, 2007; Goodwin and Young, 2007; Laaha and Blöschl, 2005; Oberlin et al., 1973*].

In both cases, although the hydrologist needing to calibrate a model will not consider rivers in question as *properly* gauged, they are not *strictly speaking* ungauged. In this paper, we will refer to

them as *almost ungauged*. It is true that hydrologists have rarely used point streamflow measurements for model parameterization (see discussion below). But as modelers, we consider it essential to make the most of whatever streamflow measurements are available. The framework we present in this paper allows identification of the parameters of hydrological models on the basis of point measurements.

1.2 Model parameterization is not a simple optimization problem

For many years, the issue of model parameterization has been addressed mostly from a numerical point of view, i.e., parameterization has been identified with optimization. It is now increasingly recognized that optimization is only one component of the parameter identification problem, and that local optima are not always the worst of the problems that hydrologists face. As shown by *Wagner et al.* [2003], the identifiability of parameters varies in time, with some periods showing a higher information content than others. Moreover, very often, the series available for calibration are simply not long enough to properly inform the calibration process – where information is lacking the *mathematical optimum* will be potentially different from the *hydrological optimum*, i.e., the parameter set that can reproduce catchment behavior over the long term.

When it comes to using not only short time series, but also individual streamflow measurements, the situation is obviously worse: a small number of flow measurements does not contain enough information to allow for a classic calibration [*Rojas-Serna et al.*, 2006; *Seibert and Beven*, 2009]. This information cannot be used alone in a classic calibration perspective. It needs to be associated with other sources of hydrological information.

1.3 A question that also applies to sites with long-term measurements

It is worth mentioning that the question of using other sources of information to guide parameterization is not specific to the *almost ungauged* catchments. Several authors have discussed this issue while working on well-gauged catchments. For example, *Koren et al.* [2003] found it beneficial to use the regionalization relationships established for ungauged catchments as initial estimates of parameter values. The calibration process would start with these estimates, and then be guided by the calibration criterion in a classic optimization objective. Another example is that of *Perrin et al.* [2008], who proposed a method to make the parameterization of rainfall-runoff models more robust, where only a short flow series was available for calibration (less than 2 years in their study). They replaced the traditional calibration by a search restricted to parameter sets previously identified in the region, and showed that this use of a library of regional parameter sets yielded more robust results than an unconstrained optimization. *Kuzmin et al.* [2008] also underline the advantage of starting calibration with already informative initial values, and then improving these a priori estimates during the calibration process. These authors mention that searching in the vicinity of the starting values using a local search algorithm yields better results than using a global search algorithm, especially where calibration data are lacking.

The above studies suggest that regional sources of information can be useful, even in the classic case of model calibration at a long-term gauging station. In the case of almost ungauged stations, on which this paper focuses, using the complementary information from regional hydrological knowledge becomes essential.

1.4 Scope of the paper

This article presents a method that allows hydrological model parameterization with limited streamflow information, through a combination of the partial information provided by point flow measurements with regional information.

After a critical literature review (Section 2), we present the proposed approach (Section 3), a data set of 705 catchments and two hydrological models used for evaluating the approach (Section 4). The results are presented and discussed in Section 5 and perspectives are presented in Section 6.

2 Relevant literature

2.1 Parameter estimation on entirely ungauged catchments

Predicting streamflows in ungauged basins (PUB) is a grand challenge for hydrology [Sivapalan, 2003]. A full review of ungauged basin research is out of the scope of this article. Readers are referred to recent reviews by Oudin *et al.* [2008], Reichl *et al.* [2009] or Oudin *et al.* [in press]. Here, we only mention the most common parameter estimation techniques: (1) prediction by regression equations linking physiographic descriptors with parameters (regressions to be calibrated on gauged catchments), (2) transfer of parameter sets obtained in gauged catchments that are similar in terms of physical characteristics, (3) transfer of parameter sets obtained in geographically close catchments. Many variants exist, including procedures of regional model calibration [Fernandez *et al.*, 2000; Hundecha *et al.*, 2008; Parajka *et al.*, 2007].

A few comparative studies have been carried out over the last few years to evaluate the relative efficiency of various parameter estimation procedures in ungauged conditions. It must be noted that these comparisons sometimes produced contradictory results. A few studies (e.g., Merz and Blöschl [2004] in Austria, Oudin *et al.* [2008] in France, Samuel *et al.* [in press] in Canada) found the approach of geographical proximity to perform best, whereas others (e.g., McIntyre *et al.* [2005] and Kay *et al.* [2007] in the UK) found better results with physical similarity. These apparently opposite conclusions may stem from a number of differences: location of test basins, spatial density of gauging stations, models used, available catchment descriptors, and unavoidable arbitrary choices (e.g., pooling options) in the application of methods [Oudin *et al.*, 2008].

2.2 Sensitivity of model parameter values to the length of streamflow time series

In gauged catchments, parameter estimation is rather straightforward because it can rely on flow observations: model errors can be quantified, an objective function can be defined, and optimization algorithms can be applied. However, the robustness of the identified parameter sets clearly depends on the information content of available flows. Generally, increasing the length of the flow series available for calibration tends to produce more robust parameter sets [see, e.g., Anctil *et al.*, 2004; Yapo *et al.*, 1996]. Similarly, Gill *et al.* [2007] showed that increasing the percentage of missing data in a calibration series will tend to produce less robust models. Indeed, longer time series generally encompass a larger variety of hydrological conditions, which makes the series more informative for calibration. However, shorter time series may also provide valuable information:

Brath *et al.* [2004] reported tests using calibration periods ranging from 1.5 to 12 months. In validation mode, they found that the best results were obtained using the parameters calibrated in the 12-month period, 3 months seeming to be a minimum requirement.

Kim and Kaluarachchi [2009] tested the sensitivity of a water balance to decreasing data availability using discontinuous series. They showed that short data series could produce robust parameter sets as long as some parts of the hydrograph, especially high and low flows, are included in the series.

Perrin et al. [2007] used flow data sampled out of long series to calibrate models. They showed that consistent parameter estimates could be obtained with only 350 flow data. However, model performance dropped significantly when fewer data were used during calibration.

This drop in performance caused by shortening calibration time series was efficiently attenuated in the discrete model parameterization (DISP) approach of *Perrin et al.* [2008]. In DISP, parameter estimation is made by searching the optimal parameter set within a collection of parameter sets previously obtained in other (gauged) catchments. The authors showed that this approach outperforms the classic calibration approaches when only short-flow time series are available.

2.3 Published attempts to deal with point flow measurements

The problem of parameter estimation in almost ungauged catchments has received little attention up to now. Only a few authors have investigated this issue:

Rojas-Serna et al. [2006] presented a method merging the use of regionalized parameter values and optimization, to parameterize a model when only a couple of point measurements were available.

Montanari and Toth [2007] proposed a practical approach to calibrate rainfall-runoff models with sparse data, using the Whittle estimator as a likelihood function and calibrating the model in the spectral domain.

Seibert and Beven [2009] used point flow measurements to constrain the choice of model parameters. They concluded that "*surprisingly little runoff data was necessary to identify model parameterizations that provided good results for the 'ungauged' test periods. These results indicated that a few runoff measurements can contain much of the information content of continuous runoff time series.*"

From a slightly different perspective, one can mention the work of *Winsemius et al.* [2009], who proposed a framework to integrate hard and soft data to constrain the estimation of parameters. Their results confirm the potentiality of their method to be used for scarcely gauged catchments.

2.4 Synthesis

Review of the hydrological literature suggests that (1) parameter estimation can be done using a limited number of flow observations (not necessarily contiguous but possibly spread in time), but that (2) classic optimization algorithms are insufficient for model parameterization when only a few flow data are available for calibration. Exploiting the prior knowledge gained at a regional level (in neighboring catchments with a perennial station) represents a valid alternative for parameter identification. The approach that we present in this paper specifically combines regional information (transferred from perennial gauging stations located in the geographical neighborhood) with local information (brought by the limited number of flow measurements).

3 Presentation of the parameter estimation method for almost ungauged catchments

This section outlines the proposed approach, and describes how it blends regional knowledge with point measurements, in order to address the question of parameter estimation in almost ungauged catchments.

3.1 Origin of the method

The proposed method builds on two existing approaches that we merged:

- The neighborhood approach [see, e.g., *Oudin et al.*, 2008] used for entirely ungauged basins, based on regional information: the parameter set for the target ungauged catchment is chosen among existing parameter sets, previously calibrated on gauged catchments. A distance between the ungauged catchment and its gauged neighbors can be defined either geographically (spatial proximity) or in the space of catchment descriptors (physical similarity).
- The DIScrete Parameterization (DISP) approach [*Perrin et al.*, 2008] proposed for gauged catchments: as in the previous case, the parameter set for the ungauged catchment is selected within a collection of existing parameter sets, but here the criterion for parameter set selection is the value of the model error obtained by running the model with an existing parameter set on the target catchment. This method has been preferred over classic optimization algorithms since *Perrin et al.* [2008] showed that they lead to overfitting situations on short calibration time series.

These two approaches use the same prior information (a library of parameter sets previously calibrated on gauged catchments) but differ in the way parameter sets are selected, i.e., the way they define the distance between the donor catchment (i.e., the gauged catchment providing the parameter set) and the target ungauged catchment: in the first case, the distance is either defined in the space of physical descriptors or geographically (regional information); in the second case, the distance is a function of the difference between model simulations and flow observations (local information).

3.2 Conditions of application

In the presentation below, we consider that the lumped continuous rainfall-runoff model we wish to apply has previously been calibrated on p gauged catchments, providing p parameter sets that constitute a parameter library. For these p catchments and for the studied ungauged catchment, a number of physical descriptors (e.g., catchment area, drainage density, mean slope, vegetation cover, etc.) and geographic coordinates are also available. Based on this regional information, the proposed approach can be applied using either physical similarity or spatial proximity as a basis for hypothesizing catchment similarity.

In these neighborhood methods, we are not limited to a single parameter set. We will consider a pool of M best parameter sets, as it often gives better results than selecting a single parameter set [*Goswami et al.*, 2007]. Pools of parameter sets can be used either by parameter averaging (an average parameter set is calculated from the M parameter sets before applying the model on the ungauged catchment) or by model output averaging (the model is applied on the ungauged

catchment with each of the M parameter sets before averaging the M outputs). In this paper, we use the model output averaging, as suggested by *Oudin et al.* [2008].

In terms of data availability, we assume that the continuous time series of areal precipitation and potential evapotranspiration (PE) are available in the ungauged catchment over the period where individual flow measurements are made, allowing the continuous rainfall-runoff model to be run over this period.

3.3 Proposed approach

The method outlined below was found to be the most efficient among many variants that are not presented here for the sake of brevity. *Rojas-Serna* [2005] provides full details of these other attempts. The proposed approach involves three steps:

- Step 1: Ranking the p parameter sets of the library

The parameter sets of the p gauged catchments in the library are ranked using the selected neighborhood approach (physical similarity or spatial proximity). The closest catchment (geographically or in the space of the selected physical descriptors) is given rank 1, the most remote is given rank p . We note r_j^{reg} the rank of the j^{th} parameter set in the library.

In parallel, the model is run on the ungauged catchment using each of the library's p parameter sets and the available precipitation and PE time series. For each flow simulation produced, a model error F is calculated on the N dates when flow observations are available. F is the root mean square error (RMSE) calculated on all the available flow measurements:

$$F = \sqrt{\frac{1}{N} \sum_{i=1}^N (\sqrt{Q_i} - \sqrt{C_i})^2} \quad \text{Eq (1)}$$

where Q_i and C_i are the observed discharge and the discharge calculated with the model, respectively, for the date of the i^{th} discharge measurement. Here we use root square transformed flows, as *Oudin et al.* [2006] showed this formulation of F yields all-purpose parameter sets (i.e., giving similar levels of quality in the simulation of low and high flows), but other model error formulations could be used instead. The p parameter sets are ranked by increasing model error. The parameter set providing the lowest model error is given rank 1, the one providing the largest error is given rank p . We note r_j^{loc} the rank of the j^{th} parameter set in the library.

- Step 2: Combining ranks and selecting a pool of M parameter sets

Two distances between the target catchment and the catchments in the library have been defined in the previous step: a distance in terms of neighborhood and a distance in terms of model error. Both distances contain information on the relevance of the donor catchments and we wish to combine them into a single index. As the distances are not expressed in the same unit and do not vary over the same range in both cases, it is difficult to use their absolute values in the combination. Instead, we propose to combine the ranks previously defined as follows:

$$r_j = \alpha.r_j^{reg} + (1 - \alpha).r_j^{loc} \quad \text{Eq (2)}$$

where r_j is the combined rank of the j^{th} parameter set in the library and α is a weighting coefficient (varying between 0 and 1) expressing the relative importance of the regional information compared to the local one. Its value needs to be determined empirically (see section 5.3). When α equals 1, the method comes down to the neighborhood approach. When α equals 0, the method only uses point flow information. The combined rank r takes values between 1 and p . Note that two parameter sets in the library may have the same combined rank, which means that they will be considered equivalent in the proposed method.

This procedure enables us to merge the regional information with the local information gained from point flow measurements. Using the combined rank r , a pool of the M closest parameter sets can be selected (i.e., parameter sets obtaining ranks r between 1 and M).

- **Step 3: Determining the flow in the almost ungauged catchment using the pool of M parameter sets**

Using the available precipitation and PE time series, the flow time series is simulated for the almost ungauged catchment on the basis of the selected pool of M parameter sets (applying either the parameter averaging or output averaging option).

3.4 Setting rules for practically applying the method

To practically apply the method when N flow measurements are available, two values must be defined:

- The number M of parameter sets selected in the pool to be applied on the target catchment.
- The value of α , i.e., the weighting coefficient that defines the relative importance of regional information.

We can expect the values of M and α to depend on the number N of flow measurements. Indeed, for large values of N , there will be a lot of information in the flow data, so that we ought to give more weight to the minimization of model error than to the neighborhood. Consequently, α should take values closer to 0. In Section 5, we analyze the sensitivity of the proposed approach to M and α values.

Last, we thought that model complexity (here seen as the number of parameters to calibrate) may also play a role. We will thus conduct our analysis using two different models.

4 Data, models, and assessment methodology

For a general evaluation of the method, we used a large set of catchments and two rainfall-runoff models.

4.1 Data set

Two reasons underlie our need to work with a large and varied catchment set. First, when a method is assessed on a large data set, we can have greater confidence in its applicability and robustness. Second, the method itself requires a library of parameter sets that should be large enough to be representative of the conditions of the studied zone [see the discussion by Perrin *et al.*, 2008].

We used a set of 609 catchments in France (Figure 2). These catchments have limited gaps in flow series (less than 10% for every year of the series). This set represents a variety of hydrometeorological conditions, as shown in Table 1.

Daily data were available over the 1995–2005 period in this study. Areal catchment rainfall was calculated using the SAFRAN gridded values provided by Météo-France [Quintana-Segui *et al.*, 2008]. Mean annual curves of potential evapotranspiration were computed using the formula proposed by Oudin *et al.* [2005] based on temperature. Streamflow time series were extracted from the HYDRO database of the French Ministry for Ecology. Data are generally of good quality.

4.2 Rainfall-runoff models and parameter library

The method outlined in this paper can be applied with any lumped rainfall-runoff model. Here we will use two models to get more robust conclusions and possibly analyze differences between models. We applied the four-parameter GR4J model [Perrin *et al.*, 2003] and the eight-parameter TOPM model [Michel *et al.*, 2003]. A sketch of the model structures is shown in Figure 3 and the meaning of the parameters is given in Table 2. These two models were previously used by Perrin *et al.* [2008] in a test of the discrete parameterization approach described in Section 3.

To build the library of model parameters, the models were calibrated on each catchment using the optimization algorithm applied by Edijatno *et al.* [1999]. The objective function used during optimization is the Nash and Sutcliffe [1970] criterion calculated on root square transformed flows as recommended by Oudin *et al.* [2006].

4.3 Assessment procedure

The proposed approach was successively applied to each catchment considered as ungauged. Each time, the parameter set of the studied catchment was excluded from the library to test the approach.

We applied the split-sample test scheme advised by Klemeš [1986] by splitting the available record into two periods (1995–2000 and 2000–2005) that were alternatively used for model parameter identification and model assessment in validation (parameter identification in period 1 and validation in period 2 and vice versa). For each period, the first year (i.e., 1995 and 2000, respectively) was used for model warm-up, which means that model performance was actually computed in the 1996–2000 and 2001–2005 periods (i.e., actually nonoverlapping periods).

Here the N flow measurements were randomly drawn in the flow series on each period. The random option was chosen because it corresponds quite well to the case where one collects point flow data without following a predefined acquisition strategy. The random drawing was done once per catchment. Since the number of catchments is large, we think this does not prevent obtaining robust results.

4.4 Evaluation criteria

The evaluation of the method was based on model performance obtained in validation mode as measured by the Nash and Sutcliffe [1970] criterion (NS). NS varies within the interval $]-\infty, 1]$. The lack of lower bound for this criterion is a problem when working on a large set of catchments in ungauged conditions, as the criterion may take highly negative values in some catchments where the

model fails, which prevents making a meaningful average of performance over the catchment test set. To circumvent this pitfall, percentiles of the distribution of performance over the test set (e.g., the median performance) could be used, but here we preferred to use a mean criterion that integrates all the results. Therefore we used the anamorphosis, C , of the Nash-Sutcliffe criterion proposed by *Mathevet et al.* [2006] and given by:

$$C = \frac{NS}{2 - NS} \quad \text{Eq. (3)}$$

The main value of this transformation is to have a bounded criterion pertaining to the interval $]-1, 1]$, which allows us to make meaningful performance averages over the test set. Note that this criterion keeps the same baseline as the Nash-Sutcliffe criterion ($C = 0$ when $NS = 0$), has the same optimum (1 means perfect simulation for both criteria), but yields lower positive values compared to the NS criterion (e.g., $C = 0.67$ when $NS = 0.80$).

The efficiency of the parameter estimation method will be assessed by the mean of C values obtained by the hydrological model over all validation tests (here 1218, i.e., twice the number of catchments), calculated on flows (C_Q), on logarithm transformed flows (C_{LQ}), which put more weight on high and low flows, respectively, and on root square transformed flows, which put same weight on all type of flows (C_{RQ}).

4.5 Reference methods

The model performance obtained by applying the proposed method was compared with the model performance obtained:

1. In ungauged conditions: we applied the approach of spatial proximity with the output averaging pooling option, as shown by *Oudin et al.* [2008] to be the most efficient method in our French test set. The distance separating ungauged catchments and neighbors considered here is a distance combining distance to the outlet and distance to the centroid. Precedent tests showed that used this distance was beneficial to transfer information between catchments of different sizes. The distance considered is thus a "hydrological distance" defined as:

$$d = 0.2 \times d_{outlet} + 0.8 \times d_{centroid}$$

2. In fully gauged conditions: we applied the optimization algorithm mentioned above using all data available in the calibration period.

5 Results

In this section, we present the results of the proposed approach for almost ungauged catchments. We discuss the sensitivity of the method to the value of the weighting factor of the regional information (α) and to the number (M) of parameter sets in the pool applied to the ungauged catchment. We use two different models and an increasing number of flow measurements (N). After

determining the optimum values, we evaluate the performance of the proposed approach, which we compare with the reference methods.

5.1 Results of reference methods

Figure 4 shows that the best results in ungauged conditions using the neighborhood approach are obtained with a small pool of parameter sets: seven parameter sets for GR4J and nine parameter sets for TOPM. This difference between the two models is similar to that found by *Oudin et al.* [2008].

Figure 4 shows the mean efficiency obtained by the two models over the catchment set for ungauged and fully gauged conditions. Efficiencies obtained by the two models are close in ungauged conditions. GR4J is somewhat better than TOPM in fully gauged conditions, which corroborates the results found by *Oudin et al.* [2008]. Note that in gauged mode, models are quite efficient on average in the 609 catchments as they reach *C2M* efficiencies of around 0.60–0.65 (equivalent to 0.75–0.79 in terms of the Nash-Sutcliffe criterion). As expected, the efficiency obtained in fully gauged conditions is far better than the one obtained in fully ungauged conditions: this sets the range of improvement that can be obtained using flow information.

5.2 Sensitivity to the number M of neighbors in almost ungauged conditions

We evaluated the sensitivity of the proposed approach to the number M of parameter sets selected in the pool applied to the ungauged catchment. To simplify the presentation, we fixed the coefficient α at 0.5 (the results are very similar for other α values). We tested the two models using an increasing number of donor sets ($M = 1, 2, \dots, 9, 10, 15$) and an increasing number of flow measurements ($N = 1, 2, \dots, 6, 7, 10, 20, 50$). Figure 5 presents the evolution of mean model efficiency over the catchments (C_Q , C_{LQ} and C_{RQ}) with the number M of donor sets in sparsely gauged conditions for the GR4J and TOPM models for an increasing number of flow measurements (from 1 to 50). It shows that as soon as a few donor sets are used, the results stabilize and are no longer very sensitive to M . The same behavior is observed whatever the number N of point flow measurements, the model, or the criterion. For a given model, the optimum number of donors does not seem to depend much on N . Therefore we chose to set the number of donor sets to the same value identified in ungauged conditions: $M = 7$ for GR4J and $M = 9$ for TOPM.

5.3 Sensitivity to the weighting factor α

We evaluated the sensitivity of the proposed approach to the value of weighting factor α , which sets the weight of the regional information. We can expect that the larger the number of point flow measurements, the lower the value of α , given that the flow measurements will yield more information. We tested values of α from 0 (no use of regional information) to 1 (no use of local information). Each value of α was tested using different numbers of flow measurements (from 1 to 50 as in the previous section). The results are shown in Figure 6.

When α equals 1 (i.e. using only regional information), the results are like those found previously for the strictly ungauged catchments. When α equals 0 (i.e. using only local flow information), model efficiency progressively increases with the number of flow measurements. The performance exceeds the performance obtained in ungauged conditions as soon as five flow measurements are available for GR4J and TOPM.

Figure 6 presents the evolution of mean model efficiency (C_Q , C_{LQ} and C_{RQ}) over the catchments with the weighting factor α of the regional information in sparsely gauged conditions for the GR4J and TOPM models for an increasing number of flow measurements (from 1 to 50). Figure 6 also shows that an optimum value of α can be identified in the efficiency curves drawn for each value of N . This value is very close to 1 when a single flow measurement is available, which means that it is better to mostly trust regional information in these conditions. However, it does not strictly equal 1: this indicates that some information can be gained even from a single flow measurement. Note that for all values of N , the increase in model performance when α departs from 1 is very rapid. This means that flow measurements yield very different and complementary information compared to solely regional information. This can be linked to the finding of *Seibert and Beven* [2009], who found that even a few flow measurements already contain valuable information.

When N increases, the optimum α value progressively decreases. This can be expected, as the regional information becomes less and less relevant in comparison to the local (flow) measurements. For each value of N , we identified this optimum value and plotted it against N (see Figure 7). It can be seen that the patterns are quite similar for both models and both efficiency criteria. However, the optimum value is not exactly the same in all cases. The differences between the two models may come from the fact that the neighborhood approach does not give equally valuable information. The differences between criteria for a given model may come from the fact that the respective value of the flow measurements and neighborhood approach may depend on the flow level.

It seems therefore that the optimal settings of the proposed approach in terms of the relative importance of the regional information may partly depend on the model and the modeling objectives. However, we can give a general formulation to determine the value of α for a given number N of flow measurements available:

$$\alpha = \exp \left[- \left(\frac{N}{a} \right)^b \right] \quad \text{Eq. (4)}$$

where a and b are two coefficients to be determined for the model and the objectives. The values of a and b for the tests shown here are given in Table 4 and the corresponding curves are shown in Figure 7.

5.4 Performance of the approach

Figure 8 shows the efficiency of the GR4J and TOPM models obtained by applying the proposed approach for almost ungauged conditions to our data set. It exploits both local flow measurements and regional information, using the number M of neighbors and the optimum values of α determined previously. It is compared with the two reference methods (the neighborhood for the ungauged case and the optimization algorithm for the fully gauged case).

On average, the proposed approach efficiently uses the information provided by regional information and local flow measurements, since it proves more efficient than the reference method for both efficiency criteria (C_Q , C_{LQ} and C_{RQ}).

These results show that the method proposed here can effectively exploit the two sources of information. It is particularly interesting when only a few flow measurements are available. Even

when a single flow measurement is available, some improvement in model performance can be obtained in comparison with sole use of the regional information.

For the criteria C_{LQ} , the proposed approach provides more efficient results than the fully gaged case for both models. This can be surprising. It can be explained by the use of different criteria between the parameters calibration (C_{RQ}) and the evaluation of the method (C_{LQ}). Here, the fully gaged value is not the optimal validation value for this criteria.

For the TOPM model, the proposed approach improves too the results of the fully gaged case for the criteria C_{RQ} . In this case, the reason is not the use of different criterion between parameters calibration and method evaluation, because they are the same. These high performances can be explained by the use of multiple sets of parameters to simulate flows on almost ungauged catchments (Figure 9). Furthermore, the period of parameters calibration is not the same as method evaluation. The optimal set of parameters of the period calibration is not inevitably the optimal set of parameters of the period validation. Then, higher performances obtained with the proposed approach can be founded by the selection in the library of set parameters, of better parameters for this evaluation period. The model robustness is no more assured. This observation is particularly observed for models with a large numbers of parameters, as TOPM model and its 8 parameters. These results have been yet obtained for this model by Perrin et al. [2008].

6 Concluding remarks

In this article, a combined approach was devised to simultaneously exploit regional information and local measurements. Tested on a set of 609 French catchments, the proposed approach weights the two sources of information depending on the availability of point flow measurements. It provides simulations intermediate between fully ungauged and fully gaged situations, and proved efficient both in high and low flows. This approach opens ways to make model parameter estimation more reliable in all the catchments where only few flow measurements are available. As more flow measurements become available, the model becomes more robust and efficient, making the parameter estimation procedure become a "learning process," as suggested by *Buytaert and Beven* [2009].

The method was tested successfully using two rainfall-runoff models. Although the settings of the method are sensitive to both the model used and the modeling objective, the method is general and flexible enough to be efficiently applied to other models. One could also imagine using other neighborhood approaches (e.g., a physical similarity-based ranking instead of a spatial proximity one). The fact that our approach was tested on a large data set is a further argument to convince end-users of its generality [*Andréassian et al.*, 2007].

Among the perspectives for this work, we see very concrete hydrological issues (such as the estimation of reference flows to describe the high, medium, and low flows in a catchment), as well as more theoretical ones. We think, for example, that the notion of *almost ungauged catchment* could be extended, for instance, in a context of nonstationarity (a long *past* record could then still be considered insufficient if a *recent* change has occurred).

As flow is the target variable in most modeling studies, we believe that acquiring point flow measurements can be the best value for money in terms of field measurements to improve parameter estimation, as also recently suggested by *Seibert and Beven* [2009]. However, further

comparative tests involving other types of observations should be made to come to definite conclusions in this respect.

In these tests, flow information was considered as taken randomly in the flow series, irrespective of the flow magnitude. As shown by *Wagener et al.* [2003], it is likely that high, medium, or low flows will not be similarly informative for the estimation of the various model parameters. Therefore, the definition of strategies for acquiring point flow data could help in maximizing the information content of each of them. Some authors have already investigated this issue [see, e.g., *Clausen*, 1995; *Juston et al.*, 2009]. This could be adapted to the context of the approach proposed here.

7 Acknowledgements

The authors wish to thank Météo-France and SCHAPI for providing meteorological and hydrological data sets, respectively.

8 References

- Anctil, F., C. Perrin, and V. Andréassian (2004), Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models, *Environmental Modelling & Software*, *19*, 357-368, DOI: 310.1016/S1364-8152(1003)00135-X.
- Andréassian, V., J. Lerat, C. Loumagne, T. Mathevet, C. Michel, L. Oudin, and C. Perrin (2007), What is really undermining hydrologic science today?, *Hydrol. Processes*, *21*, 2819-2822, DOI: 2810.1002/hyp.6854.
- Bhatt, V. K., and A. K. Tiwari (2008), Estimation of peak streamflows through channel geometry, *Hydrol. Sci. J.-J. Sci. Hydrol.*, *53*, 401-408.
- Brath, A., A. Montanari, and E. Toth (2004), Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model, *J. Hydrol.*, *291*, 232-253.
- Buytaert, W., and K. Beven (2009), Regionalization as a learning process, *Water Resour. Res.*, *45*, W11419.
- Chopart, S., and É. Sauquet (2008), Usage des jaugeages volants en régionalisation des débits d'étiage, *Revue des Sciences de l'Eau*, *21*, 267-281.
- Clausen, B. (1995), Discharge Data-Collection and Analysis Strategies in Low-Flow Studies, *Nordic Hydrol.*, *26*, 191-204.
- Corbonnois, J., D. François, I. Arts, M. Sary, E. Gille, P. Gamez, A. Devos, and J. P. Decloux (1999), Répartition des écoulements en basses eaux. Aspects méthodologiques. (Spatialization of low water flows. Methodology), paper presented at XXII IUGG General Assembly, 18-30 July 1999, Birmingham (UK).
- Decloux, J. P., and M. Sary (1991), Campagnes d'étiage. Objectifs, traitement et valorisation des données, *Mosella*, *XXIII*, 121-134.
- Edijatno, N. O. Nascimento, X. Yang, Z. Makhlof, and C. Michel (1999), GR3J: a daily watershed model with three free parameters, *Hydrol. Sci. J.*, *44*, 263-277, DOI: 210.1080/02626669909492221.
- Eng, K., and P. C. D. Milly (2007), Relating low-flow characteristics to the base flow recession time constant at partial record stream gauges, *Water Resour. Res.*, *43*.
- Fernandez, W., R. M. Vogel, and A. Sankarasubramanian (2000), Regional calibration of a watershed model, *Hydrol. Sci. J.*, *45*, 689-707.

- François, D., and M. Sary (1990), Etude méthodologique des basses eaux. estimation minimale des points de mesure, 14 pp, Cegum, Université de Metz.
- François, D., and M. Sary (1994), Etude méthodologique des débits d'étiage, 30 pp, Cegum, Université de Metz.
- Gill, M. K., T. Asefa, Y. Kaheil, and M. McKee (2007), Effect of missing data on performance of learning algorithms for hydrologic predictions: Implications to an imputation technique, *Water Resour. Res.*, 43.
- Goodwin, T., and A. Young (2007), Reducing uncertainty for the estimation of hydropower potential - the use of local data, paper presented at British Hydropower Association Conference, 11 October 2007, www.british-hydro.org, Birnam, Australia.
- Goswami, M., K. M. O'Connor, and K. P. Bhattarai (2007), Development of regionalisation procedures using a multi-model approach for flow simulation in an ungauged catchment, *J. Hydrol.*, 333, 517-531.
- Hundecha, Y., T. Ouarda, and A. Bardossy (2008), Regional estimation of parameters of a rainfall-runoff model at ungauged watersheds using the "spatial" structures of the parameters within a canonical physiographic-climatic space, *Water Resour. Res.*, 44, W01427.
- Juston, J., J. Seibert, and P. O. Johansson (2009), Temporal sampling strategies and uncertainty in calibrating a conceptual hydrological model for a small boreal catchment, *Hydrol. Processes*, 23, 3093-3109.
- Kay, A. L., D. A. Jones, S. M. Crooks, T. R. Kjeldsen, and C. F. Fung (2007), An investigation of site-similarity approaches to generalisation of a rainfall-runoff model, *Hydrology and Earth System Sciences*, 11, 500-515.
- Kim, U., and J. J. Kaluarachchi (2009), Hydrologic model calibration using discontinuous data: an example from the upper Blue Nile River Basin of Ethiopia, *Hydrol. Processes*, 23, 3705-3717.
- Klemeš, V. (1986), Operational testing of hydrological simulation models, *Hydrol. Sci. J.*, 31, 13-24.
- Konz, M., S. Uhlenbrook, L. Braun, A. Shrestha, and S. Demuth (2007), Implementation of a process-based catchment model in a poorly gauged, highly glacierized Himalayan headwater, *Hydrology and Earth System Sciences*, 11, 1323-1339.
- Koren, V., M. Smith, and Q. Duan (2003), Use of a priori parameter estimates in the derivation of spatially consistent parameter sets of rainfall-runoff models, *Calibration of Watershed Models, Water Science and Application*, 6, 239-254.
- Kuzmin, V., D. J. Seo, and V. Koren (2008), Fast and efficient optimization of hydrologic model parameters using a priori estimates and stepwise line search, *J. Hydrol.*, 353, 109-128.
- Laaha, G., and G. Blöschl (2005), Low flow estimates from short stream flow records - a comparison of methods, *J. Hydrol.*, 306, 264-286.
- Le Moine, N., V. Andréassian, C. Perrin, and C. Michel (2007), How can rainfall-runoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments, *Water Resour. Res.*, 43, W06428, doi: 06410.01029/02006WR005608.
- Mathevet, T., C. Michel, V. Andréassian, and C. Perrin (2006), A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins, in *Large sample basin experiments for hydrological model parameterisation: Results of the Model Parameter Experiment - MOPEX*, edited by V. Andréassian, et al., pp. 211-219, IAHS Red Books Series n°307.
- McIntyre, N., H. Lee, H. Wheeler, A. Young, and T. Wagener (2005), Ensemble predictions of runoff in ungauged catchments, *Water Resour. Res.*, 41.

- Merz, R., and G. Blöschl (2004), Regionalisation of catchment model parameters, *J. Hydrol.*, 287, 95-123.
- Michel, C., C. Perrin, and V. Andréassian (2003), The exponential store: a correct formulation for rainfall-runoff modelling, *Hydrol. Sci. J.*, 48, 109-124, DOI: 110.1623/hysj.1648.1621.1109.43484.
- Montanari, A., and E. Toth (2007), Calibration of hydrological models in the spectral domain: an opportunity for scarcely gauged basins?, *Water Resour. Res.*, 43, W05434.
- Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models. Part I - A discussion of principles., *J. Hydrol.*, 27, 282-290.
- Oberlin, G., G. Galéa, and J. T. Toni (1973), Intérêt des jaugeages épisodiques pour l'estimation des étiages de petits bassins non équipés, *Cahiers de l'ORSTOM, série Hydrologie*, X, 349-368.
- Oudin, L., V. Andréassian, T. Mathevet, C. Perrin, and C. Michel (2006), Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations, *Water Resour. Res.*, 42, W07410, doi: 07410.01029/02005WR004636.
- Oudin, L., V. Andréassian, C. Perrin, C. Michel, and N. Le Moine (2008), Spatial proximity, physical similarity and ungauged catchments: confrontation on 913 French catchments, *Water Resour. Res.*, 44, W03413, doi:03410.01029/02007WR006240.
- Oudin, L., F. Hervieu, C. Michel, C. Perrin, V. Andréassian, F. Anctil, and C. Loumagne (2005), Which potential evapotranspiration input for a rainfall-runoff model? Part 2 - Towards a simple and efficient PE model for rainfall-runoff modelling, *J. Hydrol.*, 303, 290-306, DOI: 210.1016/j.jhydrol.2004.1008.1026.
- Oudin, L., A. L. Kay, V. Andréassian, and C. Perrin (in press), Are seemingly physically similar catchments truly hydrologically similar?, *Water Resour. Res.*, doi:10.1029/2009WR008887.
- Parajka, J., G. Blöschl, and R. Merz (2007), Regional calibration of catchment models: potential for ungauged catchments, *Water Resour. Res.*, 43, W06406.
- Perrin, C., V. Andréassian, C. Rojas-serna, T. Mathevet, and N. Le Moine (2008), Discrete parameterization of hydrological models: evaluating the use of parameter sets libraries over 900 catchments, *Water Resour. Res.*, 44, doi:10.1029/2007WR006579.
- Perrin, C., C. Michel, and V. Andréassian (2003), Improvement of a parsimonious model for streamflow simulation, *J. Hydrol.*, 279, 275-289, DOI: 210.1016/S0022-1694(1003)00225-00227.
- Perrin, C., L. Oudin, V. Andréassian, C. Rojas-Serna, C. Michel, and T. Mathevet (2007), Impact of limited streamflow knowledge on the efficiency and the parameters of rainfall-runoff models, *Hydrol. Sci. J.*, 52, 131-151, DOI: 110.1623/hysj.1652.1621.1131.
- Quintana-Segui, P., P. Le Moigne, Y. Durand, E. Martin, F. Habets, M. Baillon, C. Canellas, L. Franchisteguy, and S. Morel (2008), Analysis of near-surface atmospheric variables: Validation of the SAFRAN analysis over France, *Journal of Applied Meteorology and Climatology*, 47, 92-107.
- Reichl, J. P. C., A. W. Western, N. R. McIntyre, and F. H. S. Chiew (2009), Optimization of a similarity measure for estimating ungauged streamflow, *Water Resour. Res.*, 45.
- Rojas-Serna, C. (2005), Quelle connaissance hydrométrique minimale pour définir les paramètres d'un modèle pluie-débit ?, Thèse de Doctorat thesis, 319 pp, Cemagref (Antony), ENGREF (Paris), France.
- Rojas-Serna, C., C. Michel, C. Perrin, and V. Andréassian (2006), Ungauged catchments: How to make the most of a few streamflow measurements?, *Large sample basin experiments for*

hydrological model parameterisation: Results of the Model Parameter Experiment, IAHS Red Books Series n°307, 230-236.

Samuel, J., P. Coulibaly, and R. Metcalfe (in press), Estimation of Continuous Streamflow in Ontario Ungauged Basins: Comparison of Regionalization Method, *Journal of Hydrologic Engineering*.

Seibert, J., and K. J. Beven (2009), Gauging the ungauged basin: how many discharge measurements are needed?, *Hydrology and Earth System Sciences, 13*, 883-892.

Wagener, T., N. McIntyre, M. J. Lees, H. S. Wheater, and H. V. Gupta (2003), Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis, *Hydrol. Processes, 17*, 455-476.

Winsemius, H. C., B. Schaefli, A. Montanari, and H. H. G. Savenije (2009), On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information, *Water Resour. Res., 45*, W12422.

Yapo, P. O., H. V. Gupta, and S. Sorooshian (1996), Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data, *J. Hydrol., 181*, 23-48.

Table 1. Main characteristics of the 609 catchments used to test the approach.

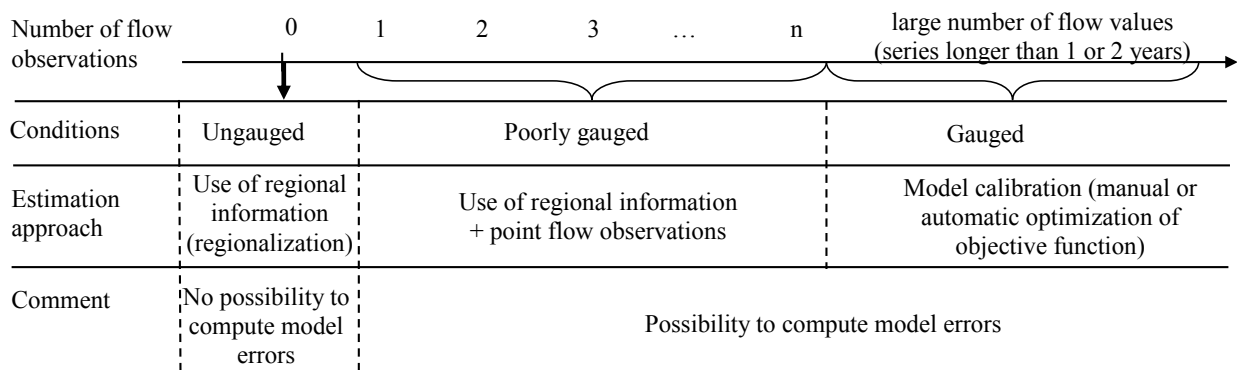
Quantiles	0.05	0.25	0.5	0.75	0.95
Catchment area (km ²)	34	109	270	833	4514
Mean elevation catchment (m)	87	180	375	781	1350
Annual rainfall, P (mm/year)	714	863	1003	1209	1688
Annual PE (mm/year)	533	616	655	687	782
Annual discharge, Q (mm/year)	159	272	411	643	1308

Table 2. List of parameters of the GR4J and TOPM models

	Parameter	Meaning
GR4J	X1	Capacity of the production store (mm; positive)
	X2	Water exchange coefficient (mm; positive or negative)
	X3	Capacity of the nonlinear routing store (mm; positive)
	X4	Unit hydrograph time base (day; ≥ 0.5)
TOPM	X1	Parameter of the exponential store (mm; positive)
	X2	Evapotranspiration parameter (mm; positive or negative)
	X3	Topographic index distribution parameter (mm; positive)
	X4	Pure time delay (day; ≥ 1)
	X5	Topographic index distribution parameter (mm; positive or negative)
	X6	Evapotranspiration parameter (mm; positive)
	X7	Capacity of the interception store (mm; positive)
	X8	Capacity of the routing store (mm; positive)

Table 3. Coefficient values to be used in Eq. (4) for the GR4J and TOPM tests with two criteria (C_{LQ} , C_{RQ} and C_Q)

Model	C_{LQ}		C_{RQ}		C_Q	
	a	b	a	b	a	b
GR4J	5.1	2.7	8.5	1.7	14	2.2
TOPM	6	1.9	7.7	1.2	13.9	1.4

**Figure 1. Comparison of the ungauged, sparsely gauged, and gauged conditions for the estimation of model parameters (n corresponds to the number of flow observations under which the application of classic estimation procedures are no more robust)**

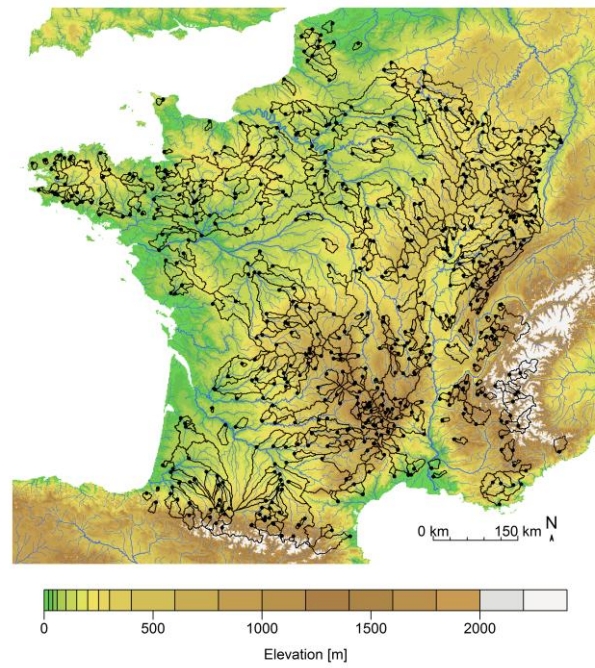


Figure 2. Location of the 609 catchments used in this study

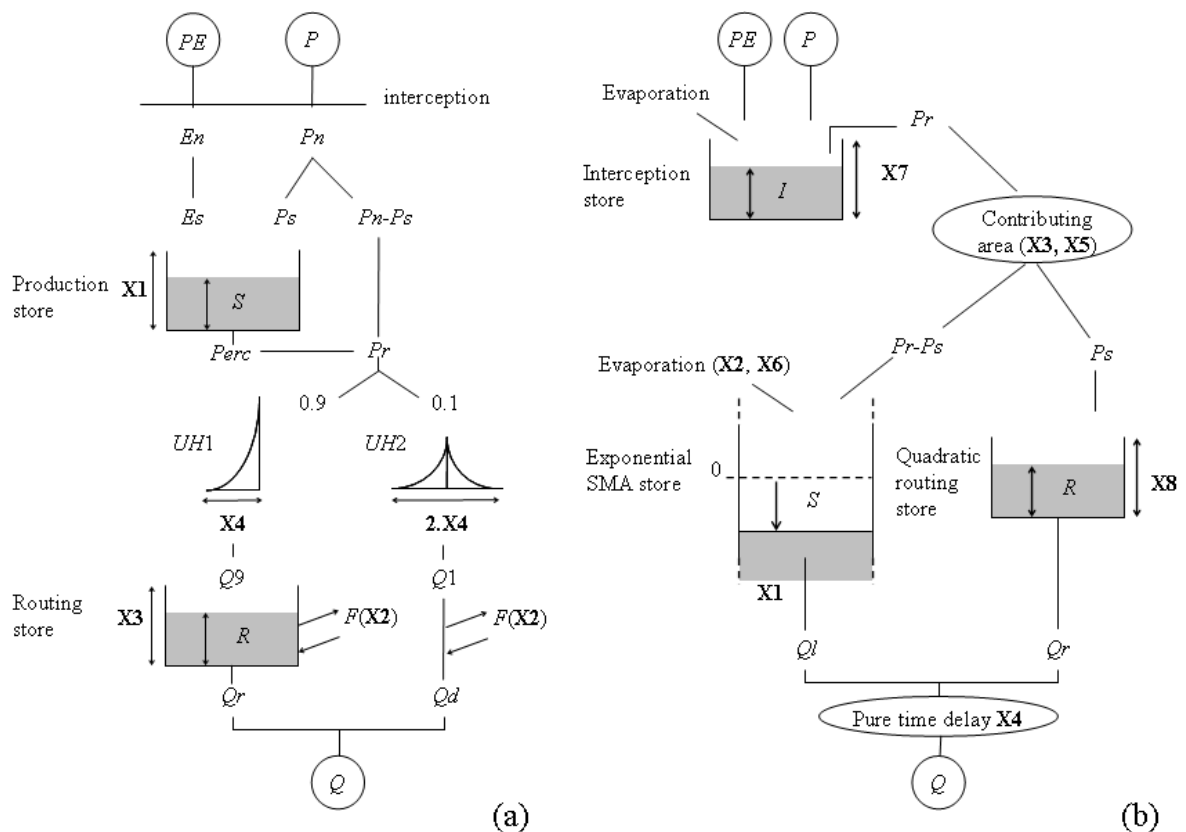


Figure 3. Schemes of the GR4J (a) and TOPM (b) rainfall-runoff models (PE: potential evapotranspiration; P: precipitation; Q: streamflow; Xi: model parameter i; other letters are internal state variables)

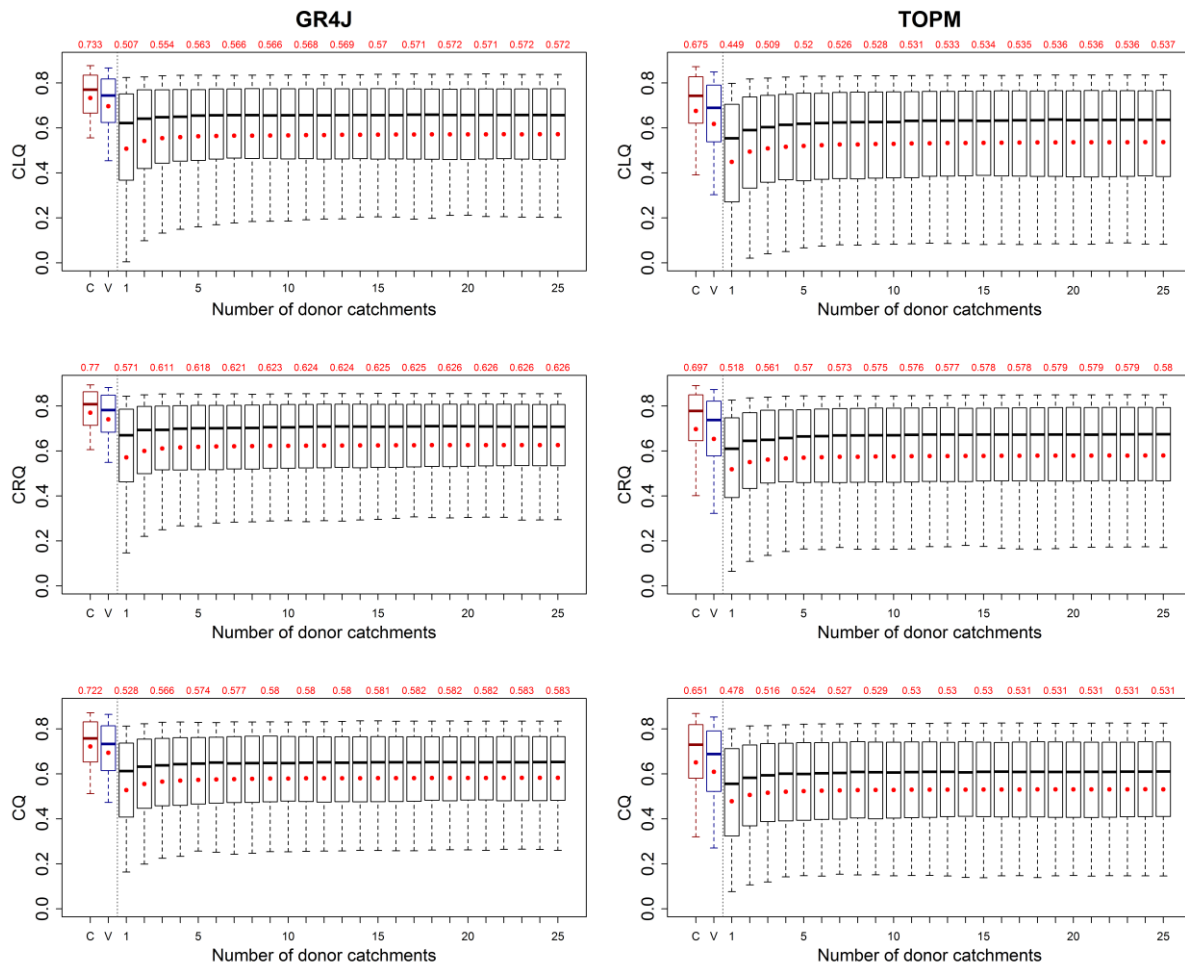


Figure 4. Distribution of the evolution of model efficiency in calibration and validation in gauged conditions and in validation in ungauged conditions with the number of donor catchments.

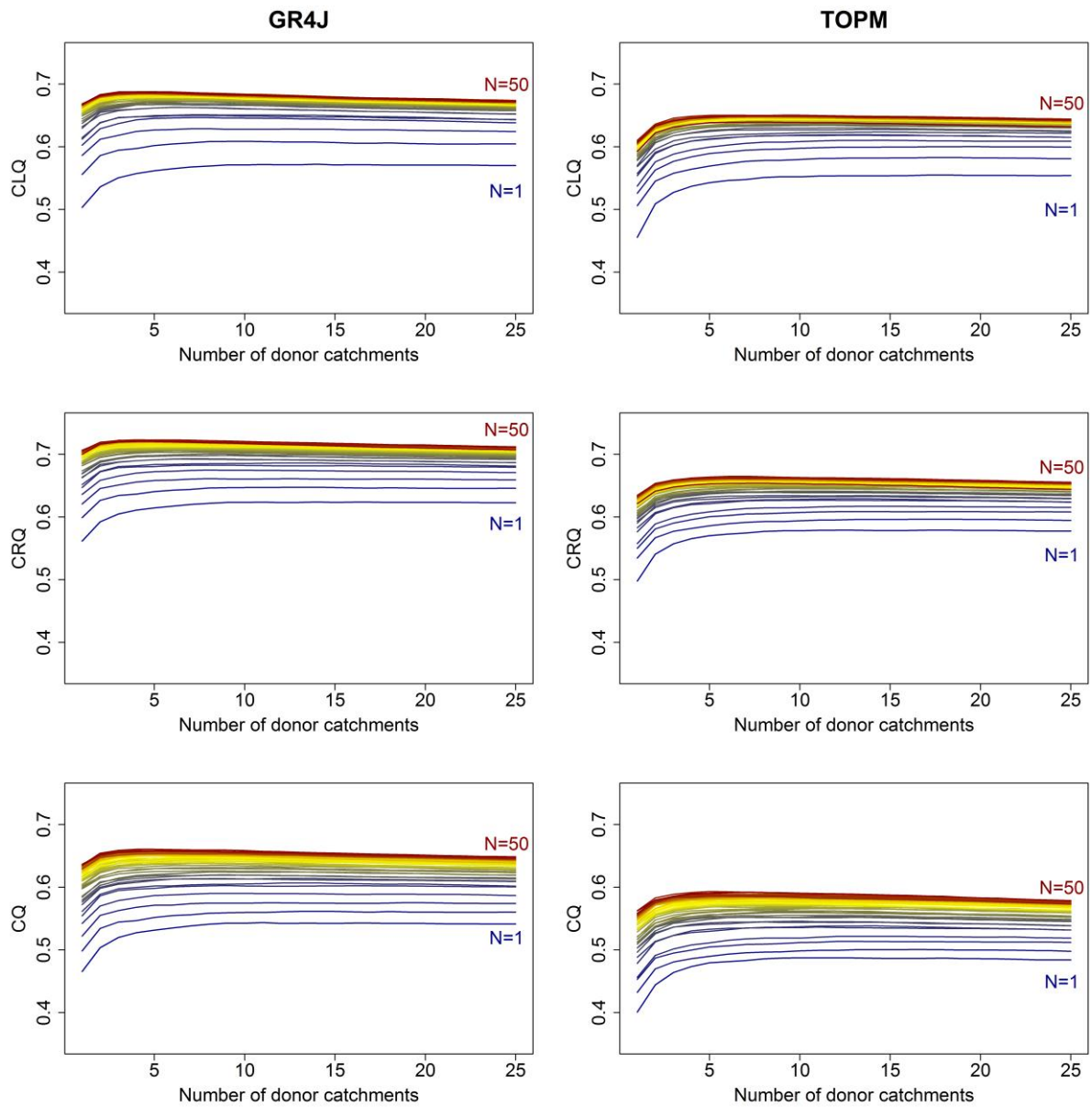


Figure 5. Evolution of the mean model efficiency over the catchments (C_{LQ} , C_{RQ} and C_Q) with the number M of donor sets in sparsely gauged conditions for the GR4J and TOPM models for an increasing number of flow measurements (from 1 to 50). Coefficient α is set to 0.5.

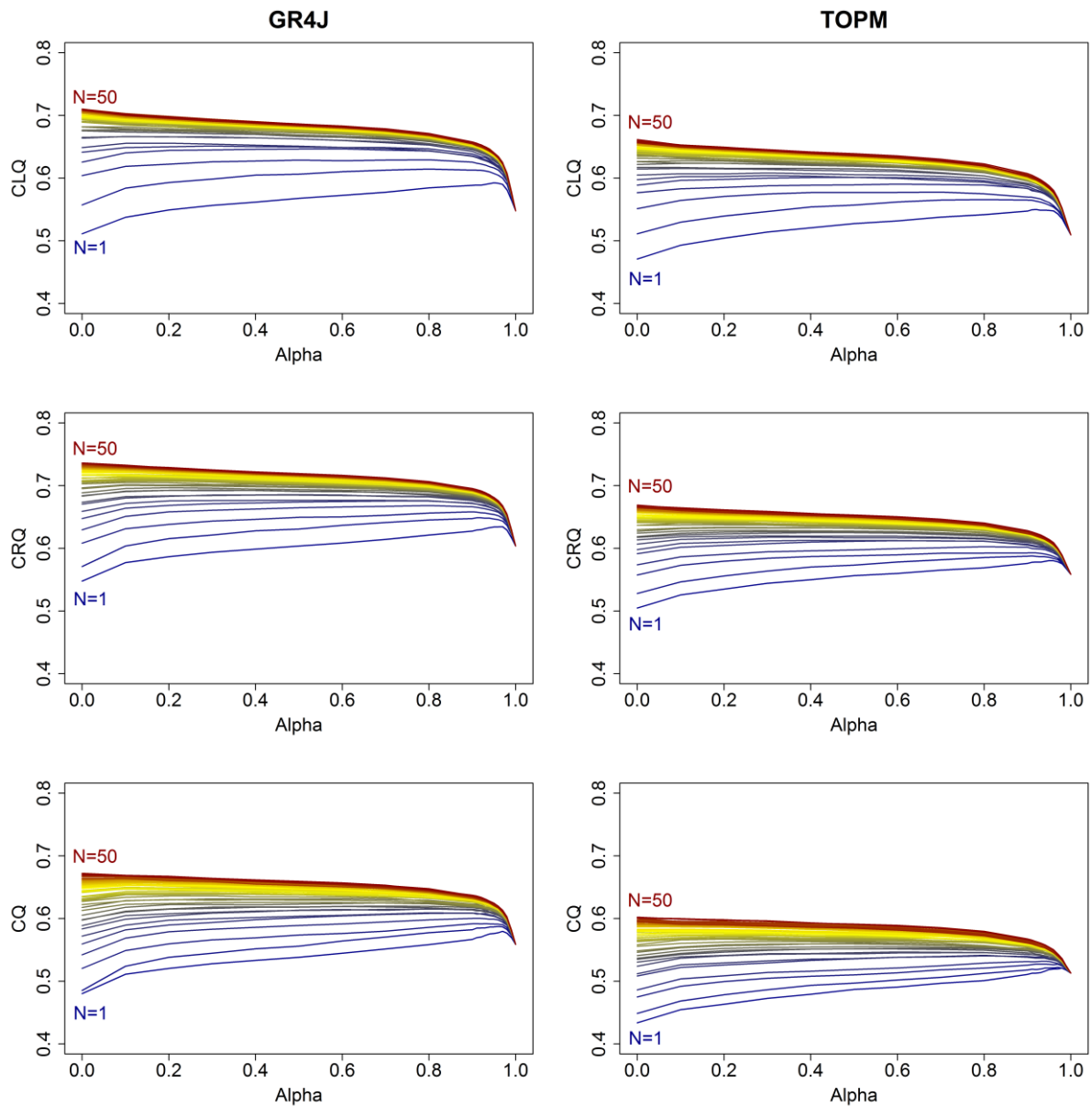


Figure 6. Evolution of mean model efficiency (C_{LQ} , C_{RQ} and C_Q) over the catchments with the weighting factor α of the regional information in sparsely gauged conditions for the GR4J and TOPM models for an increasing number of flow measurements (from 1 to 50). The number N of donor catchments is set to 7 and 9 for GR4J and TOPM, respectively.

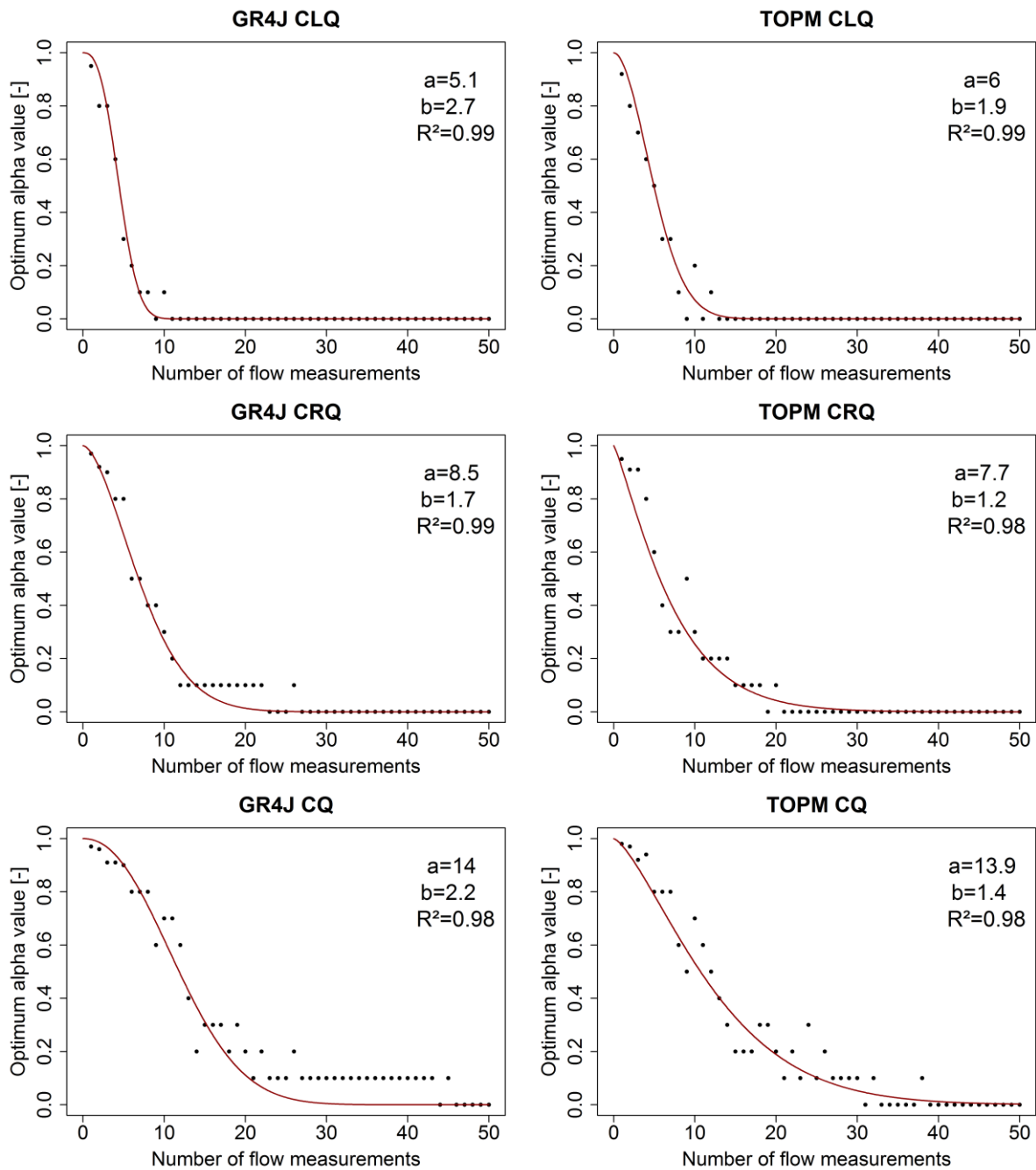


Figure 7. Optimum values of the weighting factor α of the regional information in sparsely gauged conditions for the GR4J and TOPM models for increasing number of flow measurements (from 1 to 50). The number N of donor catchments is set to 7 and 9 for GR4J and TOPM, respectively. The dashed line corresponds to the curve defined in Eq. (4).

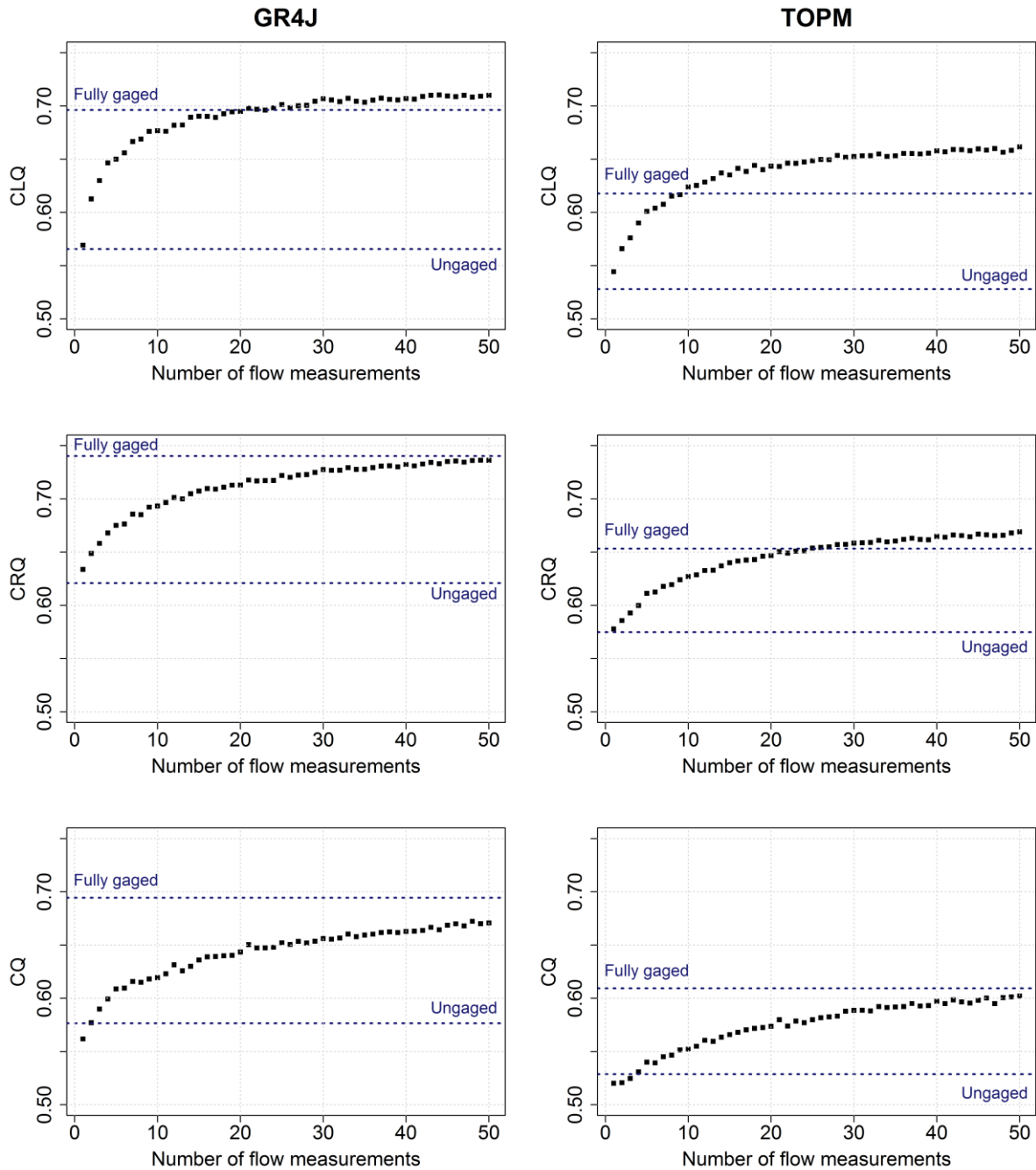


Figure 8. Efficiency of the proposed approach for sparsely gauged catchments compared to the ungauged and gauged conditions for the GR4J and TOPM models for increasing number of flow measurements (from 1 to 50). The number N of donor catchments is set to 7 and 9 for GR4J and TOPM, respectively.

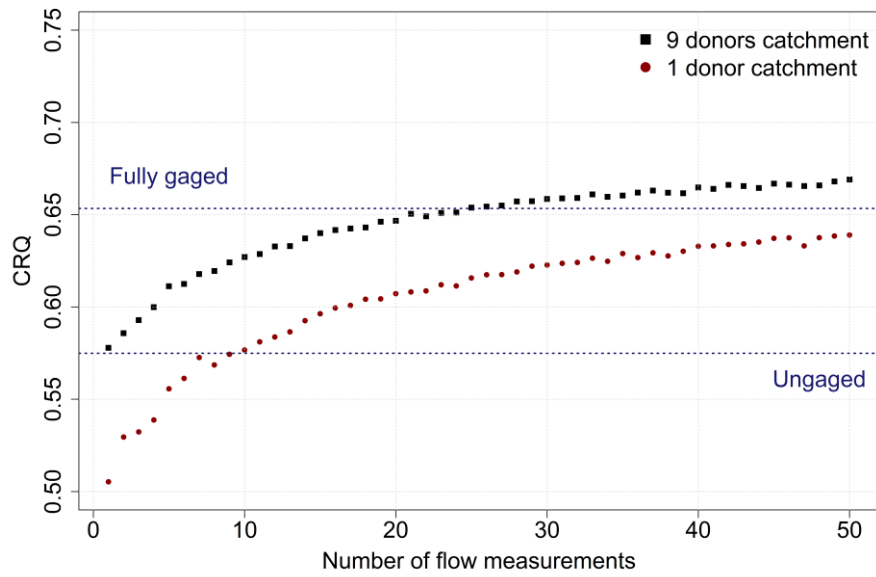


Figure 9. Efficiency of the proposed approach for sparsely gauged catchments compared to the ungauged and gauged conditions for the TOPM models and for the evaluation criteria C_{RQ} , for increasing number of flow measurements (from 1 to 50). The number N of donor catchments varies between the black points (9 donors) and the red points (only one donor).

Annexe C. Liste des 30 bassins utilisés pour la partie II

Tableau 15. Liste des 30 bassins versants utilisés pour la partie II sur l'impact de la densité du réseau pluviométrique en hydrologie. S désigne la surface, Z l'altitude moyenne des bassins versants et Q, P et ETP désignent respectivement leurs moyennes annuelles de débit, de pluie et d'évapotranspiration potentielle. FSP désigne la fraction solide moyenne des précipitations sur le bassin versant, N_{PLU} le nombre de pluviomètres disponibles pour chaque bassin et S_{PLU} la surface pluviométrique de chaque bassin.

Code	Nom de la station	S [km ²]	Z [m]	Q [mm/an]	P [mm/an]	ETP [mm/an]	FSP [-]	N _{PLU} [-]	S _{PLU} [km ²]
A1080330	L'Ill à Didenheim	668	390	309	932	659	0,051	7	5053
A5431010	Le Madon à Pulligny	948	328	357	906	641	0,044	7	641
A9091050	La Sarre à Keskastel	881	318	337	952	647	0,047	9	377
A9221010	La Sarre à Sarreinsming	1760	292	325	907	650	0,044	13	669
H0100010	La Seine à Nod-sur-Seine	373	404	389	911	653	0,044	7	184
H0321030	L'Ource à Autricourt	574	353	350	894	658	0,039	12	209
H0503010	L'Hozain à Buchères [Courgerennes]	249	170	180	760	674	0,026	7	141
H2412010	L'Armançon à Quincy-le-Vicomte	478	354	223	872	661	0,037	11	182
H6412020	La Vesle à Châlons-sur-Vesle	828	131	160	710	658	0,025	11	326
H7401010	L'Oise à Sempigny	4320	137	251	804	640	0,027	12	907
H9021010	L'Eure à Saint-Luperce	315	224	159	731	652	0,024	7	272
I3131010	L'Orne à Rabodanges	1019	213	247	817	645	0,022	12	214
J7483010	La Seiche à Bruz [Carcé]	809	70	183	737	693	0,009	7	410
K1321810	L'Arroux à Etang-sur-Arroux [Pont du Tacot]	1792	431	397	979	665	0,042	19	208
K5653010	L'Auron à Bourges [L'Ormediot]	575	192	205	775	706	0,019	15	134
L2321610	Le Clain à Vivonne [Danlot]	1824	146	247	819	723	0,01	13	479
L4653010	La Bouzanne à Velles [Forges]	438	214	229	811	708	0,018	10	146
O2344010	Le Girou à Cépet	524	215	145	776	803	0,01	8	296
O7874010	Le Dourdou à Conques	545	507	431	1084	723	0,022	8	288
P3502510	La Corrèze à Tulle [Pont des soldats]	354	601	928	1405	655	0,043	7	290
P3922510	La Corrèze à Brive-la-Gaillarde [Le Prieur]	954	473	696	1315	692	0,028	12	293
P4271010	La Vézère à Campagne	3657	371	508	1166	718	0,021	28	287
P7001510	L'Isle à Bassilac [Charrieras]	1859	289	442	1071	731	0,015	10	590
P7041510	L'Isle à Périgueux	2111	279	408	1053	735	0,014	13	581
U0924010	La Vingeane à Saint-Maurice-sur-Vingeane	397	347	314	917	656	0,042	7	327
U0924020	La Vingeane à Oisilly	609	314	310	892	668	0,038	8	403
U3214010	La Grosne à Jalogny [Cluny]	334	470	380	947	676	0,041	9	159
U3225010	La Guye à Sigy-le-Châtel [Corcelles]	271	330	240	886	689	0,031	7	312
U4644010	L'Azergues à Lozanne	798	517	301	922	686	0,043	17	130
V1774010	La Bourbre à Tignieu-Jameyzieu	696	375	343	1042	719	0,033	9	355

Annexe D. Réseau pluviométrique disponible pour chacun des 30 bassins versants utilisés pour la partie II

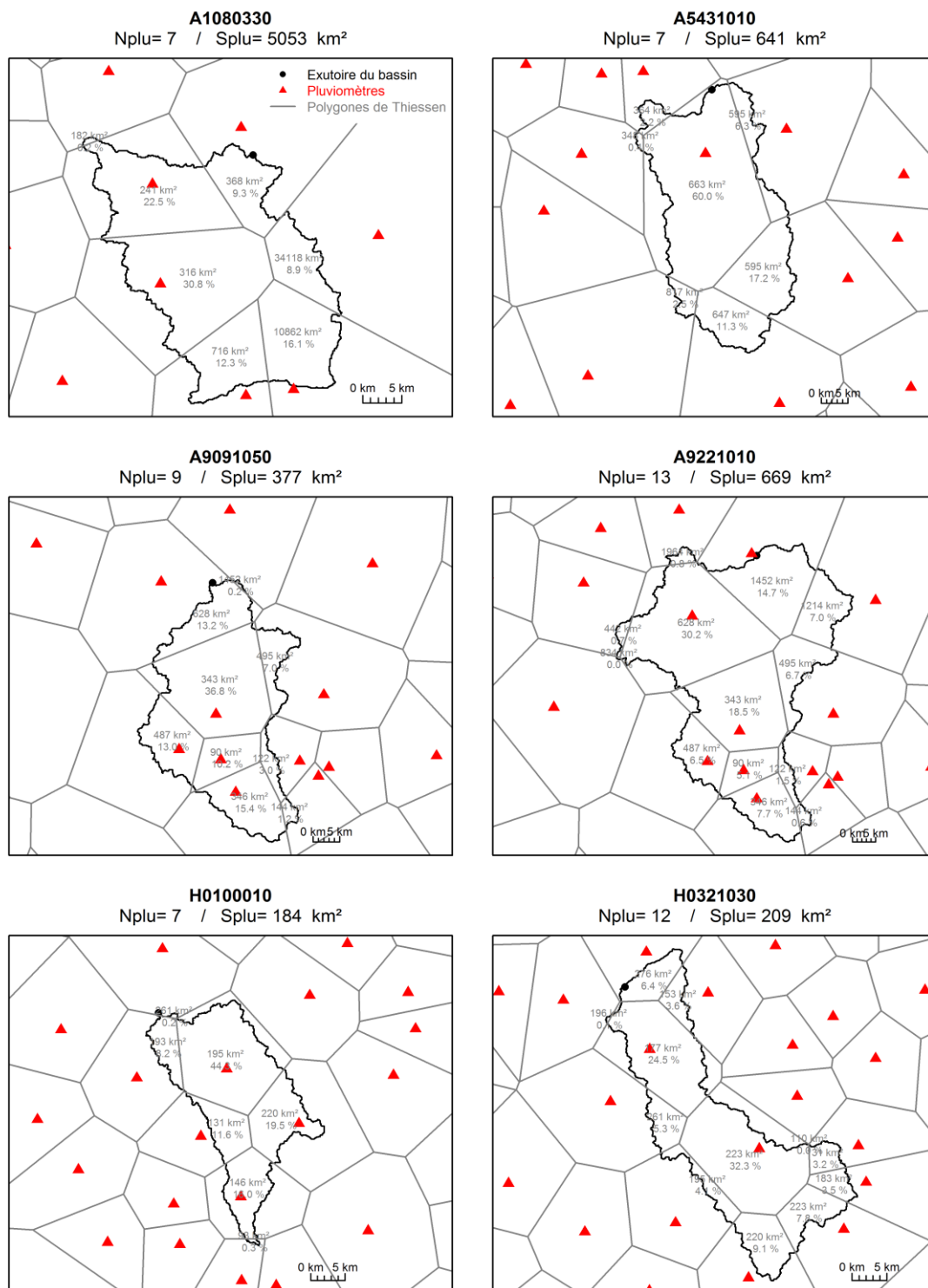


Figure 89. Réseau de mesure pluviométrique de 6 bassins versants. En gris, le contour des polygones de Thiessen. La surface des polygones de Thiessen et le pourcentage de la surface du polygone de Thiessen de chaque pluviomètre comprise dans le bassin versant sont indiqués au centre de chaque polygone intersectant le bassin versant. N_{PLU} et S_{PLU} indiquent respectivement le nombre de pluviomètres à proximité du bassin et la surface pluviométrique de chaque bassin.

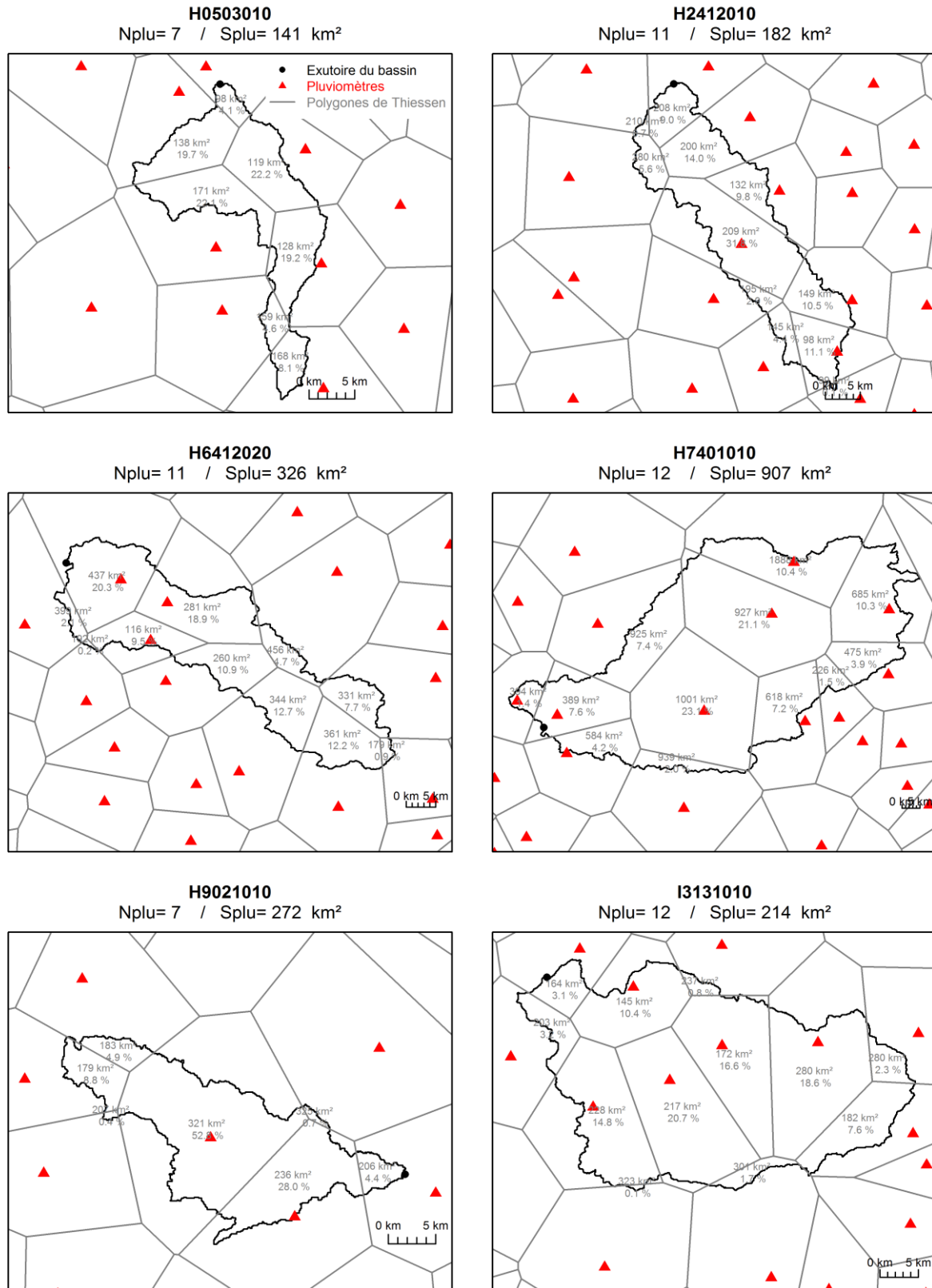


Figure 90. Réseau de mesure pluviométrique de 6 bassins versants. En gris, le contour des polygones de Thiessen. La surface des polygones de Thiessen et le pourcentage de la surface du polygone de Thiessen de chaque pluviomètre comprise dans le bassin versant sont indiqués au centre de chaque polygone intersectant le bassin versant. N_{PLU} et S_{PLU} indiquent respectivement le nombre de pluviomètres à proximité du bassin et la surface pluviométrique de chaque bassin.

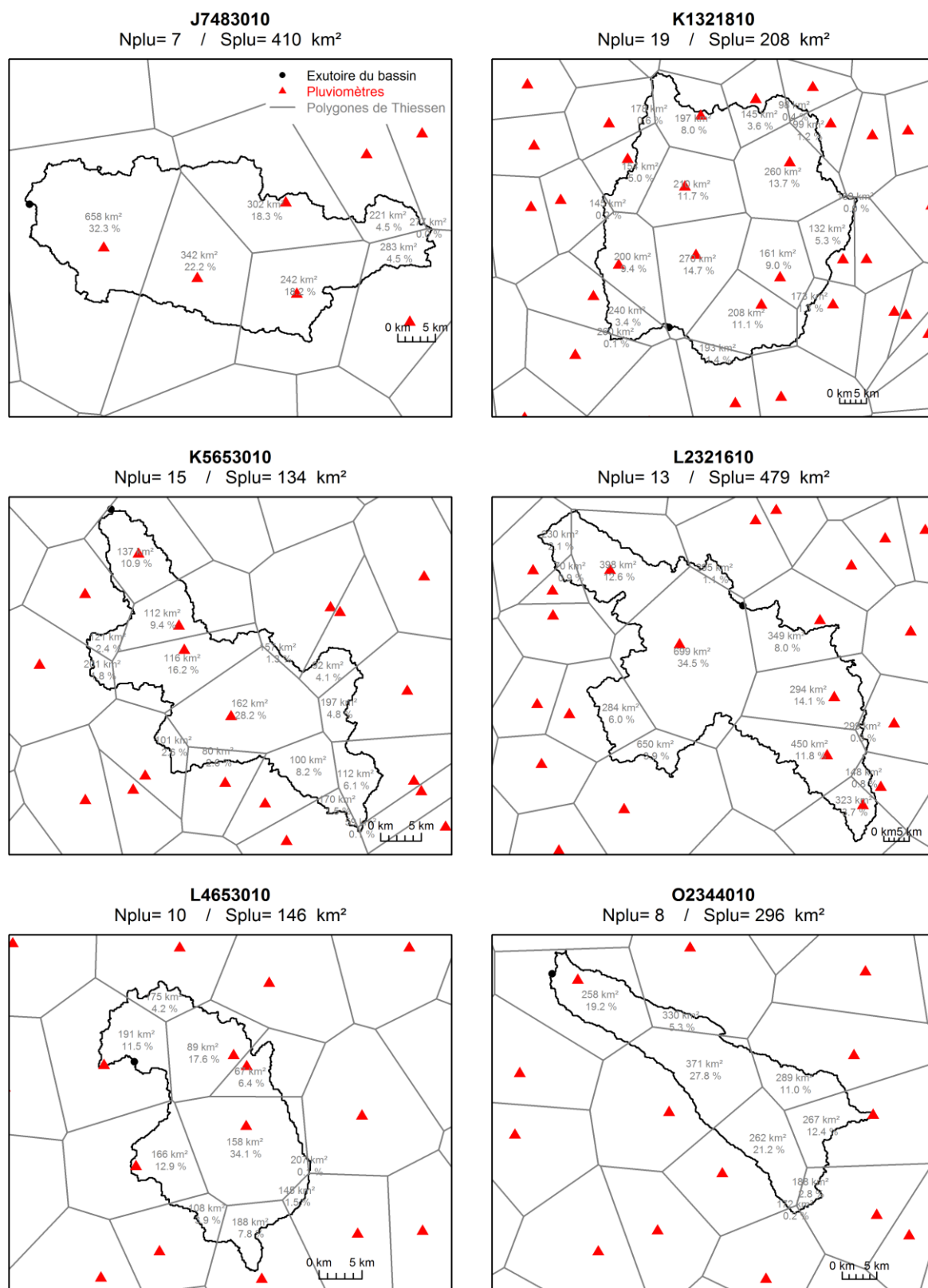


Figure 91. Réseau de mesure pluviométrique de 6 bassins versants. En gris, le contour des polygones de Thiessen. La surface des polygones de Thiessen et le pourcentage de la surface du polygone de Thiessen de chaque pluviomètre comprise dans le bassin versant sont indiqués au centre de chaque polygone intersectant le bassin versant. N_{PLU} et S_{PLU} indiquent respectivement le nombre de pluviomètres à proximité du bassin et la surface pluviométrique de chaque bassin.

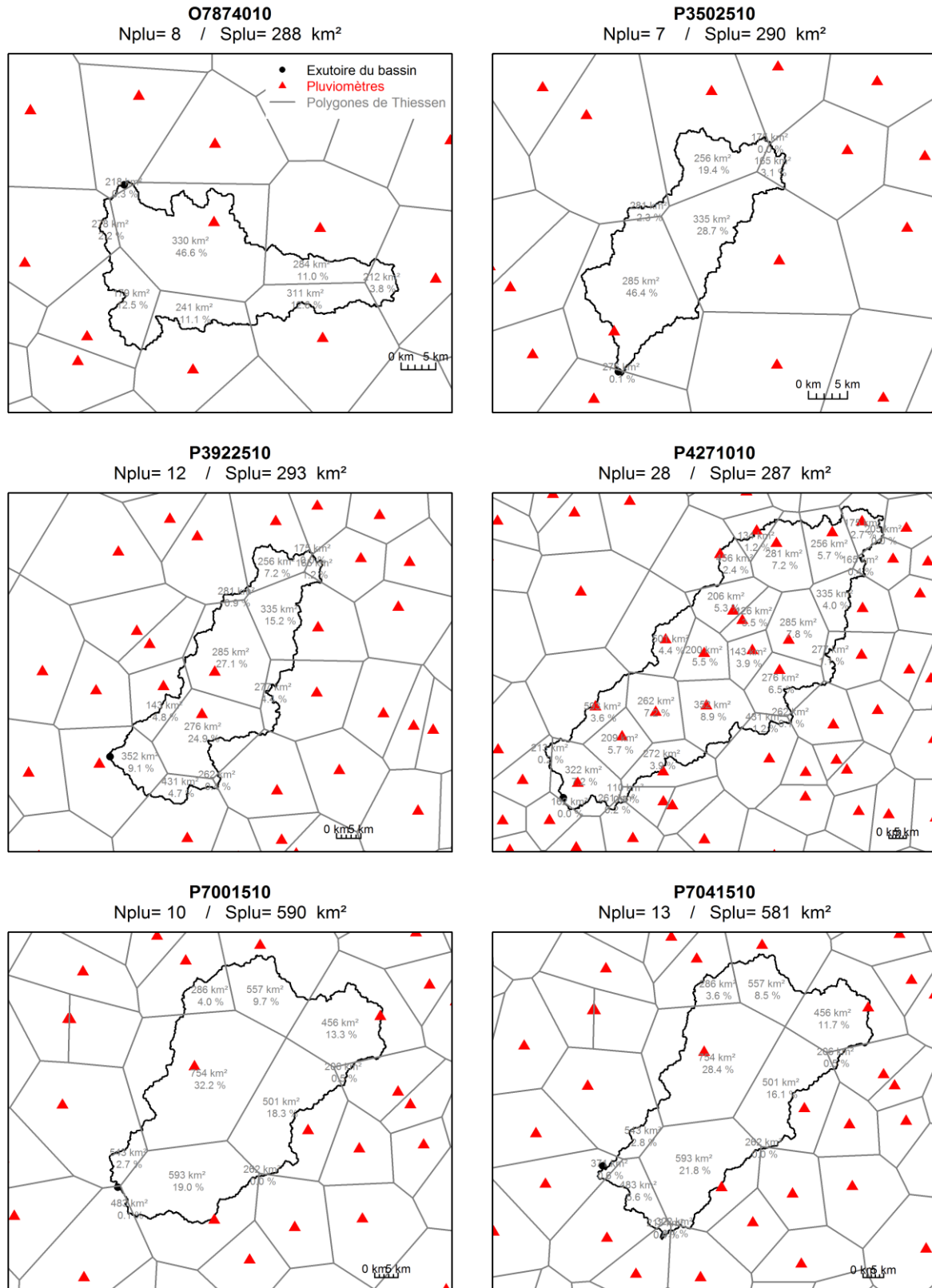


Figure 92. Réseau de mesure pluviométrique de 6 bassins versants. En gris, le contour des polygones de Thiessen. La surface des polygones de Thiessen et le pourcentage de la surface du polygone de Thiessen de chaque pluviomètre comprise dans le bassin versant sont indiqués au centre de chaque polygone intersectant le bassin versant. N_{PLU} et S_{PLU} indiquent respectivement le nombre de pluviomètres à proximité du bassin et la surface pluviométrique de chaque bassin.

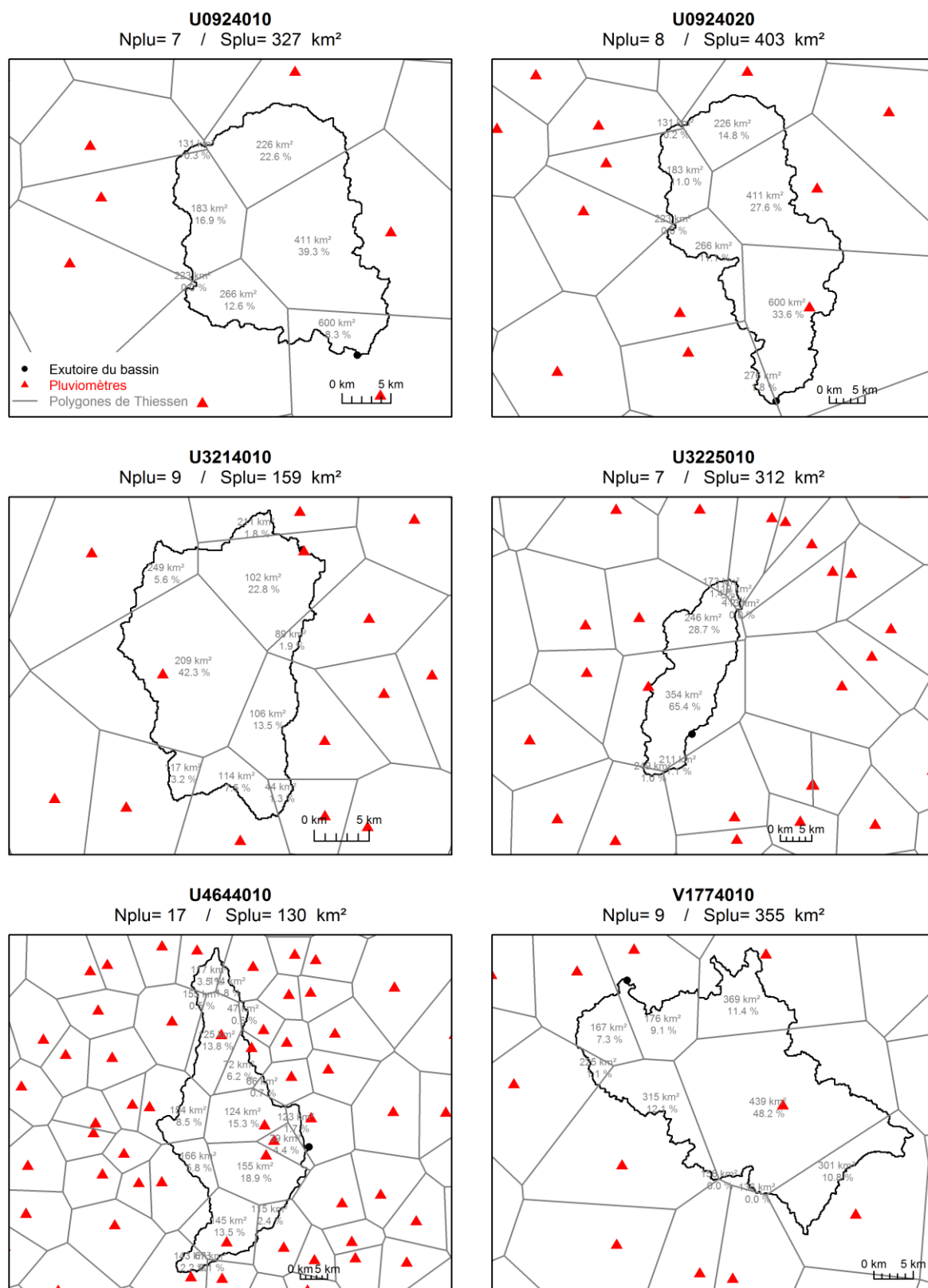
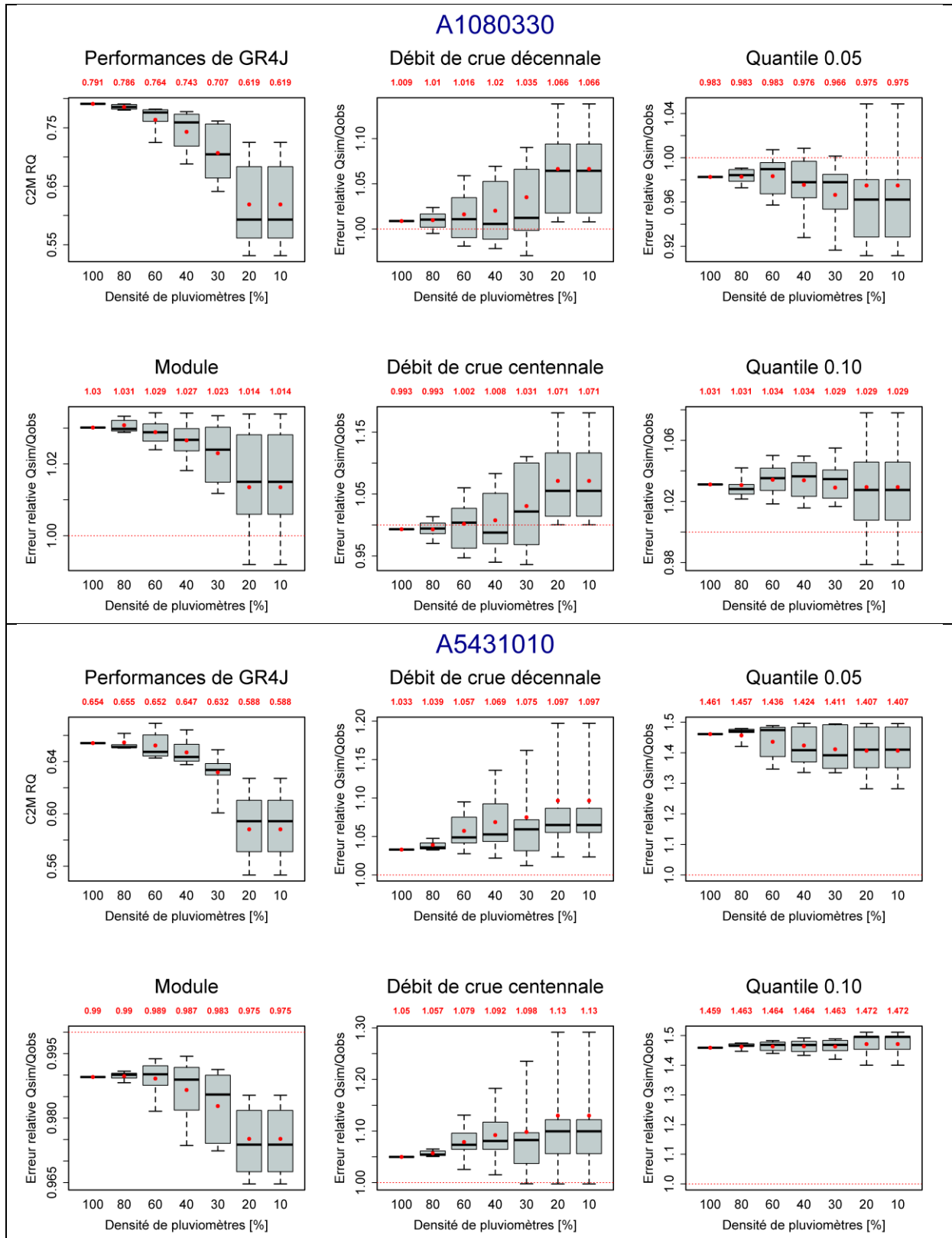
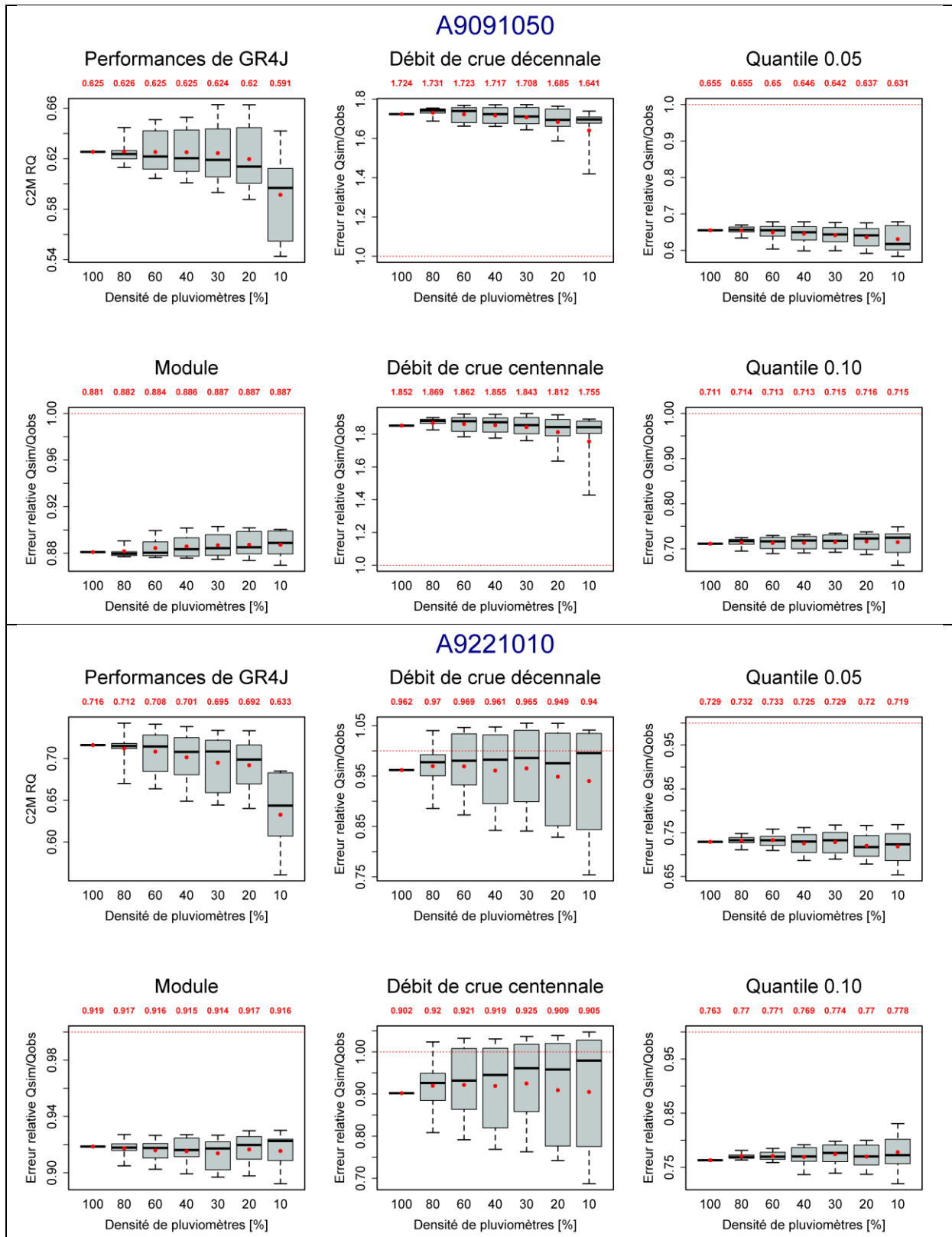
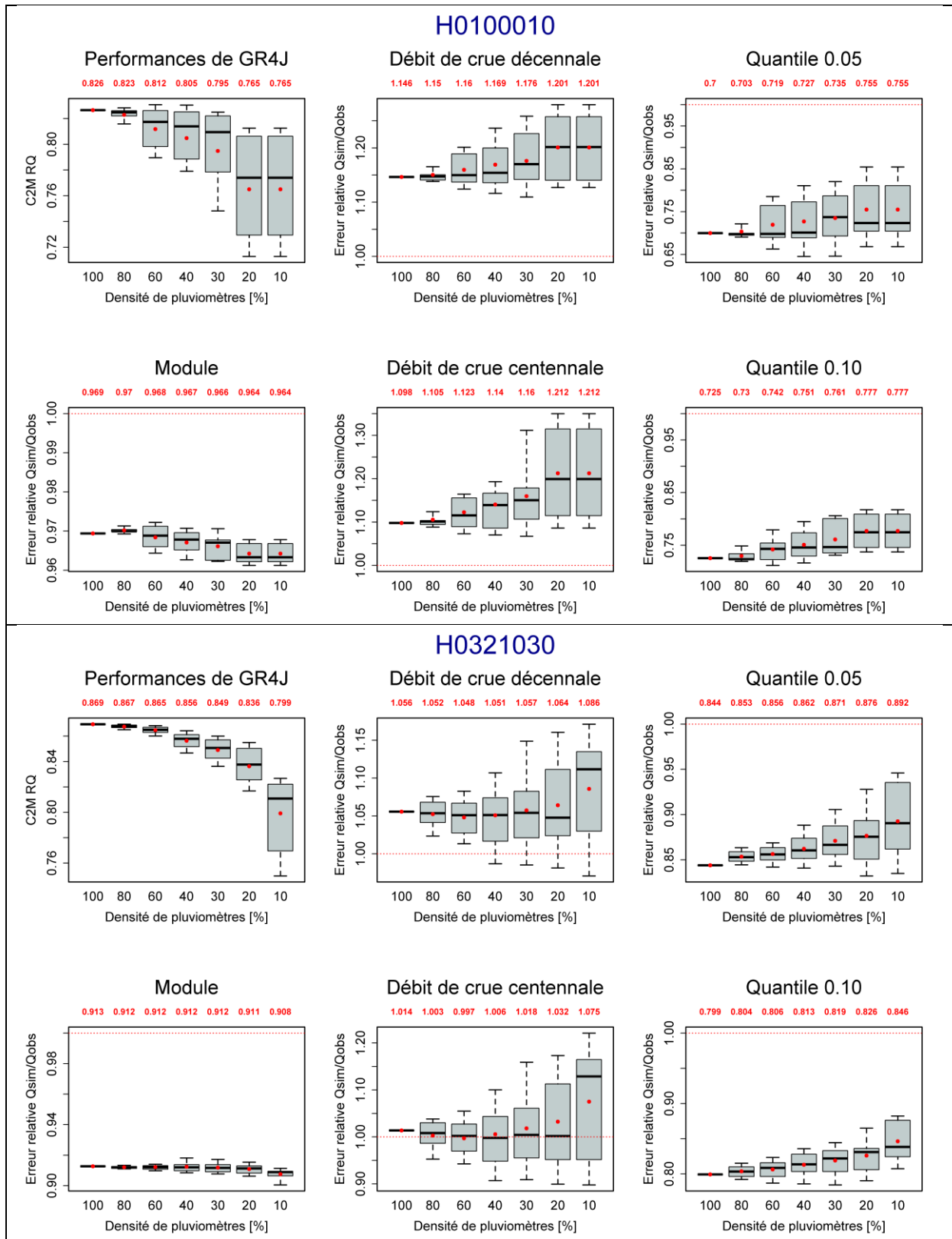


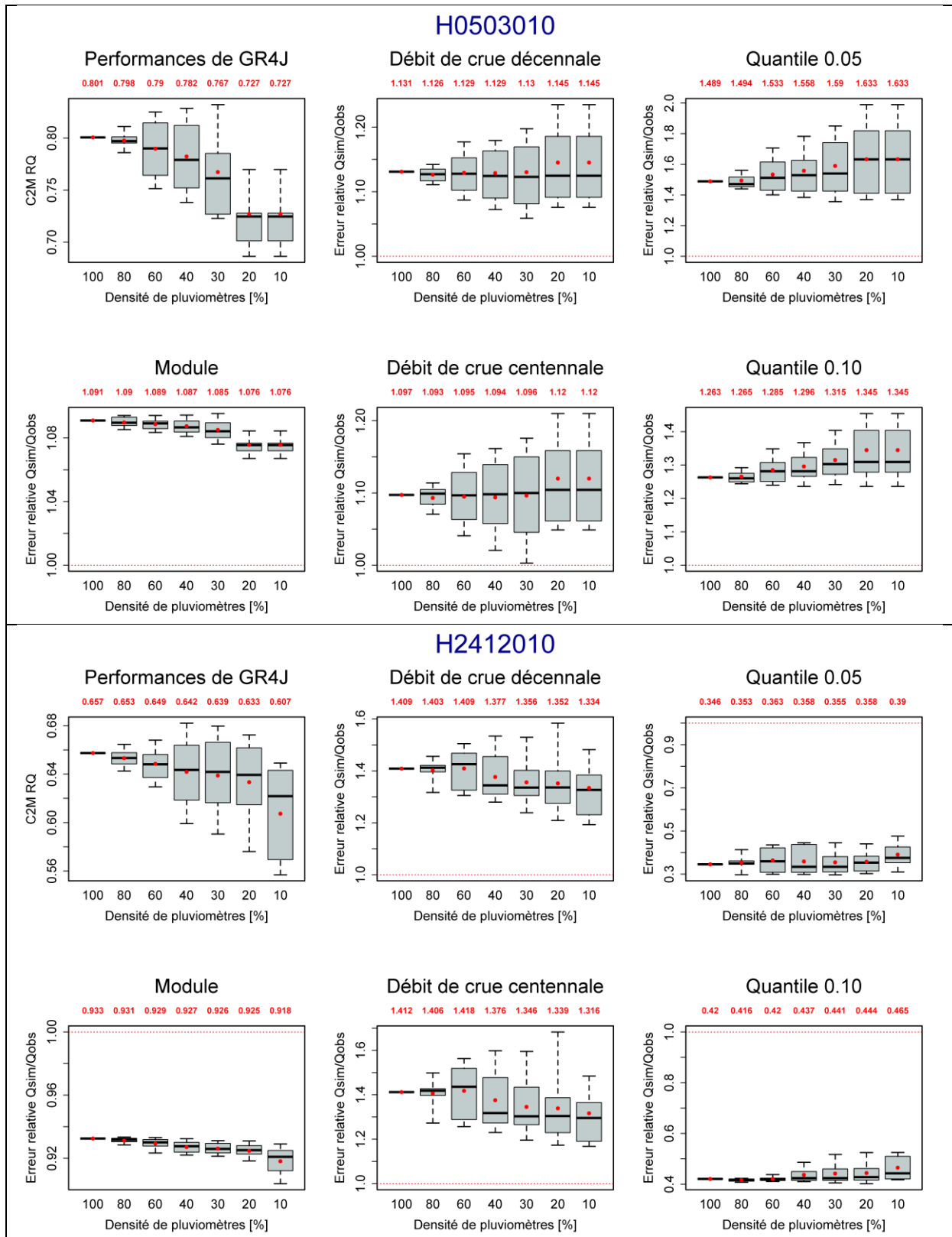
Figure 93. Réseau de mesure pluviométrique de 6 bassins versants. En gris, le contour des polygones de Thiessen. La surface des polygones de Thiessen et le pourcentage de la surface du polygone de Thiessen de chaque pluviomètre comprise dans le bassin versant sont indiqués au centre de chaque polygone intersectant le bassin versant. N_{PLU} et S_{PLU} indiquent respectivement le nombre de pluviomètres à proximité du bassin et la surface pluviométrique de chaque bassin.

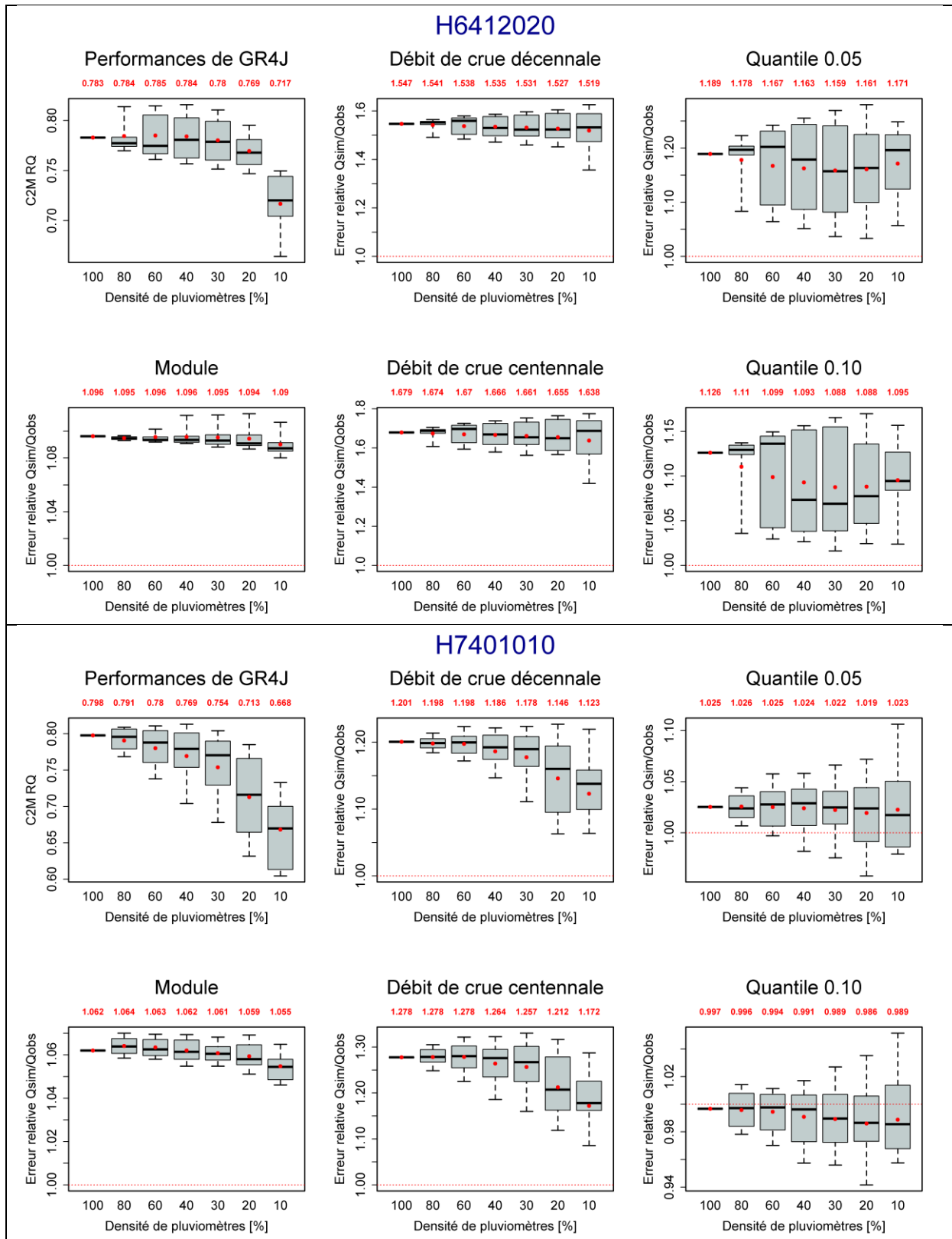
Annexe E. Impact de la densité du réseau pluviométrique sur les calculs hydrologiques par bassin versant

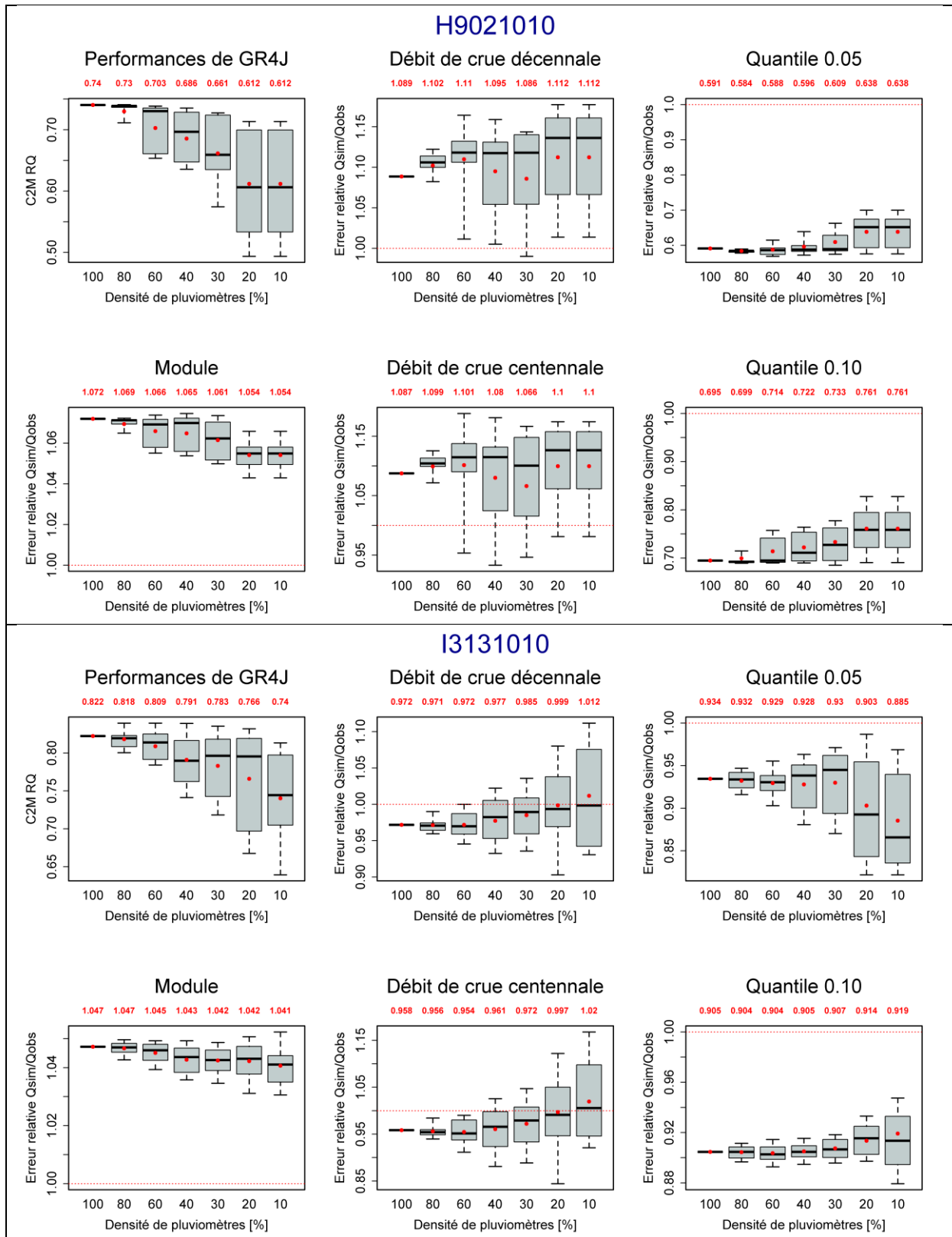


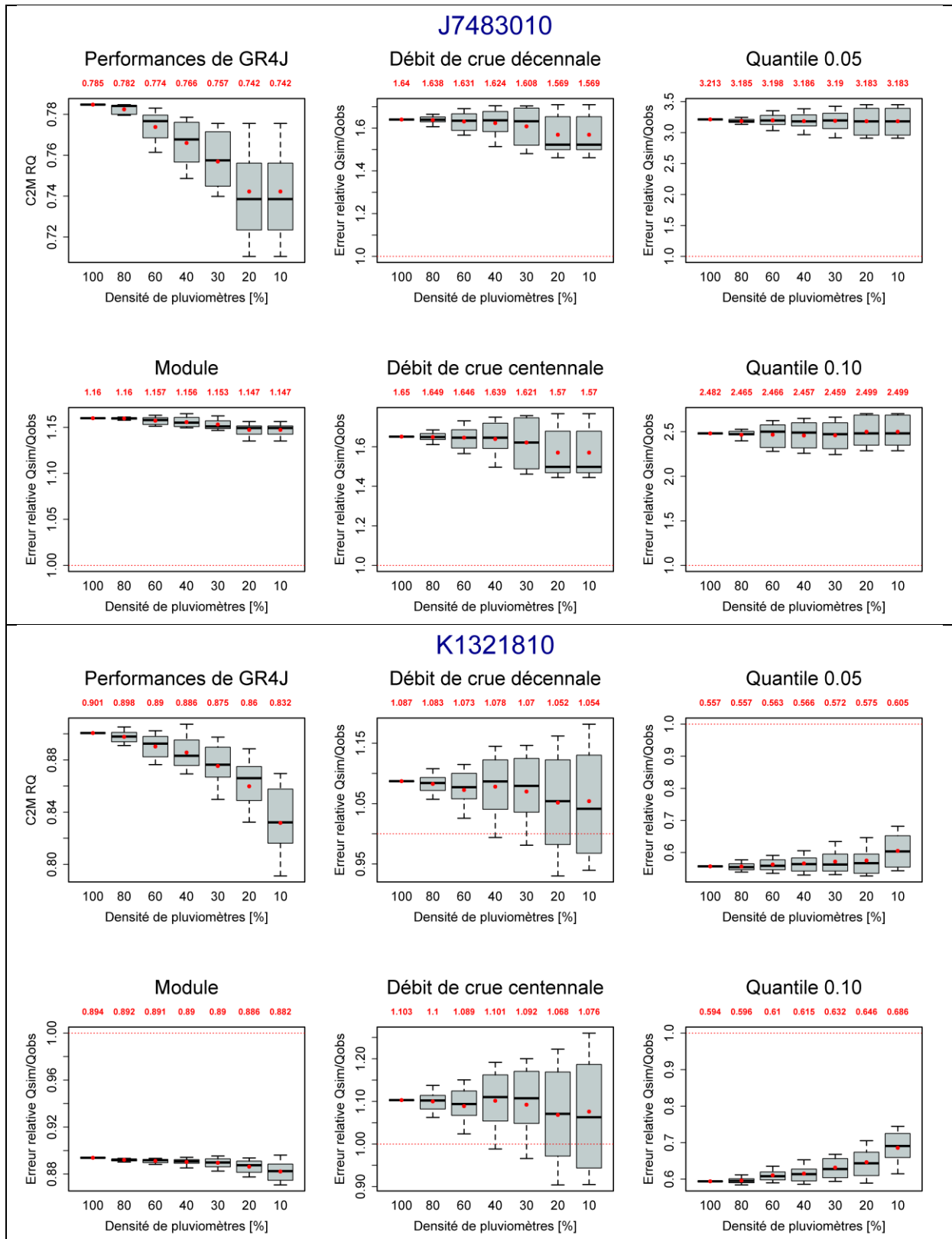


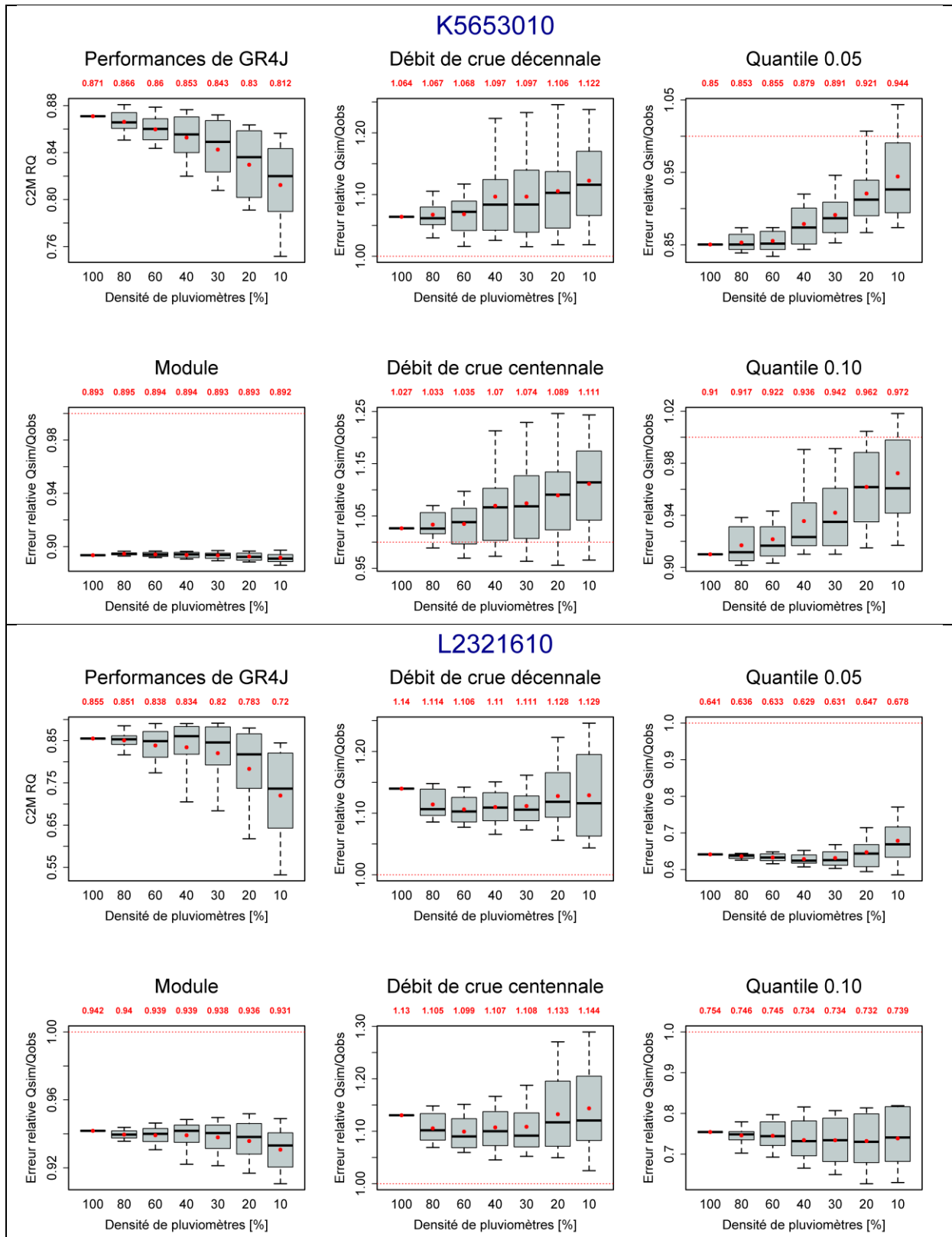


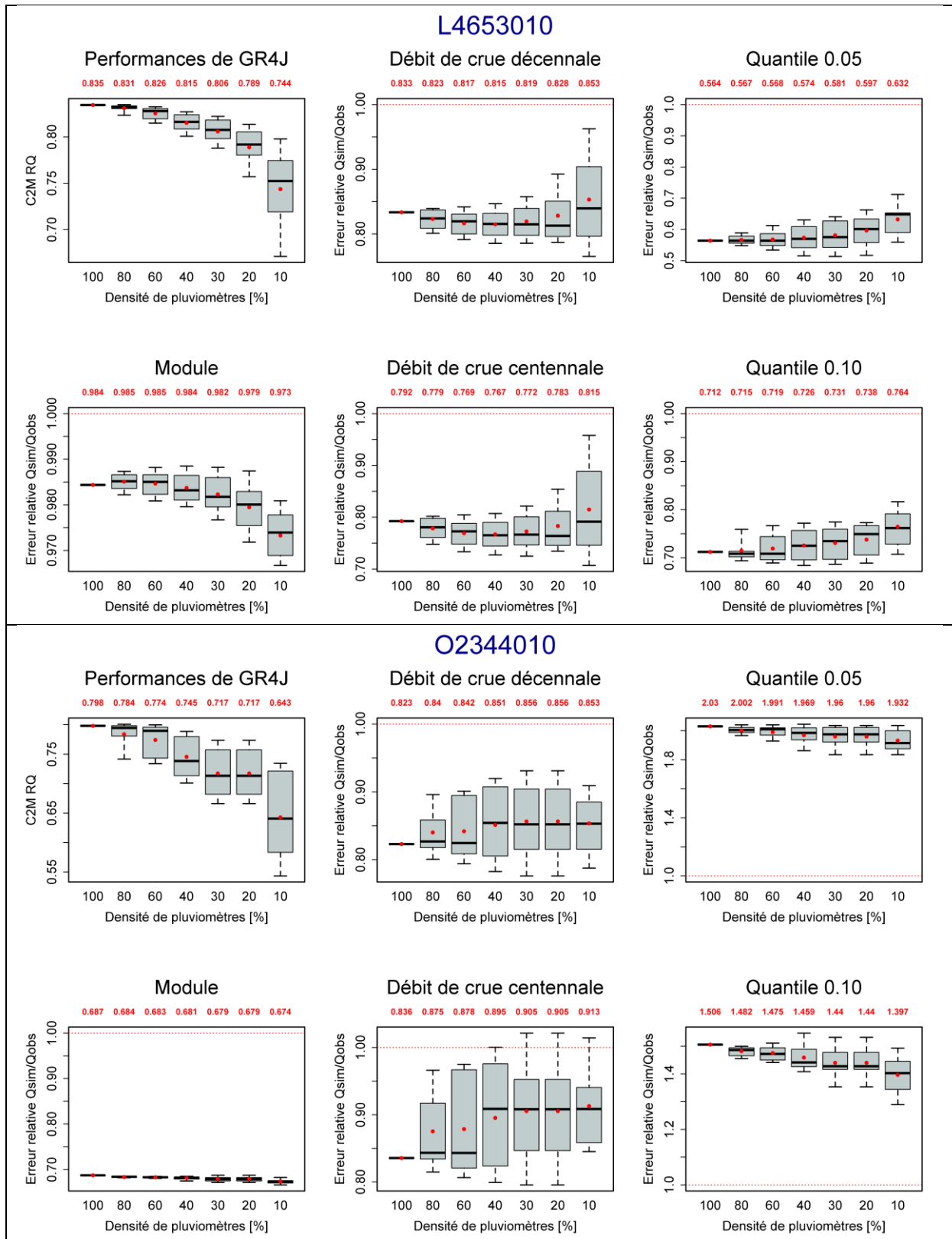


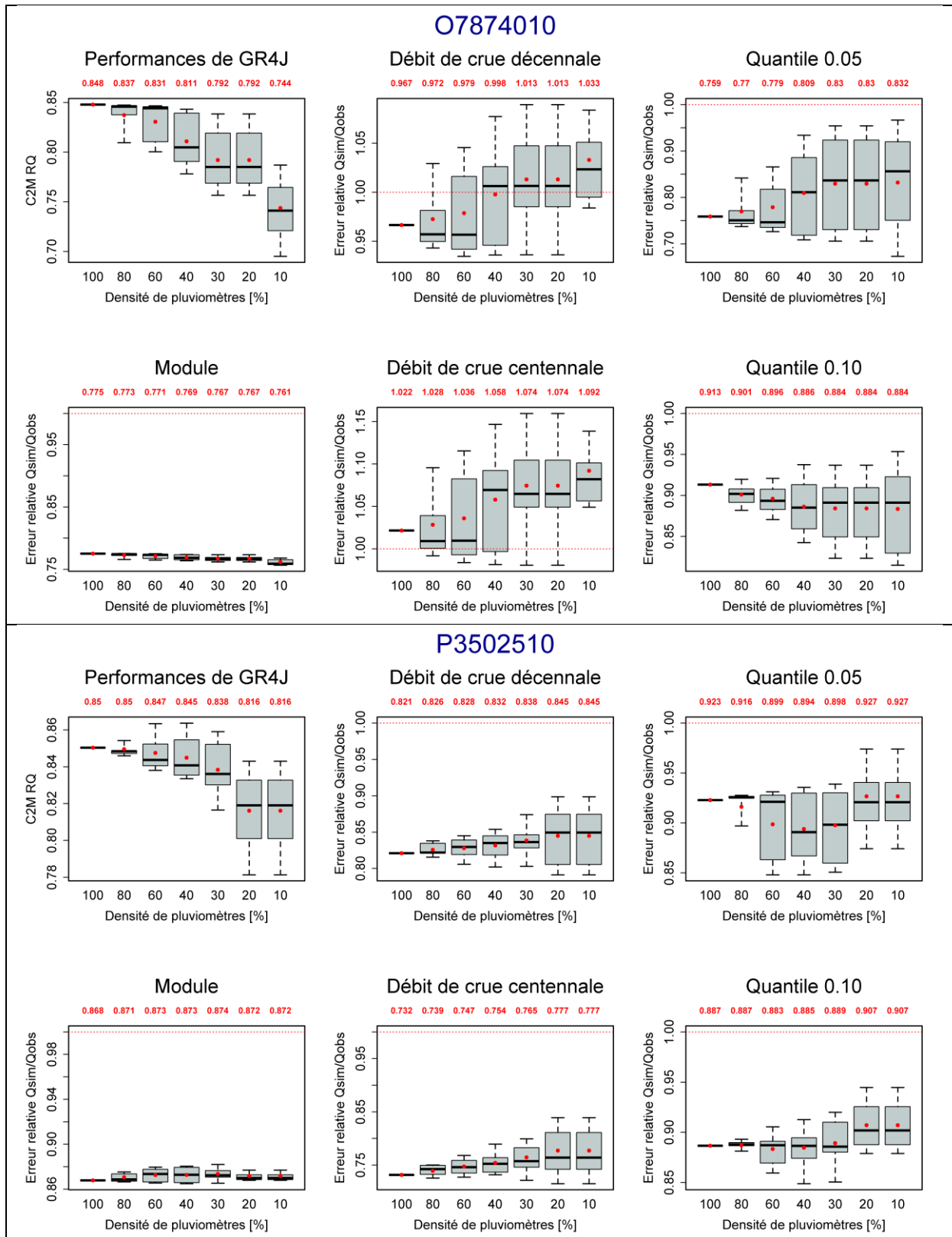


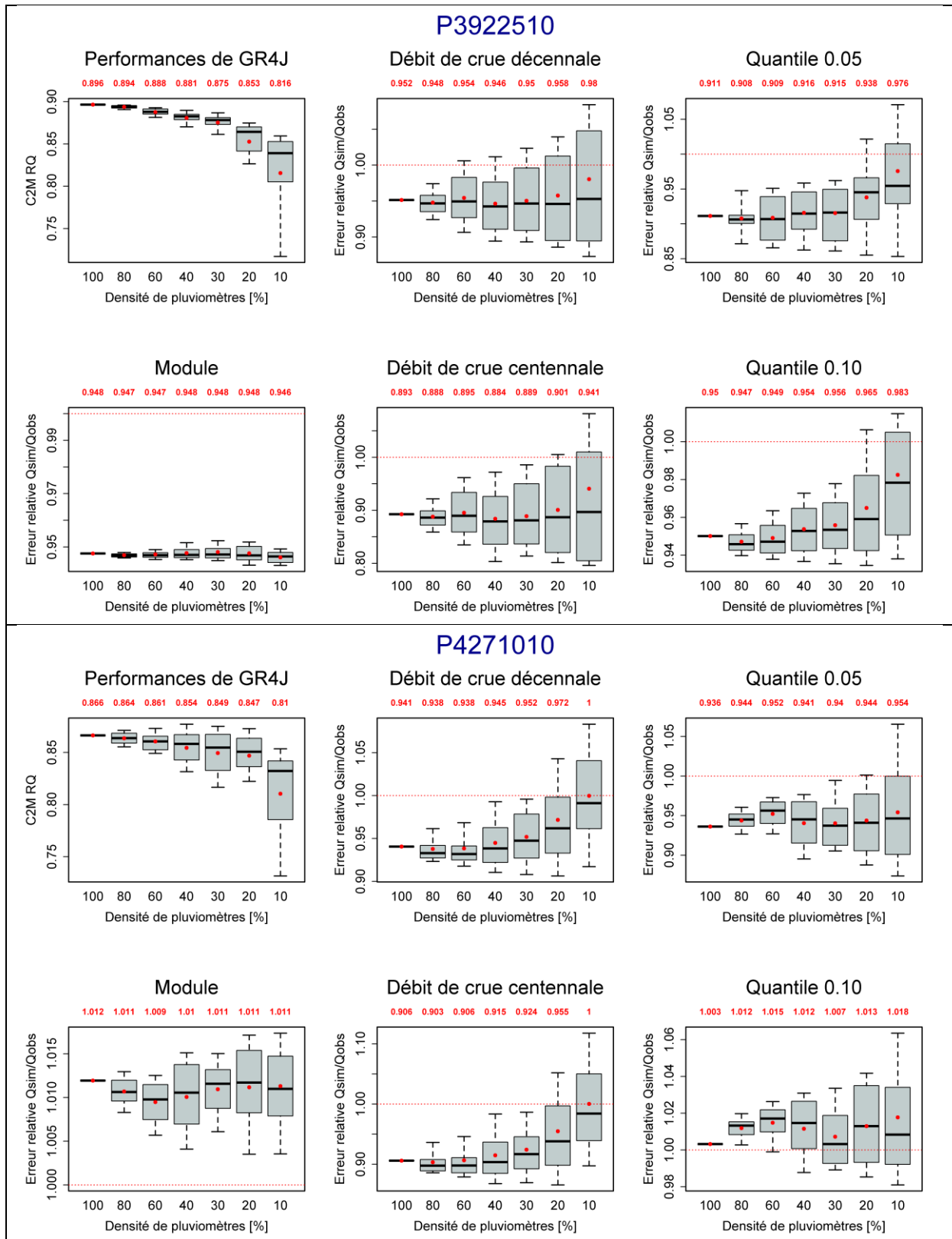


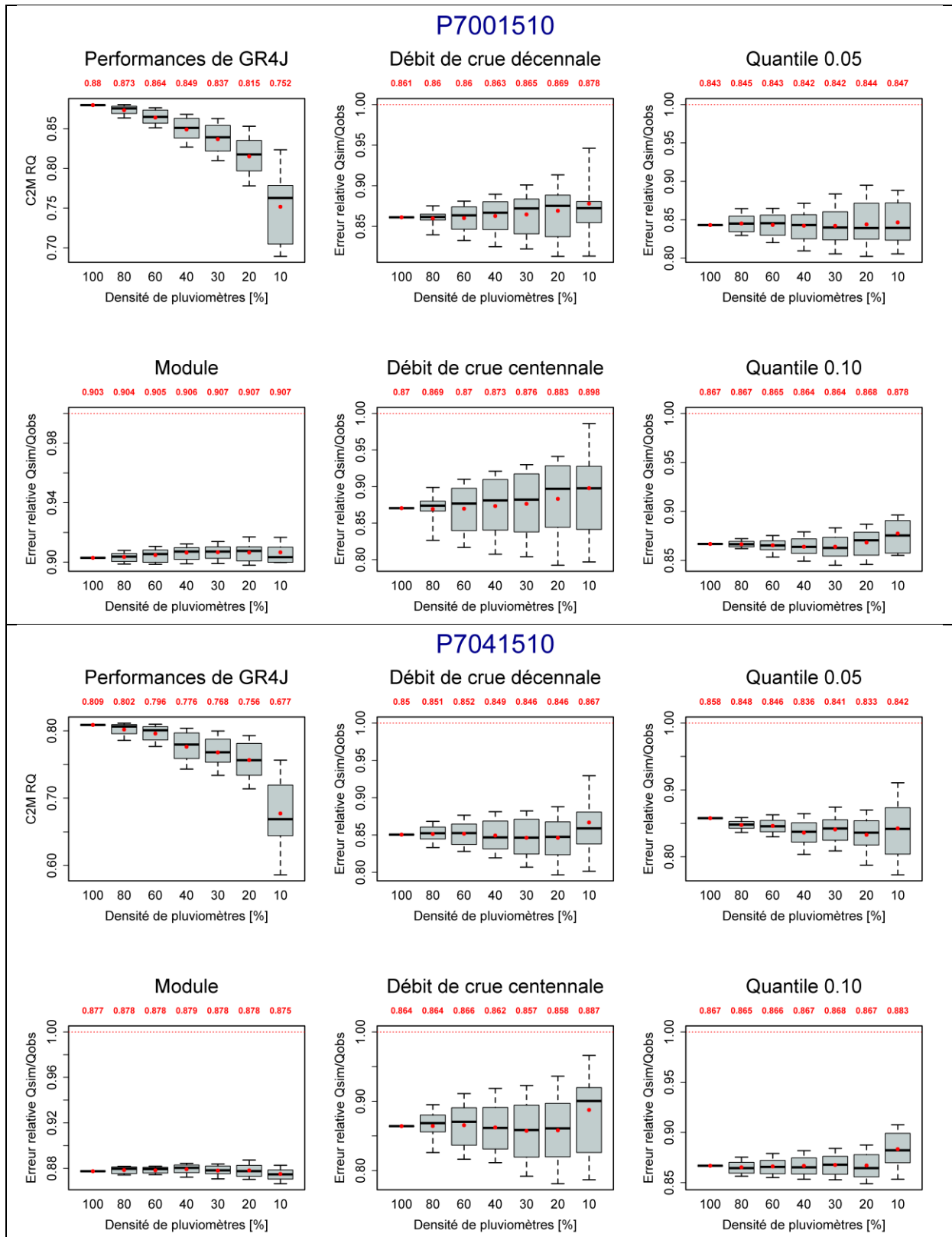


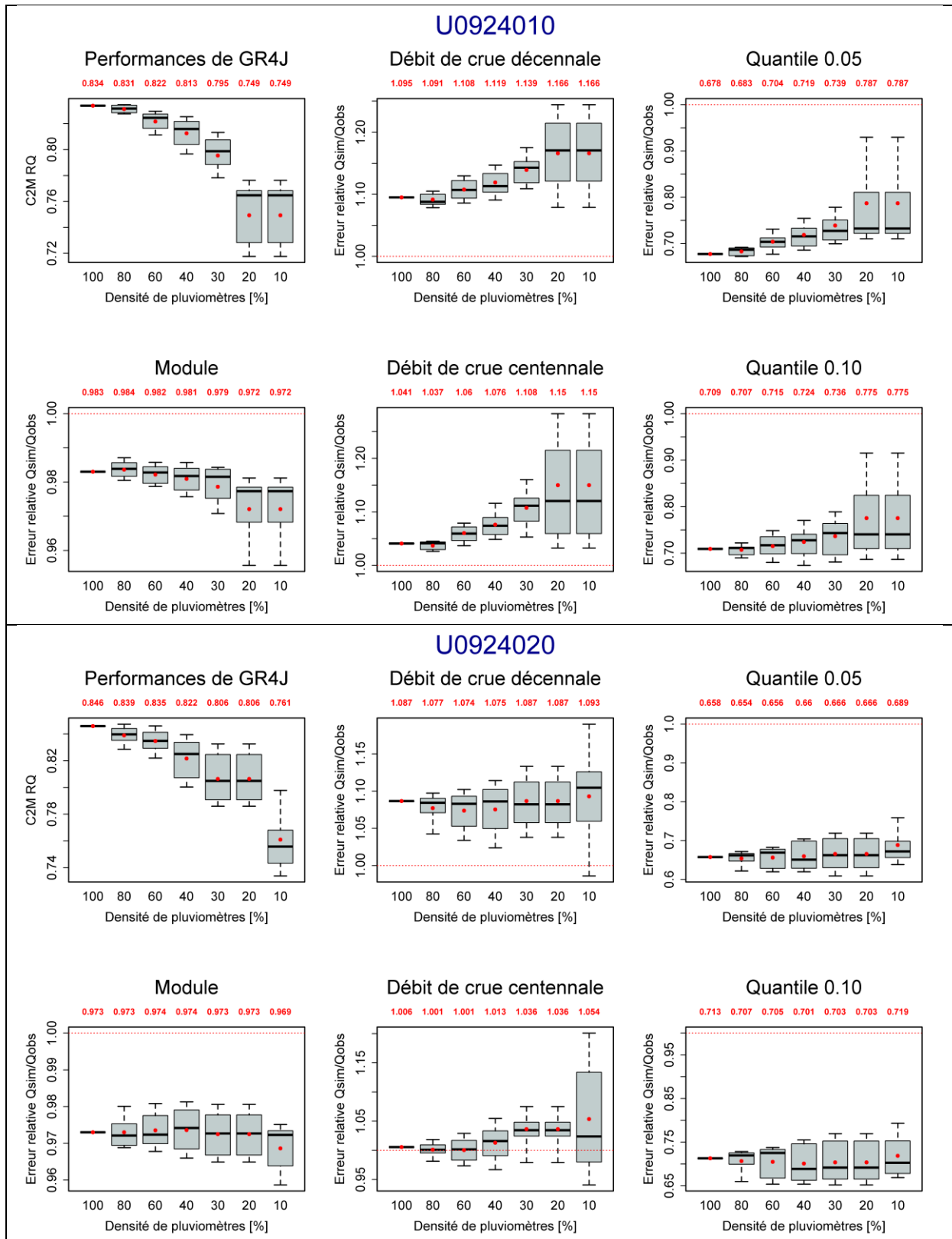


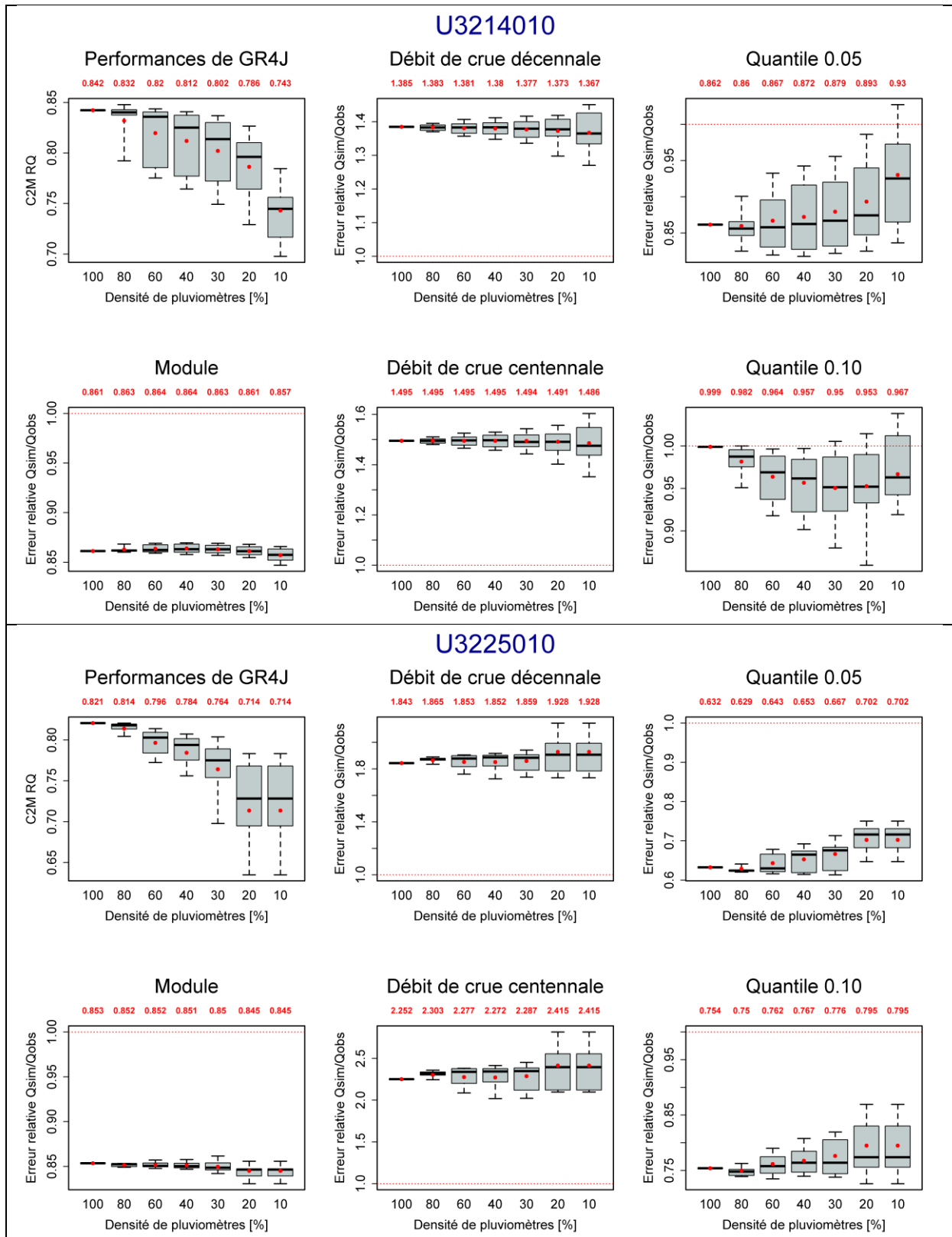












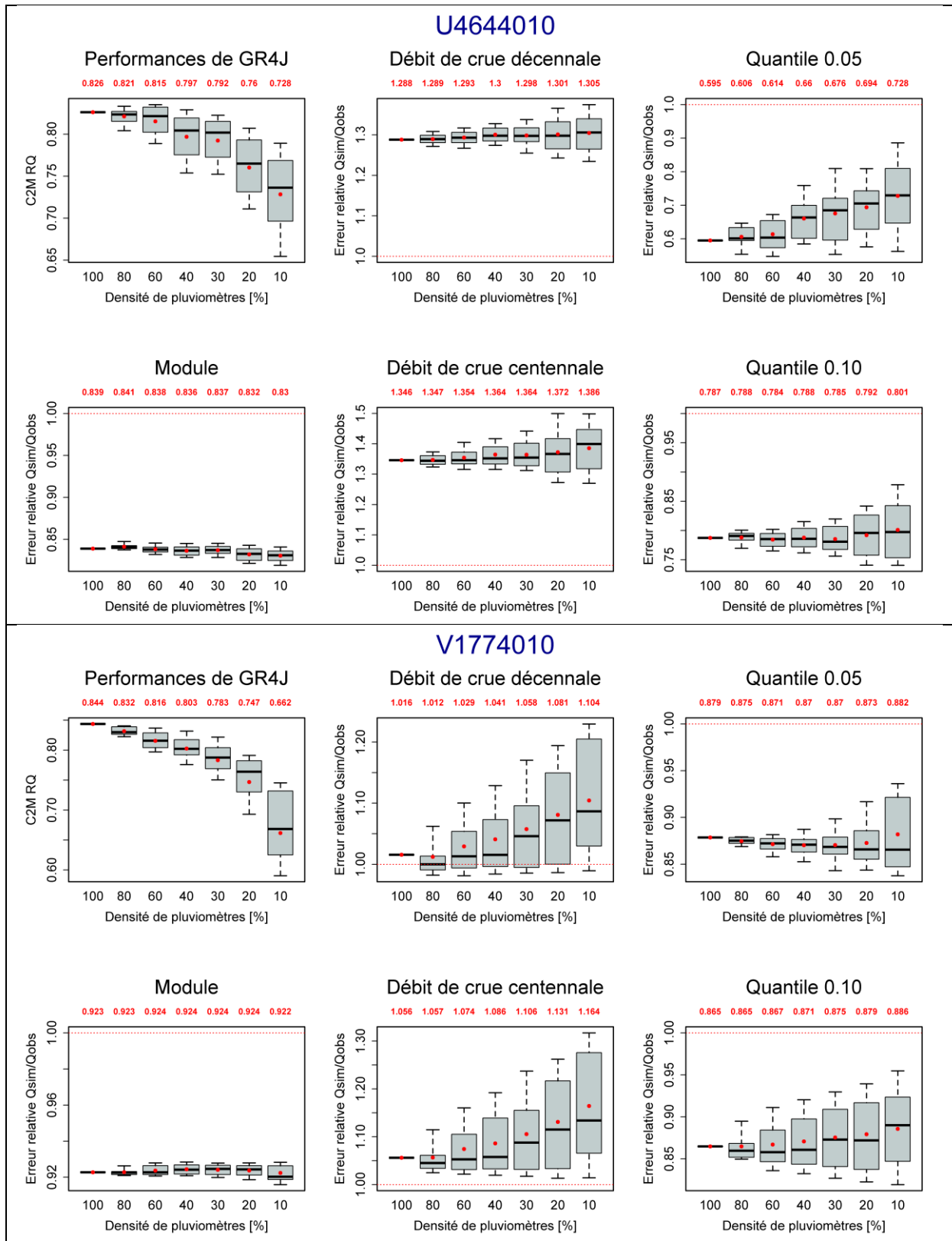
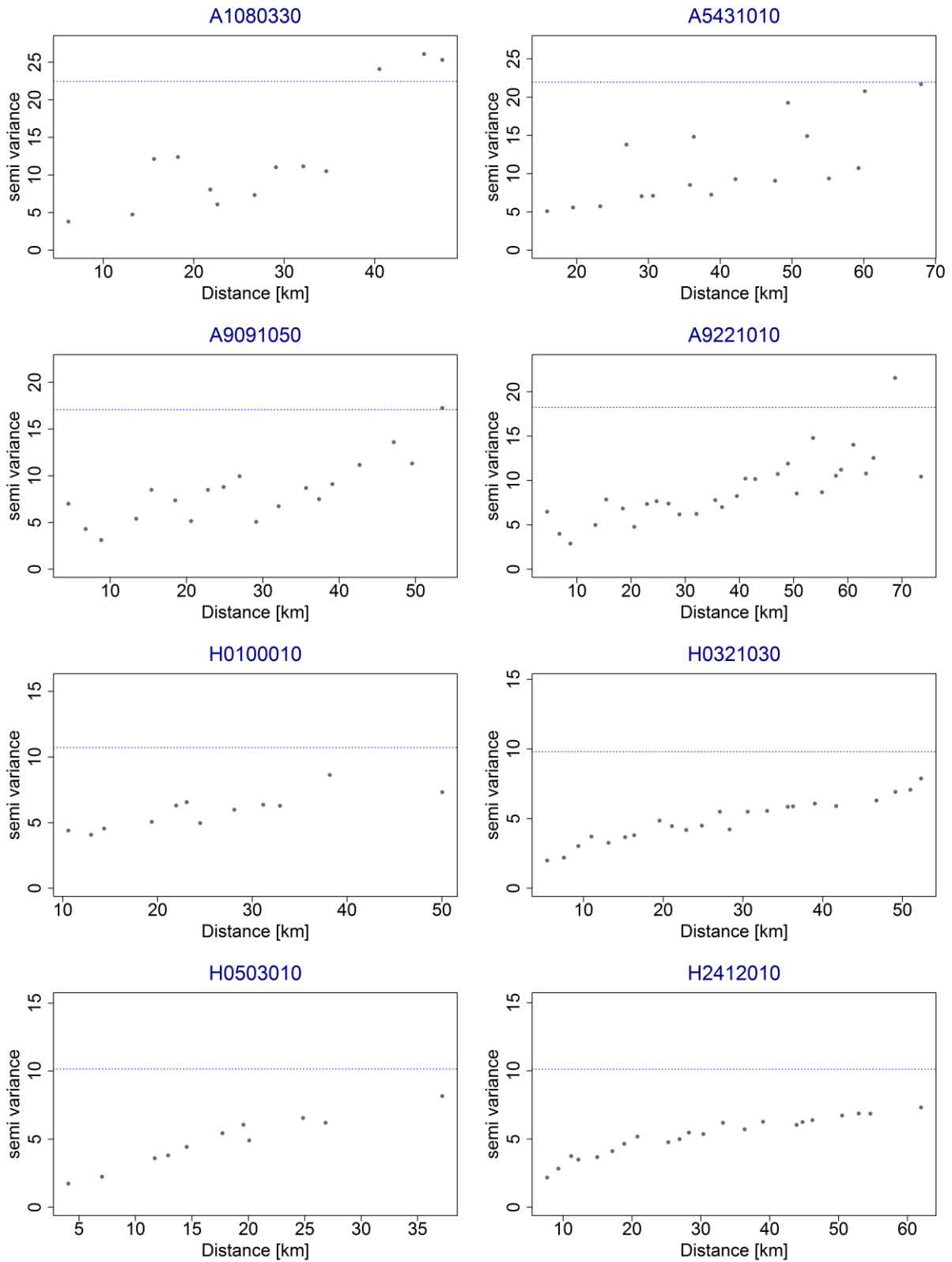
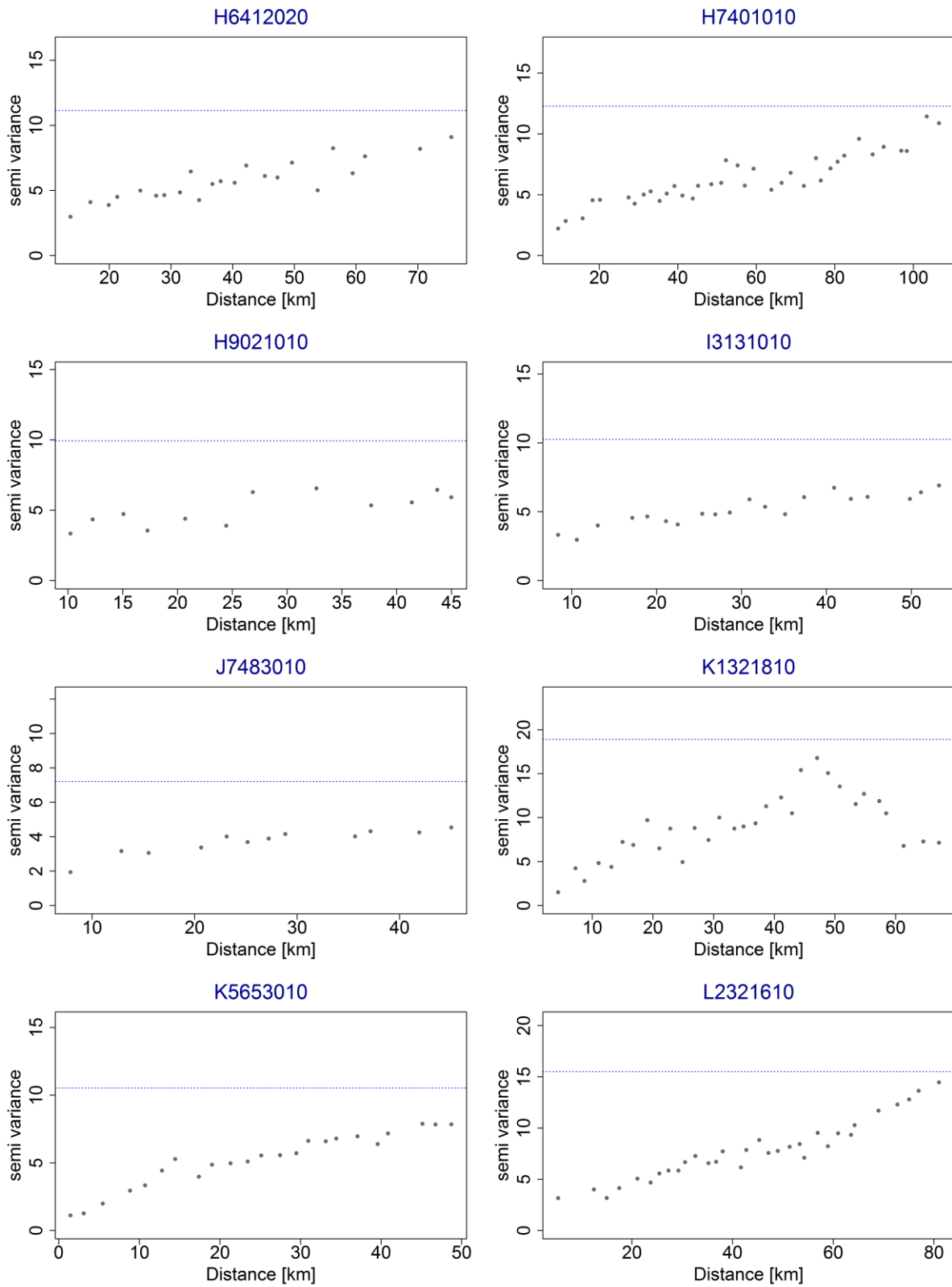
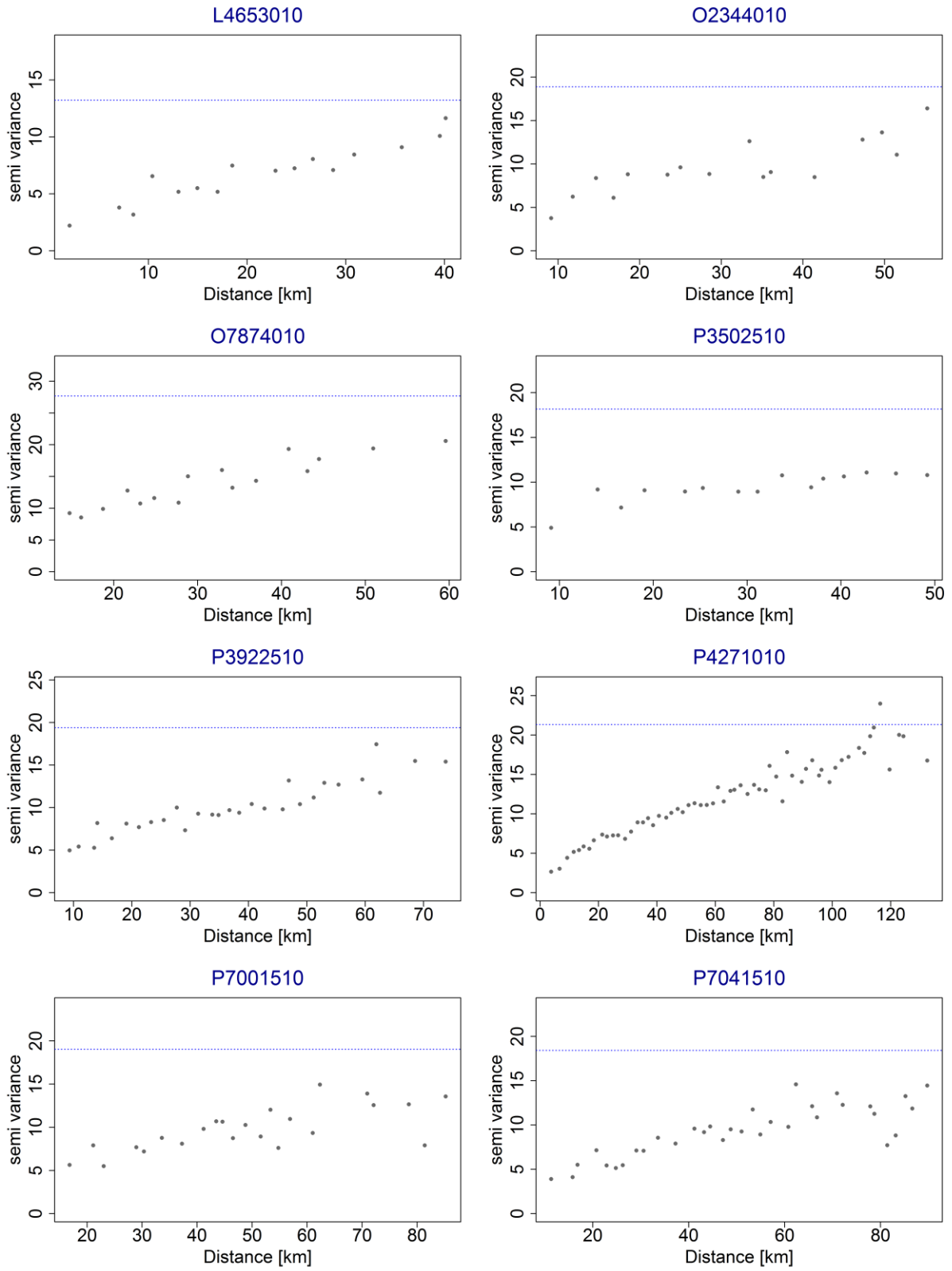


Figure 94. Performances des divers calculs hydrologiques considérés en fonction de la densité spatiale du réseau pluviométrique pour les 30 bassins versants.

Annexe F. Variogramme climatologique de chacun des 30 bassins versants de l'étude







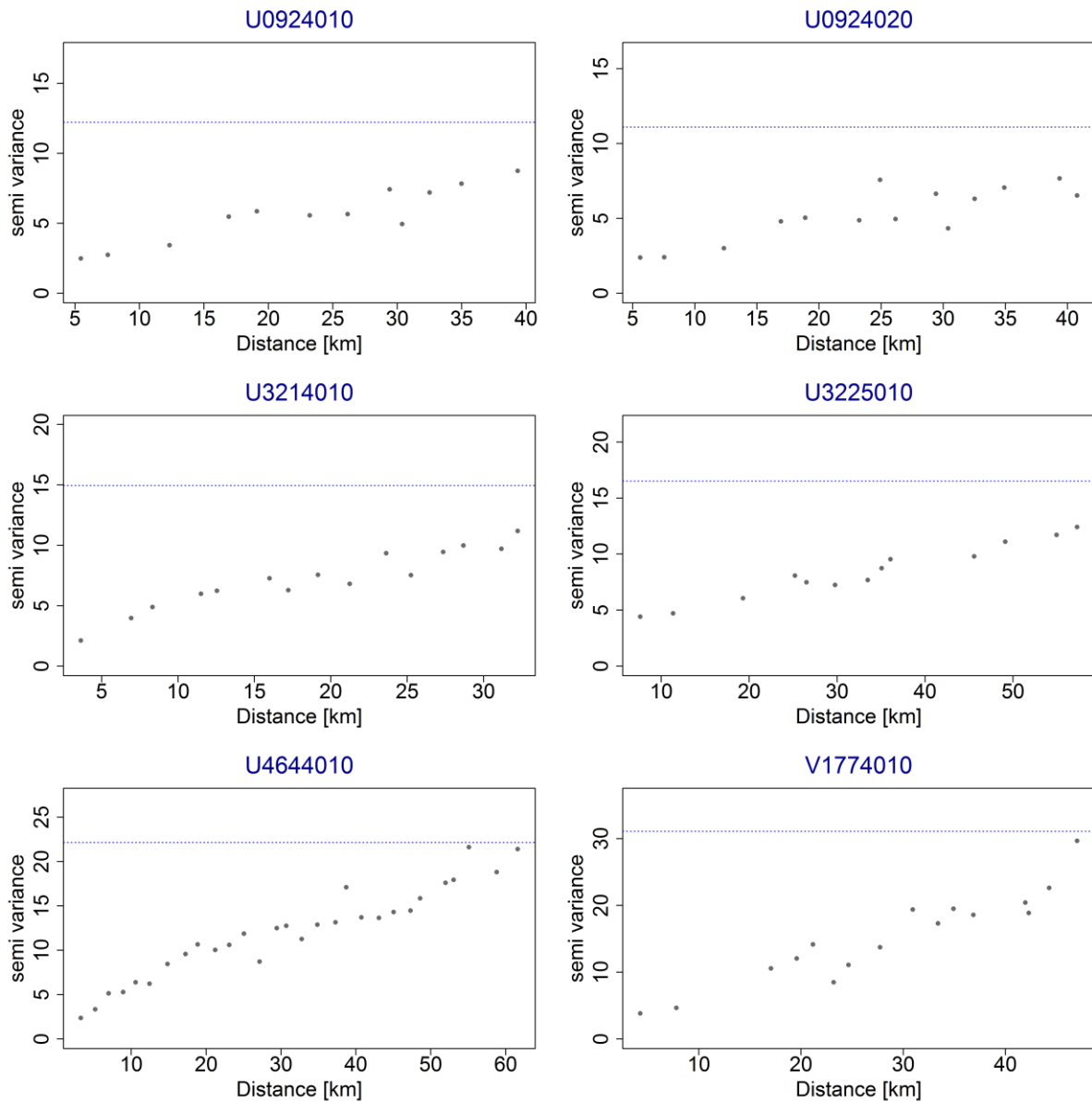


Figure 95. Variogramme climatologique de chacun des 30 bassins versants de l'étude. La ligne en pointillé bleu représente la variance moyenne de la variable aléatoire.

