



HAL
open science

Représentation de l'urbanisation dans la modélisation hydrologique à l'échelle du bassin versant

Mohamed Saadi

► **To cite this version:**

Mohamed Saadi. Représentation de l'urbanisation dans la modélisation hydrologique à l'échelle du bassin versant. Hydrologie. Sorbonne Université, 2020. Français. NNT : 2020SORUS374 . tel-03250292v2

HAL Id: tel-03250292

<https://theses.hal.science/tel-03250292v2>

Submitted on 29 Oct 2021

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

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

Doctorat Sorbonne Université

École Doctorale n° 398 : Géosciences, Ressources Naturelles et Environnement

THÈSE

pour obtenir le grade de docteur délivré par

Sorbonne Université

Spécialité : Hydrologie

présentée et soutenue publiquement par

Mohamed SAADI

le 10 Novembre 2020

Représentation de l'urbanisation dans la modélisation hydrologique à l'échelle du bassin versant

Sous la direction de **Ludovic OUDIN** et **Pierre RIBSTEIN**

Jury

Mme Isabelle BRAUD	INRAE, Lyon, France	Rapporteur
M. Éric GAUME	Université Gustave Eiffel, Nantes, France	Rapporteur
Mme Agnès DUCHARNE	CNRS-UMR METIS, Paris, France	Examinatrice
Mme Anne-Catherine FAVRE	Grenoble INP, Grenoble, France	Examinatrice
M. Markus HRACHOWITZ	Technische Universiteit Delft, Delft, Pays-Bas	Examinateur

Sorbonne Université

UMR Milieux Environnementaux, Transferts et Interactions dans les Hydrosystèmes et les Sols (METIS)

Case 105 – Tour 46/56 – 3^{ème} étage – 4 place Jussieu – 75252 Paris Cedex 05 – France

Doctorat Sorbonne Université

École Doctorale n° 398 : Géosciences, Ressources Naturelles et Environnement

THÈSE

pour obtenir le grade de docteur délivré par

Sorbonne Université

Spécialité : Hydrologie

présentée et soutenue publiquement par

Mohamed SAADI

le 10 Novembre 2020

Représentation de l'urbanisation dans la modélisation hydrologique à l'échelle du bassin versant

Sous la direction de **Ludovic OUDIN** et **Pierre RIBSTEIN**

Jury

Mme Isabelle BRAUD	INRAE, Lyon, France	Rapporteur
M. Éric GAUME	Université Gustave Eiffel, Nantes, France	Rapporteur
Mme Agnès DUCHARNE	CNRS-UMR METIS, Paris, France	Examinatrice
Mme Anne-Catherine FAVRE	Grenoble INP, Grenoble, France	Examinatrice
M. Markus HRACHOWITZ	Technische Universiteit Delft, Delft, Pays-Bas	Examineur

REMERCIEMENTS

J'éprouve une immense gratitude envers mes deux directeurs de thèse, Ludovic Oudin et Pierre Ribstein, car la réussite de ce travail tient énormément à leur encadrement exemplaire. Grâce à la confiance qu'ils m'ont insufflée, à leur disponibilité, à leur patience et aux innombrables conseils qu'ils ont sus me prodiguer avec beaucoup d'humilité, j'ai pu lutter contre les angoisses récurrentes que m'inspire l'inexactitude de la Science Hydrologique. Je peux écrire tout un livre sur nos réunions passionnantes durant lesquelles nous nous évertuions à linéariser mon travail de recherche et à me tenir le plus possible sur un chemin droit.

Je sais gré à tous les honorables membres du jury de leurs critiques constructives et leurs remarques pertinentes qui m'ont permis de prendre du recul par rapport à l'approche que j'ai adoptée. Je tiens à remercier sincèrement Agnès Ducharne, en sa qualité de présidente, Isabelle Braud et Éric Gaume d'avoir attentivement lu et soigneusement rapporté mon manuscrit de thèse, et Anne-Catherine Favre et Markus Hrachowitz d'avoir pertinemment examiné et discuté mes travaux.

La trajectoire qu'ont prise les travaux de ma thèse et la qualité du manuscrit tiennent également aux suggestions très appréciées des membres du comité de suivi, que ce soit durant ou en dehors de nos réunions annuelles. Je remercie chaleureusement Emmanuel Berthier (CEREMA Île-de-France, Trappes), François Bourgin (Université Gustave Eiffel, Nantes, puis INRAE, Antony), Pierre Brigode (Université Côte d'Azur, Polytech Nice Sophia, Nice), Charles Perrin (INRAE, Antony) et Pierre-Antoine Versini (École des Ponts ParisTech, Marne-la-Vallée). Je réitère mes remerciements à Charles Perrin d'avoir accepté de représenter l'école doctorale et de s'être occupé de rédiger les rapports de suivi.

J'ai préparé ma thèse dans un environnement technico-administratif très agréable. Je tiens donc à remercier Dovy Tristani Beguerie et Laurence Pastor, gestionnaires de l'école doctorale GRNE, et Loïc Labrousse, directeur de l'école doctorale, de leurs efforts pour former et gérer les dossiers des doctorants. Merci également à Véronique Lachasse et Christine Franke pour avoir participé à l'organisation des journées des doctorants. Côté METIS, je remercie Christophe Bonnet, Bénédicte Sabatier, Sylvain Théry et Aurélien Baro pour l'accompagnement informatique; Valérie Girard, Alban Bakhti, Francis Chabrier et Nadine Bost pour avoir facilité la gestion de toutes mes missions; et Anun Martinez pour la formation sur l'hygiène et la sécurité. Je tiens à remercier l'équipe administrative du laboratoire, notamment Jean-Marie Mouchel,

Agnès Ducharne et H  l  ne Blanchoud pour leurs efforts consid  rables surtout depuis la crise sanitaire. Merci   galement    Hafida et Thushanty de leur veille quotidienne sur la propret   du milieu de travail. Je tiens      galement remercier Marine Dangeard, C  cile Finco, Alexandre Thibault, Pierre V  quaud, Guillaume Bonnet, Benjamin Lefeuvre et surtout Flore Rembert et No  lie Molbert d’avoir repr  sent   les doctorants aupr  s du laboratoire et de l’  cole doctorale.

Je suis tr  s reconnaissant    tous ceux qui ont contribu      la mise    disposition du grand volume de donn  es et des codes en acc  s ouvert (surtout les packages en \mathbb{R} et en \LaTeX) dont les travaux de ma th  se ont amplement b  n  fici  . En particulier, je remercie Olivier Laurantin et Sylvain Denizanne de M  t  o France ainsi que Pierre Brigode pour la mise    disposition des donn  es COMEPHORE. Je remercie Jordan Read, David Blodgett, David Watkins et William Watkins de l’USGS, Tammy Walker et Michele Thornton de l’ORNL DAAC d’avoir facilit   l’acquisition des donn  es nord-am  ricaines et d’avoir   t   tr  s r  actifs en plein *Government Shutdown* de 2019.

Pour leur excellent cours sur la mod  lisation hydrologique des bassins versants, je remercie Vazken Andr  assian, Charles Perrin, Maria-Helena Ramos et Olivier Delaigue. Merci encore une fois    Olivier pour sa formation sur \mathbb{R} qui m’  tait d’une grande utilit  .

Pour m’avoir accompagn   dans ma mission d’enseignement, merci encore une fois    mes directeurs de th  se,    Val  rie Plagnes,    Nicolas Le Moine,    Dani  le Valdes et    Jamel Benmamar. Merci   galement    Julien Thiesson,    Ludovic Bodet et    Roger Gu  rin qui m’accompagnent en tant qu’ATER. Merci    Laurent Riquier et    Isabelle Bauwe pour la gestion technique et administrative.

Pendant deux ans et onze mois, j’ai partag   le Bureau 401 avec beaucoup de personnes aimables que j’aimerais remercier : Salma, ma jumelle de th  se, qui m’a tr  s amicalement accueilli lors de mon premier jour    METIS; Alexandre qui m’a appris    r  soudre mon premier *Rubik’s Cube* alors que j’  crivais mon premier article et avec qui j’ai eu des discussions tr  s instructives sur beaucoup de sujets d’histoire et de politique; Goulven qui ne m’a pas   pargn   sa bont   en me prodiguant beaucoup de conseils (et de bonbons); mon cher Thomas, qui me manque beaucoup et    qui je souhaite une tr  s bonne chance; *mi hermano* Pedro, qui ne cesse de me rappeler que tout ira bien et qu’il faut que je me repose. H  las, *ce n’est pas ainsi faire de la Science*; Thibault, dont le s  rieux annonce une excellente th  se; et finalement Radegonde, Jeanne et Anne-Sophie, qui m’ont soutenu pendant leur court passage    METIS et    qui je souhaite un tr  s bon courage pour la suite.

J’ai eu la chance de rencontrer    METIS des coll  gues tr  s sympathiques sans qui cette aventure aurait   t   beaucoup moins agr  able. Je pense    M. Paul et    tonton Sylvain,    qui je t  moigne ma profonde gratitude pour leur bienveillance et leurs amiti  s. Un grand merci    Marie (Mme Larousse) et    Vincent (M. Le Petit Robert, AAAAAAA) de leur soutien, leurs encouragements et surtout de leur patience. Je pense   galement    Audrey (l’  trang  re de l’autre bout) qui

m'a offert mon premier café à METIS et qui depuis ne cesse de veiller sur mon parcours. Je suis reconnaissant à M. Marc Le Dumont pour m'avoir soutenu et conseillé (avec beaucoup d'amour) tant sur le plan académique que sur le plan personnel. Merci à Amen pour ses conseils très pertinents souvent donnés autour d'un repas de midi à la cantine. Merci à Mounir d'être celui qui me comprend toujours mieux que les autres. Il sera très chagrinant de ne plus entendre le mélodique « Hello Momo ! » de Noëlie, de ne plus bénéficier de la bienveillance maternelle de Diane (qui se rappelle très bien du vendredi 43%), de ne plus participer aux épisodes du coin café avec Flore S., Manon, Pierre, Mme Danièle et Romane (l'experte en contrepèteries) qui ont toujours fait preuve de patience devant mes discours souvent obscurs (que moi-même je ne comprends pas toujours). J'ai eu l'honneur d'avoir décroché les titres de « Cornichon » et « Fenouil » après avoir férocelement milité auprès du Stradivarius de SIG, M. Baro, que je remercie des logos haute qualité de SU et METIS, du carnet « zéro en géographie » et de son aide dans l'acquisition des données de la banque HYDRO. Merci à Fabrice et à Laurence M. de leur amabilité et de leur soutien, ainsi qu'aux personnes agréables qui ont quitté le laboratoire durant ma thèse : Matthieu, Thomas Le Grand, Elsa, Ardalan, Marine, Julia, Ningxin et Fanny. Merci également à Josette, Gilles, Alexis, Anne, Michel, Laurence L., aux organisateurs des séminaires de vendredi, Aurélie & Damien puis Frédéric & Thomas, et à tous les chercheurs et chercheuses de METIS qui m'ont inspiré par leur dévouement et leur dynamisme. Je souhaite un très bon courage aux doctorants de METIS et de GRNE que je n'ai pas encore mentionnés, notamment les géophysiciennes Aida, Flore R. et Kawtar, les hydrogéochimistes Nouradine, Axel, Xingcheng et Gurpreet, et les hydrologues d'INRAE Daniela, Paul R.-G., Paul A. et Anthony.

Enfin, je remercie les Coste et Corinne de m'avoir gentiment et chaleureusement accueilli sous leurs toits. Merci infiniment à Sakina et à Mouad pour leur accueil et leur soutien inconditionnels. Je remercie également Mehdi pour avoir été toujours disponible et pour ses encouragements. Je dédie ce travail à ma famille et surtout à mes parents, ceux à qui je dois tout.

Cette liste est, toutefois, loin d'être exhaustive.

RÉSUMÉ

Afin de déterminer l'impact hydrologique de l'étalement urbain à venir, la modélisation hydrologique est très utile car elle permet de traduire les scénarios d'urbanisation en scénarios hydrologiques en tenant compte de la variabilité climatique. Malgré la diversité actuelle des modèles hydrologiques applicables sur les bassins urbanisés, ces modèles sont en général développés sur un nombre limité de bassins versants, ce qui pose la question de leur robustesse. L'objectif de cette thèse est de chercher une structure de modèle hydrologique capable de reproduire la relation pluie-débit pour un large échantillon de bassins versants urbanisés, situés en France et aux États-Unis, et caractérisés par différents niveaux d'étalement urbain.

Pour ce faire, nous sommes partis d'une structure de modèle conceptuel (GR4H), fonctionnant au pas de temps horaire mais non dédiée aux bassins versants urbanisés. Pour arriver à nos objectifs, nous avons considéré quatre étapes : **(1)** la collecte et la préparation d'un large échantillon de données pour plus de 2000 bassins versants situés en France et aux États-Unis. L'analyse de la variabilité des coefficients de ruissellement de 852 bassins nord-américains a montré une dépendance non seulement au taux moyen d'imperméabilité totale (TIA), mais aussi aux conditions d'humidité du sol quel que soit le niveau d'étalement urbain du bassin versant; **(2)** le test de la structure de départ sur les bassins fortement urbanisés. Cette étape a montré l'aptitude des paramètres du modèle GR4H à traduire les spécificités du comportement hydrologique en contexte urbanisé; **(3)** la régionalisation des paramètres de la structure de départ GR4H pour les lier à la fraction du couvert urbain. Les relations de régression obtenues ont été performantes mais peu sensibles à la caractérisation du couvert urbain; **(4)** la modification pas-à-pas de la structure de GR4H en rajoutant le ruissellement sur les surfaces imperméables, et en variant le partage entre écoulement rapide et écoulement lent en fonction de chaque bassin. L'évaluation de ces modifications sur 273 bassins versants urbanisés a indiqué une amélioration de la simulation des écoulements surtout lors des périodes sèches et des périodes de crue. Les paramètres rajoutés ont montré une forte dépendance au taux moyen d'imperméabilité totale.

Les travaux de cette thèse montrent un exemple de démarche constructive de développement de modèle hydrologique, par laquelle nous avons réussi à améliorer la représentation de l'urbanisation au sein du modèle tout en obtenant de meilleurs résultats en comparant avec des séries observées. Un effort de régionalisation de la nouvelle structure en exploitant des mesures de fragmentation du paysage urbain permettra une étude d'impact de différents schémas d'urbanisation future sur la réponse hydrologique.

ABSTRACT

In order to determine the hydrological impact of future urban development, hydrological modeling is very useful because it allows obtaining hydrological scenarios from urbanization scenarios by considering climate variability. Despite the current diversity of applicable hydrological models to urbanized catchments, those models were generally tested on few catchments, which raises the question of their robustness. The objective of this thesis is to look for a hydrological model structure that is able to reproduce the rainfall-runoff relationship for a large sample of urbanized catchments, located in France and the United States, and characterized by a wide variety of urban sprawl.

To do so, we started from an hourly conceptual model structure (GR4H) which was not specifically developed nor tested on highly urbanized catchments. We followed four steps: **(1)** The collection and preparation of a large sample of data for more than 2000 catchments located in France and the United States. The analysis of the variability of event runoff ratio for 852 US catchments showed the dependence of catchment response not only on mean total impervious area (TIA), but also on antecedent soil moisture conditions regardless of the level of urban sprawl in the catchment; **(2)** The test of the starting model structure on highly urbanized catchments. This step showed the ability of GR4H parameters to reflect the behavioral specificities of urbanized catchments; **(3)** The regionalization of GR4H parameters in order to relate them to the fraction of urban cover in the catchment. The obtained regression relationships performed well but showed weak sensitivity to change in the fraction of urban cover, which impeded their use for scenario prediction; **(4)** The step-by-step modification of the GR4H model structure by including excess rainfall on impervious surfaces, and by varying the partitioning between quick flow and slow flow depending on each catchment. The evaluation of these modifications on 273 urbanized catchments indicated an improvement in reproducing the observed flows in comparison to the original structure, especially during dry and flooding periods. The added parameters showed a strong dependence on the mean total impervious area.

This thesis work illustrates a constructive modeling approach by which we have succeeded in both **(1)** improving the representation of urbanization within the model and **(2)** obtaining better simulation of catchment response. An effort of regionalizing the new model structure using measures of urban landscape fragmentation will allow to better estimate the impact of different urban planning schemes on the hydrological behavior at the catchment scale.

TABLE DES MATIÈRES

Résumé	v
Abstract	vii
Table des matières	ix
Table des figures	xv
Liste des tableaux	xix
Introduction générale	1
1 Modélisation hydrologique des bassins versants urbanisés	11
1.1 Introduction	13
1.2 Quel est l'impact de l'urbanisation sur les processus hydrologiques?	14
1.2.1 Qu'est-ce que la réponse d'un bassin versant?	14
1.2.2 Impact de l'urbanisation sur les processus hydrologiques à l'échelle locale	16
1.2.3 Impact de l'urbanisation sur le comportement hydrologique à l'échelle du bassin versant	27
1.3 Aperçu des approches de modélisation des bassins versants urbanisés	40
1.3.1 Notions générales et classifications classiques	40
1.3.2 Panorama des modèles hydrologiques en milieu urbanisé	42
1.3.3 Vers une approche hybride	57
1.4 Synthèse	60
2 Échantillon des bassins versants	63
2.1 Introduction	65
2.2 Données hydroclimatiques	67
2.2.1 Données de débit	68
2.2.2 Données de précipitation	69
2.2.3 Données d'évapotranspiration potentielle	72
2.2.4 Récapitulatif des sources et des disponibilités de données hydroclima- tiques	77

2.3	Sélection des bassins versants	77
2.3.1	Critères de sélection des bassins versants	78
2.3.2	Étapes de sélection	79
2.3.3	Aperçu des caractéristiques événementielles et des bilans hydrologiques des bassins versants sélectionnés	83
2.3.4	Est-il nécessaire d'effectuer un tri supplémentaire dans les données? .	88
2.4	Caractéristiques d'occupation du sol et d'urbanisation	89
2.4.1	Mesures utilisées	89
2.4.2	Variabilité spatiale des cas d'urbanisation et différences d'occupation du sol entre la France et les États-Unis	97
2.4.3	Évolution temporelle de l'urbanisation de chacun des bassins versants sélectionnés	100
2.5	Caractéristiques hydroclimatiques de l'échantillon	100
2.5.1	Caractéristiques hydroclimatiques choisies	102
2.5.2	Variabilité spatiale des caractéristiques hydroclimatiques en France et aux États-Unis	103
2.5.3	Contrastes hydroclimatiques entre les bassins situés en France et les bassins situés aux États-Unis	108
2.5.4	Caractéristiques hydroclimatiques des bassins fortement urbanisés . .	109
2.6	Caractéristiques morphologiques de l'échantillon	111
2.6.1	Caractéristiques morphologiques choisies	111
2.6.2	Variabilité spatiale des caractéristiques morphologiques en France et aux États-Unis	113
2.6.3	Différences morphologiques entre les bassins situés en France et les bassins situés aux États-Unis	116
2.6.4	Caractéristiques morphologiques des bassins fortement urbanisés . .	117
2.7	Caractéristiques géopédologiques de l'échantillon	118
2.7.1	Caractéristiques géopédologiques choisies	118
2.7.2	Variabilité spatiale des caractéristiques géopédologiques en France et aux États-Unis	119
2.7.3	Contrastes géopédologiques entre les bassins situés en France et les bassins situés aux États-Unis	123
2.7.4	Caractéristiques géopédologiques des bassins fortement urbanisés . .	123
2.8	Synthèse	126
2.8.1	Un large échantillon avec des caractéristiques climatiques et d'urbanisa- tion très variées	126
2.8.2	Discussion sur les choix de modélisation : représentation globale au pas de temps horaire	130
2.8.3	Échantillonnage des bassins versants pour les prochains chapitres . .	130

3	Processus de génération du ruissellement en milieu urbanisé	133
3.1	Introduction	136
3.2	Dataset and methods	137
3.2.1	Catchment sample and spatial characteristics	137
3.2.2	Time-varying event and catchment characteristics	145
3.2.3	Relevance of catchment imperviousness and ASM as controls of the spatial and temporal variability of RR	150
3.2.4	Investigation of the relationship between RR and ASM for different levels of catchment imperviousness	151
3.3	Results	153
3.3.1	Assessing the relevance of catchment imperviousness and ASM as explanatory variables of event-scale RR	153
3.3.2	Analyzing RR variability using TIA and SMAR as a surrogate of ASM	155
3.4	Discussion	158
3.4.1	Importance of ASM as a control of RR in urbanized catchments	158
3.4.2	Relevance of other catchment and event characteristics in explaining the variability of RR	160
3.5	Conclusion	161
4	Comment les modèles conceptuels pluie-débit gèrent-ils les spécificités des bassins versants urbanisés?	165
4.1	Introduction	167
4.1.1	From contrasting landscapes to contrasting hydrological modeling practices	167
4.1.2	Scope of the study	168
4.2	Catchments and methodology	169
4.2.1	Sample description	169
4.2.2	Hydrological signatures used to assess the differences between rural and urbanized catchments	172
4.2.3	Assessment of the differences between urbanized and rural catchments from a rural modeling standpoint	173
4.3	Results	177
4.3.1	Are urbanized catchments hydrologically different from their rural neighbors?	177
4.3.2	Are there differences between urbanized and rural catchments observable from a rural modeling standpoint?	179
4.4	Discussion	186
4.4.1	Impact of urbanization on the hydrological behavior of catchments	186

4.4.2	Differences between urbanized and rural catchments from a modeling standpoint	187
4.5	Conclusion	188
4.6	Supporting information	189
5	Les modèles conceptuels pluie-débit sont-ils régionalisables vis-à-vis de l'urbanisation?	195
5.1	Introduction	197
5.1.1	Why and how do we regionalize hydrological model parameters? . . .	197
5.1.2	Random forest: a potentially useful tool for regionalization	199
5.1.3	Application of RF for model regionalization in multiple land-use environments	199
5.1.4	Context and scope of the study	200
5.2	Data	201
5.2.1	Sample selection	201
5.2.2	Catchment descriptors	203
5.3	Methods	206
5.3.1	Model parameters and calibration	206
5.3.2	Estimating the model parameters at ungauged locations using RF . . .	209
5.3.3	Benchmark regionalization techniques	210
5.4	Results	211
5.4.1	Model performances and estimated parameters	211
5.4.2	Descriptor importance	215
5.5	Discussion and conclusions	219
5.5.1	Regionalization with RF: what is appreciated and what is depreciated?	219
5.5.2	Weak sensitivity of the RF-derived relationships with the urbanization measure	221
5.5.3	Conclusions and perspectives	221
6	Un modèle pluie-débit conceptuel et physiquement cohérent pour les bassins versants urbanisés	225
6.1	Introduction	227
6.1.1	Hydrological models as valuable tools for the assessment of urbanization impact	227
6.1.2	Overview of hydrological models for urbanized catchments	228
6.1.3	Novelty and scope of the study	229
6.2	Catchment sample	230
6.3	Tested model modifications	232
6.3.1	Original model structure	232

6.3.2	Tested modifications to account for urbanization features and improve model performances	234
6.3.3	Evaluation of tested modifications	237
6.4	Results	241
6.4.1	Calibration performances and distributions of calibrated parameters .	241
6.4.2	Test performances	244
6.4.3	Were the improvements greater with increasing urbanization level? Investigation of the continuous assessment performances of MU5H . . .	248
6.5	Discussion and conclusion	248
6.5.1	Comparison with existing model testing experiments	248
6.5.2	Potentials of newly added parameters and further improvements . . .	251
	Conclusion générale	255
	Références bibliographiques	267
A	Méthode de séparation des événements	297
B	Équations du réservoir de l'humidité du sol	303
C	Compléments au Chapitre 4	307
C.1	Équations du modèle GR4H	308
C.1.1	Fonction de production	308
C.1.2	Fonction de transfert	310
C.2	Équations du modèle IHACRES	313
C.2.1	Fonction de production	313
C.2.2	Fonction de transfert	314
D	Une méthode d'estimation du taux moyen de l'imperméabilité effective	317
E	Compléments au Chapitre 6	321
E.1	Modèles testés	322
E.1.1	6 structures conceptuelles non dédiées aux bassins urbanisés	322
E.1.2	6 structures issues de l'optimisation de paramètres fixes et de la suppression de l'échange entre le réservoir souterrain et les branches de routage	323
E.1.3	12 structures avec prise en compte du taux moyen d'imperméabilité totale TIA	324
E.1.4	15 structures avec optimisation de X_5 et X_6	325

E.1.5	4 structures comprenant des formulations alternatives du réservoir de production pour prendre en compte l'intensité de précipitation au pas de temps horaire	326
E.2	Stratégie de calage et d'évaluation	329
E.3	Classement des 43 modèles	331

TABLE DES FIGURES

1.1	Échelle spatiale et temporelle des processus hydrologiques en milieu urbanisé.	28
1.2	Représentation d'un bassin versant urbanisé et des principaux processus hydrologiques.	30
1.3	Distribution du nombre des bassins versants utilisés pour les applications des 45 modèles.	57
1.4	Synthèse des approches actuelles de modélisation en milieu urbanisé.	58
2.1	Calcul du débit moyen horaire à l'aide des mesures de débit instantanées. . . .	69
2.2	Coefficients de Thiessen pour le bassin versant drainé par Le ru de Choqueuse à Doue.	71
2.3	Coefficients de Thiessen choisis comme poids pour déterminer la température moyenne à l'échelle du bassin versant.	73
2.4	Exemple de l'ETP moyenne horaire pour deux bassins versants à deux longitudes différentes.	76
2.5	Contours et localisations des exutoires des 870 bassins versants nord-américains.	80
2.6	Contours et localisations des exutoires des 1355 bassins versants français. . . .	82
2.7	Distributions des caractéristiques des événements extraits des séries pluie-débit des 2225 bassins versants.	85
2.8	Bilan hydrologique interannuel des 2225 bassins versants sélectionnés.	87
2.9	Relations empiriques entre <i>CPD</i> et (a) <i>TIA</i> , (b) <i>IMP80</i> et (c) <i>IrIMP</i>	96
2.10	Courbes cumulées des 6 métriques utilisées pour la description de l'occupation du sol.	98
2.11	Évolution temporelle des courbes cumulées de la fraction du couvert urbain <i>CPD</i> pour les bassins français et nord-américains.	101
2.12	Caractéristiques hydroclimatiques des 1355 bassins versants français.	105
2.13	Caractéristiques hydroclimatiques des 870 bassins versants nord-américains.	107
2.14	Courbes cumulées des caractéristiques hydroclimatiques pour l'échantillon total, les bassins situés en France et les bassins situés aux États-Unis.	109
2.15	Courbes cumulées des caractéristiques hydroclimatiques pour l'échantillon total vs. les bassins versants fortement urbanisés.	110
2.16	Caractéristiques morphologiques des 1355 bassins versants situés en France. .	114
2.17	Caractéristiques morphologiques des 870 bassins versants situés aux États-Unis.	115

2.18	Courbes cumulées des caractéristiques morphologiques de l'échantillon total, du sous-échantillon des bassins situés en France et des bassins situés aux États-Unis.	116
2.19	Courbes cumulées des caractéristiques morphologiques pour l'échantillon total vs. les bassins fortement urbanisés.	117
2.20	Caractéristiques géopédologiques des 1355 bassins versants situés en France.	120
2.21	Caractéristiques géopédologiques des 870 bassins versants situés aux États-Unis.	122
2.22	Courbes cumulées des caractéristiques géopédologiques de l'échantillon total, des bassins situés en France et des bassins situés aux États-Unis.	124
2.23	Courbes cumulées des caractéristiques géopédologiques pour l'échantillon total vs. les bassins fortement urbanisés.	125
3.1	Location and aridity index of the 852 US catchments.	138
3.2	Relative importance of space-varying and time-varying factors in predicting the event-scale runoff ratio.	154
3.3	Runoff ratio dependence on antecedent soil moisture conditions, assessed using the level of the soil moisture accounting reservoir (<i>SMAR</i> – 50), and the catchment total impervious area (<i>TIA</i>).	156
3.4	Scatter plots showing the variability of mean runoff ratio as a function of total impervious area (<i>TIA</i>) and antecedent soil moisture (<i>ASM</i>), assessed using the level of the soil moisture accounting reservoir (<i>SMAR</i> -50).	157
3.5	(a) Evolution of the parameters of the WLS regression line representing mean runoff ratio as a function of antecedent soil moisture conditions for different levels of total impervious area (<i>TIA</i>), and (b) evolution of the corresponding lines for different levels of <i>TIA</i>	159
4.1	Location of the United States urbanized catchment sample and their rural neighbors.	170
4.2	Location of the French urbanized catchment sample and their rural neighbors.	171
4.3	(a) GR4H and (b) IHACRES model structures.	174
4.4	Values of the hydrological signatures for urbanized catchments vs. their rural neighbors.	178
4.5	Distributions of GR4H and IHACRES performances in calibration and test over the whole test period, with a focus on the winter and the summer periods. . .	180
4.6	Distributions of GR4H parameters over the urbanized catchments and their rural neighbors.	182
4.7	Distributions of IHACRES parameters in the urbanized catchments and their rural neighbors.	183
4.8	Comparison between different GR4H and IHACRES internal fluxes and states.	185
5.1	The location and mean Catchment Percent Developed (<i>CPD</i>) of the 870 U.S. catchments.	202

5.2	The location and mean Catchment Percent Developed (<i>CPD</i>) of the 1355 French catchments.	202
5.3	GR4H model structure.	208
5.4	Random forest (RF) construction using the 2105 rural and urban catchments.	210
5.5	Validation scores on P2 in terms of (a) <i>KGESR</i> (b) Kling-Gupta efficiency (<i>KGE</i>) (c) <i>NSESR</i> and (d) Nash-Sutcliffe efficiency (<i>NSE</i>) using calibrated parameters on P1, estimated parameters using random forest, transferred parameters from the closest catchments, and transferred parameters from the most similar catchments.	212
5.6	Observed hourly flow time series and simulated flow using parameters transferred from calibration over P1, estimated parameters using random forest, and transferred parameters from the closest and the most similar catchment to the Suwanee Creek catchment, as response to measured precipitations between 1 March and 31 August 2013.	215
5.7	Rankings of catchment descriptors based on their importance in predicting (a) X_1 , (b) X_2 , (c) X_3 , and (d) X_4	217
5.8	Distribution of parameters (a) X_1 , (b) X_2 , (c) X_3 , and (d) X_4 . <i>CAL_RUR</i> are the rural calibrated parameters (with respect to <i>KGESR</i>), <i>RF_RUR</i> are the RF-estimated rural parameters, <i>RF_UPDATED</i> are the RF-estimated parameters over the rural sample with transferred <i>CPD</i> value from the corresponding urban catchment, <i>RF_URB</i> are the RF-estimated parameters over the urban sample, and <i>CAL_URB</i> are the calibrated parameters.	218
6.1	Location of the 205 US catchments.	231
6.2	Location of the 68 French catchments.	231
6.3	Description of GR4H model structure, parameters, states, and internal fluxes.	233
6.4	Illustration of the effect of different values of X_5 , representing conceptually the proportion of sealed surfaces, on the relationship between the excess rainfall rate $\frac{p_{eff}}{p_{th}}$, and $\frac{Prod}{X_1}$, the filling rate of the soil moisture-accounting reservoir.	236
6.5	Distributions of improvements in calibration performances obtained for the six tested model structures, expressed in terms of relative R_{KGESQ}	241
6.6	Calibrated proportion of sealed surfaces X_5 vs. estimated total impervious area from land cover databases.	242
6.7	Calibrated quick-flow/slow-flow partitioning parameter X_6 vs. estimated total impervious area from land cover databases.	243
6.8	Distributions of improvements in test performances for the five tested modifications, using six criteria for continuous assessment.	245
6.9	Summary of the improvements in median assessment criteria.	247
6.10	Improvements in reproducing the catchment response function of the total impervious area of the catchment during the test subperiod.	249
A.1	Examples of correlograms computed for the catchment drained by the USGS station 01658000 Mattawoman Creek near Pomonkey, Maryland.	299

B.1	Description of the soil moisture accounting reservoir (<i>SMAR</i>) equations. . . .	305
E.1	Variabilité du taux de ruissellement $1 - \frac{p_s}{p_{th}}$ en fonction de l'état du réservoir de production $\frac{Prod_{GR}}{X_{1,GR}}$ pour différentes valeurs de α	328

LISTE DES TABLEAUX

1.1	Récapitulatif de l'impact de l'urbanisation sur les processus hydrologiques à l'échelle locale.	23
1.2	Revue bibliographique de l'impact de l'urbanisation sur différents aspects du comportement hydrologique des bassins versants.	35
1.3	Revue bibliographique des modèles hydrologiques utilisés pour les bassins versants urbanisés.	44
2.1	Récapitulatif des sources de données hydroclimatiques et de leurs périodes de disponibilité pour les bassins situés en France et aux États-Unis.	77
2.2	Classes d'occupation du sol dans la base de données NLCD.	90
2.3	Classes et sous-classes d'occupation du sol pour le produit CORINE Land Cover.	91
2.4	Synthèse des caractéristiques de l'échantillon des 2225 bassins versants sélectionnés.	128
2.5	Synthèse des échantillons de bassins versants utilisés dans chacun des prochains chapitres.	132
3.1	Description of catchment characteristics, their estimation methods, and their sources.	140
3.2	Summary of the time-varying event characteristics and antecedent soil moisture (<i>ASM</i>) proxies.	147
4.1	Description of GR4H and IHACRES model parameters and their corresponding roles.	176
4.2	Comparison of GR4H and IHACRES internal fluxes and states.	177
4.3	Results of statistical tests investigating distribution and mean equality between the urbanized and rural samples in terms of hydrological characteristics, model performances, and model parameters.	191
5.1	Description of the different climatic, topographic, land-use, and geopedological characteristics estimated for each catchment.	204
5.2	R^2 computed between the calibrated parameters with respect to different objective functions (<i>KGESR</i> , <i>KGE</i> , <i>NSESR</i> , and <i>NSE</i>) and the estimated parameters via RF, transferred parameters from the close and from the similar catchments for the 120-urban-catchments sample.	213

5.3	Sets of parameters transferred from calibration over P1, estimated using <i>RF_ALL</i> and <i>RF_URB</i> , and transferred from the closest catchment and the most similar catchment for the Suwanee Creek catchment.	216
6.1	Summary of tested model modifications.	238
6.2	Median values of event-based assessment criteria for the six tested models over 45,025 events.	246
A.1	Summary of the steps of the event separation method.	301
E.1	Classement des 43 modèles testés.	332

INTRODUCTION GÉNÉRALE

Contexte sociétal et scientifique des travaux

Une urbanisation récemment galopante

L'urbanisation est un processus socio-économique qui s'accompagne de deux mutations principales : une migration de la population des zones rurales vers des zones densément anthropisées et une conversion du paysage naturel à un paysage rendu plus complexe par les constructions et les différents aménagements. À travers ce processus, l'homme aboutit au plus large et au plus compliqué objet qu'il pourrait construire : les villes (Hillier et Penn, 1992). Dès l'aube des civilisations, les villes sont des carrefours où les échanges sont maximisés, ce qui est à la fois racine et fruit de la richesse des services offerts, en si peu d'espace, par les différents établissements citadins : nourriture, santé, protection, commerce, éducation, etc. Ces bénéfices de l'agglomération marquent la prospérité de la vie en ville par rapport à la campagne, ce qui réduit le taux de mortalité et contribue à l'augmentation de la population urbaine. Attractive, la ville est la destination de la jeune population aspirant à améliorer son cadre de vie. Cette augmentation de la population urbaine conjuguée à un désir fort de propriété privée et à un rejet de la promiscuité ont contribué à l'accélération de l'étalement urbain (Hillier et Penn, 1992; United Nations, 2019). Depuis l'entrée de l'homme dans l'ère de l'anthropocène (Crutzen, 2002), caractérisée par des avancées techniques et scientifiques inédites, le pourcentage de population mondiale vivant dans les villes ne cesse d'augmenter : de 30% en 1950, le pourcentage a grimpé jusqu'à 55% en moins de 70 ans, et on s'attend à 68% en 2050. L'Europe et l'Amérique du Nord figurent parmi les régions du globe les plus urbanisées, avec 74% et 82% de leurs populations respectives vivant en milieu urbain (United Nations, 2019).

L'eau comme ressource primordiale à la ville

Si le nombre et la taille des agglomérations urbaines ont évolué au cours de l'histoire, la ville n'a jamais cessé d'être dépendante de l'eau. L'existence de l'homme comme espèce consommatrice est dépendante d'une eau de bonne qualité mobilisée en quantité suffisante. Une brève lecture de l'histoire montre que la santé publique est synonyme d'une eau citadine propre (Vuorinen, Juuti et Katko, 2007). La question hygiéniste ne concerne pas uniquement l'alimentation en eau potable, mais également la présence d'un milieu aquatique récepteur des déchets des agglomérations pour en diluer les éléments chimiques nuisibles. À l'ère moderne, les activités industrielles en milieu urbain ont amplifié le besoin en eau comme matière indispensable pour les différents procédés industriels. De surcroît, l'eau a servi comme moyen de transport des gens et des marchandises (Carré et Deutsch, 2015, p. 7). D'après ce panorama

des services rendus par l'eau, on ne s'étonne pas que les berceaux des civilisations fussent toujours les grands fleuves.

La ville est un lieu de concentration des usages des différentes ressources et de la génération des déchets. Ainsi, la satisfaction des besoins citadins en matière d'eau ne se fait pas dans l'autosuffisance mais en dépendance à l'hinterland rural. Force est donc de considérer un périmètre pour le cycle de l'eau plus large que la ville, comme le bassin versant hydrographique. Considérer le cycle de l'eau à cette échelle s'impose d'autant plus que la ville engendre des modifications drastiques sur le comportement hydrologique de l'ère préurbaine.

Prise de conscience de l'impact de l'urbanisation sur l'hydrologie

Si la ville bénéficie amplement des ressources hydriques du voisinage, ceci s'accomplit avec une modification du cycle de l'eau, au moins au niveau local. La forte densité de population conduit à une surexploitation des ressources en eaux souterraines et superficielles avoisinantes. Les rejets des agglomérations constituent un flux qui impacte le bilan hydrologique et compromet surtout la qualité des cours d'eau. En outre, l'urbanisation représente un changement d'occupation du sol, ce qui se traduit par une modification des processus hydrologiques à l'échelle locale, favorisant le ruissellement sur l'infiltration (Hollis et Ovenden, 1988; Ragab *et al.*, 2003b; Ragab *et al.*, 2003a; Ramier, Berthier et Andrieu, 2004). La mise en place de réseaux de drainage artificiels et l'aménagement des cours d'eau naturels accélèrent les vitesses de parcours des eaux ruisselantes, ce qui rend la réponse hydrologique plus violente et accroît le risque lié à l'eau en milieu urbain.

À partir de la fin des années 1960, un bouleversement d'opinion poussa à questionner la dimension écologique des activités anthropiques, en particulier l'impact de l'urbanisation sur le cycle de l'eau (Jacobson, 2011). Ceci a fait émerger le concept de développement durable selon lequel la planification urbaine est désormais confrontée à deux contraintes, en général contradictoires (Cortinovis et Geneletti, 2018; United Nations, 2019) : améliorer le bien-être des citoyens en maximisant les bénéfices offerts par la ville, mais aussi préserver les écosystèmes naturels en minimisant l'impact de l'agglomération sur le régime naturel, défini par le régime qu'il y aurait s'il n'y avait pas eu d'urbanisation. En particulier, la planification cherche à minimiser l'impact de l'urbanisation sur le cycle hydrologique et de la présence de la ville sur les ressources en eau. Ceci se fait grâce à l'adoption de plusieurs techniques de gestion de l'eau à la source, désignées communément par *techniques alternatives*, dont l'implémentation est encouragée/imposée par plusieurs législations (comme la loi n°92-3 de 1992 sur l'eau, ou la loi n°2014-366 de 2014 sur l'urbanisme rénové). Les décideurs doivent alors intégrer cet objectif de gestion à la source dans leurs choix politiques d'aménagement. Pour cela, ils font appel aux hydrologues pour (1) comprendre et quantifier l'impact de l'urbanisation sur le cycle de l'eau

à partir des cas historiques, (2) définir et évaluer l'efficacité des mesures d'atténuation, et (3) se servir de la compréhension du fonctionnement hydrologique d'un bassin versant urbanisé pour prédire l'impact de schémas d'urbanisation future, ce qui guidera le choix de tel ou tel schéma de planification en fonction de son impact. L'échelle exigée pour la compréhension puis la prévision ne peut plus se contenter de la ville seule; les approvisionnements et les rejets des citadins se faisant à une échelle qui dépasse la ville, une planification adaptée doit se définir à l'échelle du bassin versant hydrographique (Carré et Deutsch, 2015, p. 33).

Les hydrologues face aux défis de l'échelle

En dépit d'une bonne connaissance de l'impact de l'urbanisation sur les processus hydrologiques à l'échelle locale, le passage à l'échelle du bassin versant rend plus incertaine cette connaissance à cause de l'existence de mécanismes compensatoires qui peuvent neutraliser l'impact (Salavati, 2015).

À l'échelle locale, le comportement des surfaces urbaines est très variable selon la nature et le type de la surface et de son âge (Redfern *et al.*, 2016). Ensuite, la connectivité des surfaces urbaines au réseau de drainage conditionne la propagation de leur impact jusqu'à l'exutoire du bassin versant. Enfin, l'existence de surfaces naturelles à l'aval des surfaces urbaines (dans le sens du parcours hydraulique), ou même la présence de techniques alternatives pour une gestion adaptée, contribuent à la variabilité de l'impact de l'urbanisation d'un cas d'étude à l'autre.

À l'échelle du bassin versant, l'évidence de l'impact de l'urbanisation sur le comportement hydrologique est attestée par une littérature abondante (Fletcher, Andrieu et Hamel, 2013; Jacobson, 2011; McGrane, 2016). Un consensus existe sur une augmentation des forts débits à l'échelle du bassin versant, contrairement à une controverse actuelle sur l'évolution des bas débits et de l'écoulement moyen. De plus, la quantification et le jugement de la significativité du changement hydrologique par l'urbanisation reste encore une question ouverte. Il est très important de signaler ici que si l'exercice de détection peut être qualitatif et se contenter d'un ou plusieurs cas, l'exercice de projection exige une compréhension et une quantification de l'impact sur le plus grand nombre possible de cas d'urbanisation. Or, les études hydrologiques des milieux urbanisés se caractérisent par un aspect *monographique*, que l'on peut expliquer d'une part par le caractère individuel des formes urbaines, et d'autre part par la difficulté d'accès aux données requises pour étudier plusieurs cas à la fois.

Les modèles hydrologiques comme outils de compréhension et de prédiction de l'impact de l'urbanisation

Pour mettre en évidence l'impact de l'urbanisation sur le comportement hydrologique, on a recours à des méthodes principalement statistiques de détection et d'attribution en exploitant des mesures de terrain. L'exercice d'attribution est amélioré par l'utilisation de modèles hydrologiques, permettant notamment de distinguer l'effet de l'urbanisation de l'effet de la variabilité climatique sur l'évolution de la réponse hydrologique. Par ailleurs, l'exercice de projection se fait en priorité avec des modèles hydrologiques, grâce à la prise en compte des caractéristiques physiques du bassin versant à travers les paramètres du modèle, ce qui permet de passer de scénarios d'urbanisation à des scénarios de comportement hydrologique.

À l'échelle du bassin versant, on se sert de modèles hydrologiques, ou modèles pluie-débit, pour appréhender le cycle de l'eau et les processus hydrologiques associés. Historiquement, les objectifs de modélisation en relation avec l'urbanisation se penchaient sur le dimensionnement des réseaux d'alimentation en eau et d'assainissement et sur la protection contre les inondations. L'objectif de comprendre et quantifier l'impact de l'urbanisation sur le cycle hydrologique est relativement récent. En quelque sorte, cet objectif a conduit à une évolution (1) de la focalisation sur le ruissellement vers l'inclusion d'autres processus hydrologiques, comme l'évapotranspiration, (2) de la modélisation de la réponse hydrologique uniquement en temps de pluie, dans un cadre événementiel, vers une modélisation continue, incluant les périodes sèches, et (3) de la représentation du cycle de l'eau à l'échelle résidentielle vers une représentation à l'échelle du bassin versant.

La plupart des modèles hydrologiques actuels sont conçus à l'échelle du quartier ou de la ville, et privilégient une description détaillée du paysage pour traduire le mieux possible son hétérogénéité (Salvadore, Bronders et Batelaan, 2015). On parle alors d'une approche de modélisation « bottom-up » (ou approche ascendante) : en décrivant suffisamment les processus hydrologiques à faible échelle, avec des équations dont les paramètres sont déterminés *a priori*, on espère reproduire le comportement hydrologique à l'échelle du bassin versant. Cette approche permet un lien direct entre les propriétés physiques du milieu urbanisé et la structure du modèle grâce à des paramètres mesurables. La prévision de l'impact futur est donc possible et directe : les paramètres du modèle peuvent être modifiés d'une manière qui correspond aux scénarios d'urbanisation et les simulations de modèle résultantes permettent d'estimer les impacts des aménagements projetés sur le comportement hydrologique des bassins versants. En revanche, il est souvent difficile de valider ce type d'approche sur un large échantillon de bassins versants, car la mise en disponibilité de données exigées pour développer et s'assurer de la robustesse de ce type de modèles demande beaucoup d'investissement, et n'est disponible en pratique que pour peu de cas d'étude.

L'approche inverse « top-down » (ou approche descendante) cherche des lois de comportement hydrologique à l'échelle du bassin versant (Hrachowitz et Clark, 2017). Cette recherche se fait en testant des outils conceptuels relativement simples dont l'estimation des paramètres se fait à partir d'un calage numérique. L'approche top-down présente l'avantage d'être relativement moins exigeante en données et plus simple à appliquer par rapport à l'approche bottom-up. Ceci encourage son application pour un grand nombre de cas, ce qui permet d'élargir le panel des situations climatiques et d'urbanisation sur lesquelles les modèles sont développés et testés. Pour autant, la littérature montre que la simplicité des modèles top-down n'a pas été testée intensivement sur de larges échantillons de bassins versants urbanisés (Salvadore, Bronders et Batelaan, 2015), ce qui met en question la crédibilité des modèles actuels pour prédire l'impact d'une urbanisation future sur le comportement hydrologique d'un bassin versant.

Objectif de la thèse et questions scientifiques abordées

L'objectif de cette thèse est de chercher une structure de modèle capable de reproduire la relation pluie-débit pour un large échantillon de bassins versants caractérisés par différents niveaux d'urbanisation. Comme nouveautés par rapport aux travaux existants qui portent sur le développement de modèles en milieu urbanisé, cette thèse propose d'employer un large échantillon de données au pas de temps horaire, collecté pour des bassins versants situés en France et aux États-Unis. Elle propose également de poursuivre une approche de développement pas-à-pas : en partant d'une structure de modèle initialement non dédiée aux bassins versants urbanisés, cette structure est complexifiée au fur et à mesure pour améliorer à la fois (1) la prise en compte des spécificités des bassins urbanisés, et (2) la capacité du modèle à reproduire le comportement hydrologique des bassins versants.

Le développement d'un modèle hydrologique est souvent perçu comme un travail à portée uniquement opérationnelle. En réalité, un modèle hydrologique est fondé sur un jeu d'hypothèses essayant d'expliquer la transformation des précipitations en débit observé à l'exutoire (Blöschl, 2017). Le test du modèle permet de rejeter ou de rectifier le jeu d'hypothèses duquel il découle en mesurant sa vraisemblance vis-à-vis des mesures de terrain. Cette thèse a donc pour ambition de chercher un jeu d'hypothèses qui soit valable pour un bassin versant quel que soit le niveau d'urbanisation. L'emploi d'un large échantillon est donc nécessaire pour travailler sur un panel de cas suffisamment riche. Le choix de partir d'une structure de départ non dédiée aux bassins urbanisés a été inspiré par des travaux précédents (Kjeldsen, Miller et Packman, 2013; Singh, Maheshwari et Malano, 2009). Ce choix permet de garder une continuité entre les bassins non urbanisés et les bassins urbanisés, ce qui amène à se poser plusieurs questions.

La formulation des hypothèses sous-jacentes au modèle hydrologique nécessite une étude

préalable du fonctionnement hydrologique des bassins versants urbanisés. Ceux-ci se composent *a priori* de surfaces très peu perméables, caractérisées par un taux de ruissellement élevé, et des surfaces naturelles, où une part relativement importante des précipitations s'infiltré en fonction de l'état d'humidité du sol antérieur à l'événement, et en fonction des caractéristiques de l'orage (en particulier, le volume précipité et l'intensité). **Comme l'humidité du sol joue un rôle important dans la réponse hydrologique des bassins majoritairement non urbanisés, faut-il la prendre en compte pour représenter la réponse hydrologique des bassins urbanisés? Sinon, existe-il un seuil d'urbanisation au-delà duquel seule l'imperméabilité du bassin versant prédétermine sa réponse (c'est le cas de nombreuses approches simplifiées en contexte urbain)?** Une étude événementielle permettra de regarder la variabilité temporelle (d'un événement à un autre) et spatiale (d'un bassin versant à un autre) des coefficients de ruissellement en fonction des conditions antécédentes d'humidité du sol et du taux moyen d'imperméabilité totale du bassin. Cette étude abordera le fonctionnement hydrologique d'un bassin caractérisé par une mosaïque de couverts naturel et urbain en se basant uniquement sur une analyse de données.

Ensuite, on analysera la structure du modèle de départ et on inspectera ses limites afin d'identifier les composantes susceptibles d'être modifiées pour mieux représenter les bassins urbanisés. **Notamment, est-ce que ce modèle de départ, contenant un jeu d'hypothèses déjà testé sur un large échantillon de bassins non urbanisés, mais ne représentant explicitement aucune spécificité liée à l'urbanisation, est capable de reproduire le comportement hydrologique d'un bassin fortement urbanisé?** En se servant d'un échantillon témoin de bassins versants non urbanisés, on examinera si un même jeu d'hypothèses est valable pour différentes situations d'urbanisation, ce qui justifiera par la suite le recours ou non à une modification de la structure du modèle de départ.

Supposant que le modèle de départ a montré des performances aussi bonnes sur les bassins fortement urbanisés que sur les bassins non urbanisés, **est-ce que les distributions de ses paramètres permettent de distinguer un bassin urbanisé ou non?** Si oui, on pourrait alors imaginer des relations qui permettent de faire le passage des scénarios de l'urbanisation du bassin versant aux scénarios de comportement hydrologique *via* les paramètres du modèle hydrologique de départ. **Mais comment relier les paramètres du modèle à plusieurs caractéristiques du bassin versant, en particulier la mesure de son couvert urbain?** Il s'agit d'une première tentative de prise en compte des spécificités de l'urbanisation dans la structure du modèle hydrologique de départ. Pour cela, on établira des relations entre les paramètres du modèle calés et les caractéristiques climatiques, morphologiques, lithologiques et d'occupation du sol de plusieurs bassins versants, en particulier la fraction du couvert urbain. Les performances de ces relations et leur sensibilité à la mesure du couvert urbain permettront de trancher sur le potentiel de leur utilisation pour faire le lien entre les scénarios d'urbanisation et les scénarios hydrologiques.

Une autre manière de prendre en compte les spécificités de l'urbanisation dans la structure du modèle sera d'adapter celle-ci pour mieux reproduire la relation pluie-débit pour un large échantillon de bassins versants avec différentes situations d'urbanisation. Pour ce faire, **quelles modifications faut-il effectuer sur la structure de départ?** L'analyse du fonctionnement hydrologique aura donné une idée des spécificités de la réponse hydrologique en fonction du taux moyen d'imperméabilité totale du bassin versant et de son état d'humidité du sol. L'analyse de la structure de départ aura indiqué la partie du modèle de départ où il faut introduire les modifications. On jugera de la pertinence d'une modification en fonction de l'amélioration des performances du modèle de départ sur un large échantillon de bassins versants. Des questions méthodologiques se poseront quant à la méthode à envisager pour quantifier et qualifier la significativité de l'amélioration des performances du modèle.

Organisation du mémoire

Ce mémoire est réparti en six chapitres et clôturé par une conclusion générale.

Le premier chapitre dresse un panorama bibliographique pour situer les travaux de cette thèse. Il est axé sur deux questions principales : une revue de l'impact de l'urbanisation sur les processus hydrologiques à l'échelle locale et à l'échelle du bassin versant, puis une revue des modèles hydrologiques applicables et appliqués sur les bassins versants urbanisés.

Le deuxième chapitre détaille la collecte d'un large échantillon de données hydroclimatiques, morphologiques, lithologiques et d'occupation du sol pour des bassins versants situés en France et aux États-Unis. Il décrit les méthodes de calcul de chacune des variables. L'objectif principal est de mesurer la richesse de l'échantillon pour juger de la diversité des situations hydroclimatiques et d'occupation du sol sur lesquelles le modèle sera développé.

Sous forme d'un article publié, le troisième chapitre traite des rôles de plusieurs contrôles spatiaux et temporels dans la variabilité des coefficients de ruissellement pour un large échantillon de bassins versants nord-américains. En particulier, l'importance des conditions antérieures d'humidité du sol par rapport au taux moyen d'imperméabilité totale du bassin et aux caractéristiques de l'événement pluviométrique est analysée. Ce chapitre montre également l'importance relative d'autres facteurs caractérisant le bassin versant (morphologiques, climatiques et lithologiques).

Sous forme d'un article publié, le quatrième chapitre compare les performances de deux modèles conceptuels, GR4H (Ficchi, Perrin et Andréassian, 2019) et IHACRES (Jakeman, Littlewood et Whitehead, 1990), sur deux échantillons de bassins versants : le premier échantillon contient des bassins fortement urbanisés, et le deuxième échantillon est constitué de leurs voisins

faiblement urbanisés. Une vérification de l'impact de l'urbanisation sur le comportement hydrologique est montrée grâce à des signatures hydrologiques. À la suite de l'analyse des performances, une analyse des différences en termes de paramètres et de flux internes des deux modèles sur les deux échantillons éclaircit les parties des modèles (production, routage) où l'effet de l'urbanisation se manifeste le plus.

Sous forme d'un article publié, le cinquième chapitre montre un travail de régionalisation des paramètres du modèle de départ (GR4H; Ficchi, Perrin et Andréassian, 2019) pour les lier avec plusieurs caractéristiques du bassin versant dont fait partie la fraction de son couvert urbain. Les relations de régression obtenues à l'aide de l'algorithme de *random forest* sont testées sur un échantillon de bassins versants fortement urbanisés en les comparant avec des méthodes classiques de transfert de paramètres. La sensibilité de ces relations à la mesure de la fraction du couvert urbain est également analysée.

Sous forme d'un article en révision, le sixième chapitre compare cinq modifications du modèle de départ (GR4H; Ficchi, Perrin et Andréassian, 2019) sur un large échantillon de bassins versants représentant différents niveaux d'urbanisation. Les performances correspondant à chacune des modifications sont montrées. L'analyse de la variabilité des paramètres rajoutés en fonction du taux moyen d'imperméabilité totale du bassin est effectuée.

La conclusion générale dresse un bilan des résultats obtenus et discute des futures voies de recherche.

CHAPITRE 1

MODÉLISATION HYDROLOGIQUE DES BASSINS VERSANTS URBANISÉS

Sommaire

1.1	Introduction	13
1.2	Quel est l'impact de l'urbanisation sur les processus hydrologiques?	14
1.3	Aperçu des approches de modélisation des bassins versants urbanisés . . .	40
1.4	Synthèse	60

1.1 Introduction

Il existe un grand nombre de modèles hydrologiques utilisés pour représenter le cycle de l'eau dans un milieu urbanisé. Si ces modèles diffèrent peu, en termes d'équations physiques et d'outils conceptuels, de ceux utilisés pour un milieu non urbanisé, les objectifs de leurs applications peuvent être très différents voire exclusifs aux milieux urbanisés (Fletcher, Andrieu et Hamel, 2013; Salvadore, Bronders et Batelaan, 2015). Parmi ces objectifs, la prédiction de l'impact de l'urbanisation sur le comportement hydrologique des bassins versants est une préoccupation actuelle tant sur le plan sociétal que scientifique. D'un point de vue sociétal, l'urbanisation accrue depuis la deuxième moitié du 20^{ème} siècle a engendré des impacts sur le bilan et la qualité des ressources en eau, ce que les planificateurs urbains sont obligés de prendre en compte dans leurs choix de développement urbain. Pour la communauté scientifique, l'urbanisation est un « changement », d'origine anthropique, qui se manifeste non seulement sur le paysage mais aussi sur l'hydrologie, et fait donc partie des préoccupations de la décennie 2013-2022 de l'Association Internationale des Sciences Hydrologiques « *Panta Rhei—tout s'écoule* » (Montanari *et al.*, 2013). Dans leur liste des 23 problèmes ouverts en hydrologie, Blöschl *et al.* (2019b) soulignent que bien que cette question du changement d'occupation du sol a été posée depuis des décennies, la prévision de son impact sur l'hydrologie reste largement incertaine, à cause notamment de sa dépendance à l'échelle spatiale d'étude.

Le présent chapitre sert à positionner les travaux de cette thèse dans le contexte de la prévision de l'impact de l'urbanisation sur l'hydrologie à travers l'utilisation de modèles hydrologiques. Il dresse l'état de l'art sur deux questions principales : **(1)** quel est l'impact de l'urbanisation sur les processus hydrologiques à l'échelle locale (quartier, ville) et à l'échelle du bassin versant? **(2)** Quelles approches de modélisation sont actuellement utilisées pour représenter les milieux urbanisés? Ces questions seront brièvement reprises dans les chapitres qui traitent du test des modèles sur les bassins urbanisés (Chapitres 4 et 6). D'autres questions abordées dans cette thèse sont moins présentes dans le présent chapitre mais seront revues plus tard dans l'introduction du chapitre les concernant, comme l'utilisation d'un large échantillon de bassins versants (Chapitre 2), le rôle de l'humidité du sol dans la variabilité des coefficients de ruissellement sur un bassin urbanisé (Chapitre 3), ou la régionalisation des paramètres d'un modèle hydrologique (Chapitre 5).

Chacune des deux questions énoncées est traitée dans une section indépendante (Section 1.2 pour la première question sur l'impact hydrologique de l'urbanisation, et Section 1.3 pour la seconde question sur les outils de modélisation). Nous commençons par décrire le changement opéré par l'urbanisation sur les processus hydrologiques à l'échelle locale (Section 1.2.2), puis nous analysons comment ces changements se traduisent à l'échelle du bassin versant, en ciblant trois composantes du régime hydrologique : l'écoulement moyen, l'étiage et la crue (Section 1.2.3). Ensuite, nous faisons un bilan des modèles hydrologiques développés et/ou

utilisés pour les environnements urbanisés (Section 1.3.2). Ce bilan illustrera le besoin en outils de modélisation robustes et permettant de représenter explicitement, de par la structure du modèle ou ses paramètres, le lien entre les caractéristiques urbaines du bassin versant et sa réponse hydrologique. Finalement, nous présentons la méthodologie suivie dans cette thèse pour répondre à ce besoin (Section 1.3.3).

1.2 Quel est l'impact de l'urbanisation sur les processus hydrologiques ?

Avant d'aborder les outils de modélisation utilisés en milieu urbanisé, il est nécessaire d'analyser dans quelle mesure les processus hydrologiques sont modifiés à cause de l'urbanisation. Cette revue permettra de comprendre les spécificités paysagères et hydrologiques d'un milieu urbanisé par rapport à un milieu non urbanisé, ce que les modèles hydrologiques devraient retraduire dans leurs structures/paramètres. La question d'échelle est centrale dans cette thèse ; alors que les travaux traditionnels de l'hydrologie urbaine se focalisent sur l'échelle de la ville ou du quartier, nous proposons d'analyser également l'impact de l'urbanisation sur le comportement hydrologique à l'échelle du bassin versant. Cette proposition a été poursuivie notamment dans les travaux de thèse de Salavati (2015), qui ont porté sur les méthodes de détection et de quantification de l'impact de l'urbanisation en utilisant un large échantillon de bassins versants urbanisés.

Cette section comprend trois parties. La première partie (Section 1.2.1) sert d'introduction en signalant le rôle de l'occupation du sol dans la réponse hydrologique. L'impact de l'urbanisation (comme cas particulier du changement d'occupation du sol) sur les processus hydrologiques à l'échelle locale est détaillé dans la deuxième partie (Section 1.2.2). Enfin, la troisième partie (Section 1.2.3) montre l'état des connaissances actuelles sur l'impact de l'urbanisation à l'échelle du bassin versant.

1.2.1 Qu'est-ce que la réponse d'un bassin versant ?

L'hydrologie est la science du cycle de l'eau. Elle s'emploie à décrire les flux hydriques entre les différents stocks d'eau dans un lieu donné au cours d'un temps défini. Dans son métier, l'hydrologue regarde souvent l'eau : (1) soit comme étant une ressource à gérer pour l'irrigation, l'alimentation en eau potable et industrielle, ou pour la préservation d'un écosystème aquatique, (2) soit comme un danger dont il faut se protéger dans le cas de la prévision du risque d'inondation, ou de la protection des ouvrages et des infrastructures.

Nous nous intéressons aux flux hydrologiques sur les continents, c'est-à-dire les échanges entre l'atmosphère, le sol et l'aquifère souterrain. Nous étudions les bilans d'eau sur des entités relativement indépendantes que sont les bassins versants. Un bassin versant est une surface topographique, délimitée par une ligne de crête, appelée également ligne de partage des eaux. Sur cette surface topographique, la partie des précipitations qui ne s'évapore ni s'infiltré dans le sous-sol ruisselle sur la surface pour rejoindre le cours d'eau, puis converge en un point donné de ce cours d'eau, appelé exutoire du bassin versant (Roche, Miquel et Gaume, 2012, pp. 42-47). Cette définition sous-entend que le flux observé à l'exutoire, appelé débit (exprimé en volume d'eau récolté au cours d'un intervalle de temps), ne provient que du ruissellement de surface. En réalité, une partie du débit observé à l'exutoire, désignée comme le débit de base, peut provenir de la nappe sous-jacente au cours d'eau. Cet échange peut également se manifester dans le sens inverse, et dans ce cas-là c'est la rivière qui alimente la nappe sous-jacente (Le Moine, 2008, pp. 50-52).

On peut concevoir le bassin versant comme un système soumis à un forçage climatique auquel il répond par un débit à l'exutoire, qui est fonction des propriétés physiques du bassin versant. Ces propriétés peuvent être regroupées en trois types (Gao *et al.*, 2018; McDonnell *et al.*, 2007; Poncelet, 2016, p. 22; Singh et Woolhiser, 2002) :

- La morphologie regroupe tout ce qui caractérise la forme du bassin versant, à savoir sa taille, son altitude moyenne, sa pente, et la densité de son réseau hydrographique. Elle conditionne les processus de génération du ruissellement et la dynamique du transfert des quantités ruisselantes jusqu'à l'exutoire du bassin versant (Beven, Wood et Sivapalan, 1988; Rodriguez-Iturbe et Valdés, 1979).
- La pédologie et la géologie (regroupées ci-après sous le terme *géopédologie*) comportent les caractéristiques du sol et du sous-sol, notamment la porosité, la perméabilité, la texture du sol, etc. Elles conditionnent les stocks en eau dans le sol et en souterrain, en contrôlant la variabilité spatiale de l'humidité du sol et du ruissellement (Dunne et Black, 1970; Grayson *et al.*, 1997; Merz et Blöschl, 2009). Elles impactent également les échanges de l'eau avec l'atmosphère et les échanges nappe-rivière (Brunner, Cook et Simmons, 2009; Sophocleous, 2002; Winter *et al.*, 1998).
- L'occupation du sol caractérise ce qui occupe la surface du bassin versant notamment la végétation, la forêt, les plans d'eau et les surfaces artificialisées. Au-dessus des couches géologiques et pédologiques, la couverture du sol est la première couche (du haut vers le bas) se trouvant dans la zone critique et constitue l'interface entre le sol et l'atmosphère. Elle impacte fortement la réponse d'un bassin versant à un événement en répartissant le volume précipité entre l'interception, l'infiltration, et le ruissellement (Bronstert, Niehoff et Bürger, 2002). Elle joue un rôle crucial dans le bilan hydrologique d'un bassin versant en contrôlant l'évapotranspiration (Foley *et al.*, 2005; Guimberteau *et al.*, 2017; Lian *et al.*, 2018).

Ainsi, la variabilité de la réponse du bassin versant est influencée par la variabilité à la fois des caractéristiques climatiques, morphologiques, géopédologiques et d'occupation du sol (Troch *et al.*, 2015). Les études du comportement hydrologique s'intéressent à l'impact du changement climatique (Blöschl *et al.*, 2019a; Bronstert, Niehoff et Bürger, 2002; Etchevers *et al.*, 2002; Praskievicz et Chang, 2009) comme à celui du changement d'occupation du sol, par exemple du fait de la déforestation (Andréassian, 2004) ou de l'urbanisation (Fletcher, Andrieu et Hamel, 2013; Jacobson, 2011; Li *et al.*, 2020; Oudin *et al.*, 2018). En Sciences de l'Environnement, l'étude du changement hydrologique est cruciale vu son importance comme conducteur des changements écologiques et géomorphologiques (Paul et Meyer, 2001; Poff, Bledsoe et Cuhacyan, 2006; Walsh *et al.*, 2005). Cette thèse s'intéresse à l'étude de l'urbanisation comme cause du changement de comportement hydrologique, sur lequel une revue est présentée dans les deux prochaines parties (i) au niveau local (quartier, ville), puis (ii) à l'échelle du bassin versant topographique.

1.2.2 Impact de l'urbanisation sur les processus hydrologiques à l'échelle locale

L'urbanisation est un cas particulier du changement d'occupation du sol. Elle survient en réponse à une augmentation de la densité de population qui s'accompagne de la construction de bâtiments et de routes en utilisant des matériaux généralement moins perméables à l'eau que les surfaces naturelles. Cette artificialisation des sols est la conséquence d'un besoin hygiénique d'évacuation rapide des eaux usées et pluviales pour éviter leur stagnation et d'un besoin hydraulique de protection contre les inondations (Chocat *et al.*, 2001; Carré et Deutsch, 2015, p. 22). L'utilisation de matériaux peu perméables à l'eau est justifiée par la stabilité des infrastructures, pour lesquelles les eaux d'infiltration sont nuisibles (Chocat, 1997, p. 14). Le changement d'un paysage rural à un paysage fortement urbanisé est facilement perceptible : les couleurs de la saison manifestées par la végétation et les sols cultivés font place à la couleur grise du bâti et des routes asphaltées.

En revanche, il n'est pas aussi facile d'appréhender le changement qui s'opère dans le comportement hydrologique d'un bassin versant qui s'urbanise. Commençons par regarder l'impact sur les principaux processus hydrologiques intervenant dans la transformation des précipitations en débit à l'exutoire.

1.2.2.1 Les précipitations

À cause de l'artificialisation des sols, l'urbanisation modifie le bilan énergétique à la surface en introduisant des matériaux ayant des propriétés thermiques différentes des surfaces naturelles,

notamment l'albédo et l'emmagasinement de chaleur. Cette modification contribue à la création du phénomène de l'îlot de chaleur urbain (Oke, 1982) qui se matérialise par des différences de température entre la ville et les surfaces rurales qui l'entourent. Les différences d'humidité du sol entre les surfaces urbaines et rurales impactent les échanges de chaleur latente (Dixon et Mote, 2003; Husain, Bélair et Leroyer, 2014). À ces perturbations du bilan énergétique s'ajoute le changement de rugosité de surface à cause des constructions, ce qui déplace la couche limite atmosphérique et affecte entre autres le vent et la formation des nuages (Dou *et al.*, 2015; Shepherd, 2005). Les émissions importantes d'aérosols du fait de la densité des activités anthropiques en ville (comme le transport) impactent la composition atmosphérique en particules, ce qui modifie le bilan énergétique et la formation des nuages (Jin, Shepherd et King, 2005).

La revue récente de Liu et Niyogi (2019) a constaté une augmentation des précipitations de 14% à 22% à cause de l'urbanisation surtout en aval de la ville (suivant la direction du vent) et de 11% à 21% au-dessus de la ville. Toutefois, les villes figurant dans cette revue sont majoritairement situées en Chine et aux États-Unis, et aucune ville européenne n'y est prise en compte. Dans une moindre mesure, l'augmentation des précipitations a été également observée en bordure de la ville, ce qui est expliqué par le changement de la direction des tempêtes lorsqu'elles rencontrent les barrières constituées par les bâtiments (Dou *et al.*, 2015). L'impact de l'urbanisation sur les précipitations est dépendant de la saison été/hiver et du contraste jour/nuit, avec un effet plus important sur les événements estivaux (Changnon, Shealy et Scott, 1991; Shepherd, Pierce et Negri, 2002) et les événements ayant lieu pendant le jour, à cause de l'importance de l'effet de l'îlot de chaleur urbain pendant le jour et la faiblesse des interactions sol-atmosphère pendant la nuit (Niyogi *et al.*, 2011). La complexité du phénomène de la formation des précipitations rend difficile la classification des mécanismes par lesquels l'urbanisation impacte les précipitations, ce qui empêche d'en cerner les plus importants (Shepherd, 2005).

1.2.2.2 L'évapotranspiration et l'interception

À côté des précipitations, l'évapotranspiration constitue le deuxième forçage climatique le plus important dans le bilan hydrologique. Nous considérons l'interception avec l'évapotranspiration sous l'hypothèse que toute l'eau interceptée s'évapore.

L'urbanisation se traduit par un remplacement des sols naturels et du couvert végétal par des surfaces peu perméables. Ceci se traduit par une diminution du taux de la chaleur latente en faveur d'une augmentation du taux de la chaleur sensible (Weng et Lu, 2008), entraînant ainsi une diminution de l'évapotranspiration. Cet effet est d'autant plus important qu'il y a un drainage rapide des eaux de ruissellement, ce qui équivaut à une diminution de l'eau disponible pour l'évapotranspiration (Grimmond et Oke, 1991; Taha, 1997).

Notons cependant que dans certains cas, ce schéma général de diminution de l'évapotranspiration n'a pas lieu ou est contrebalancé par d'autres facteurs. Par la création de l'îlot de chaleur urbain (décrit dans la partie précédente), l'urbanisation modifie les conditions climatiques locales, ce qui se traduit par une augmentation de la température en ville par rapport aux surfaces rurales environnantes (Oke, 1982 ; Santamouris, 2015). Ceci peut contribuer à l'augmentation de l'évaporation du sol, dont la source d'humidité peut être dans certains cas assurée localement par les fuites récurrentes dans les réseaux d'eau potable (Xu *et al.*, 2014). En outre, l'augmentation de la température conjuguée à la sur-irrigation des espaces verts urbains contribue à l'augmentation de la transpiration (Salvador, Bautista-Capetillo et Playán, 2011).

Les toitures des bâtiments interceptent une large proportion des précipitations selon leur pente, leur orientation et leur hauteur. L'interception peut alors dépasser 90% des précipitations, ce qui permet d'avoir des taux d'évaporation importants variant de 9% à 56% des précipitations annuelles (Ragab *et al.*, 2003a), avec des taux plus importants sur les toitures ayant une faible pente. L'évaporation peut compter pour une bonne proportion du bilan hydrologique sur les routes et les chaussées (Ragab *et al.*, 2003b ; Ramier, Berthier et Andrieu, 2011), atteignant jusqu'à ~25% des précipitations (Ragab *et al.*, 2003b ; Ramier, Berthier et Andrieu, 2004).

1.2.2.3 Le ruissellement et l'infiltration

Le ruissellement est peut-être le processus le plus étudié et le plus associé aux surfaces urbaines, à cause de l'imperméabilisation des surfaces (Arnold et Gibbons, 1996). Cette imperméabilisation vient de l'utilisation de matériaux garantissant la qualité des infrastructures urbaines (comme l'asphalte et le béton), en particulier leur étanchéité contre toute infiltration de l'eau qui peut compromettre leur stabilité et leur durabilité.

En comparaison avec les surfaces naturelles, le stockage dépressionnaire est très faible sur les surfaces urbaines (Salvadore, Bronders et Batelaan, 2015). Les plus importants taux de ruissellement sont enregistrés sur les toitures, et varient de 30% jusqu'à 90%, selon le mois de l'année (Hollis et Ovenden, 1988 ; Ragab *et al.*, 2003a). Les taux de ruissellement sur les chaussées et les routes sont également très variables : ils peuvent dépasser 90% des précipitations, comme ils peuvent se limiter à moins de 10% (Mansell et Rollet, 2006 ; Ramier, Berthier et Andrieu, 2004). Les mesures *in situ* montrent des taux moyens plutôt faibles, atteignant en moyenne ~10% pour les faibles événements et ~30% pour les événements à fortes précipitations (Hollis et Ovenden, 1988). En outre, les sols urbains étant plus compacts que les sols naturels, ils sont caractérisés par une faible porosité et par conséquent une faible capacité de rétention d'eau (Gregory *et al.*, 2006 ; Jim, 1998), ce qui réduit leur taux d'infiltration et favorise le ruissellement.

Si l'urbanisation contribue à la réduction de l'infiltration (en privilégiant le ruissellement),

elle ne la supprime pas complètement. Même avec un fort taux de ruissellement, les toitures peuvent ne pas être connectées au réseau de drainage, et donc déverser l'eau collectée dans le périmètre du bâtiment, ce qui pourrait contribuer localement à la recharge des aquifères et à l'humidité des sols (Redfern *et al.*, 2016). Comme le ruissellement, l'infiltration sur les chaussées et les routes dépend du type de matériau et de l'âge de la structure. En vieillissant, ces structures deviennent altérées, ce qui engendre des fissures permettant des chemins préférentiels pour l'infiltration (Fernández-Barrera *et al.*, 2008; Wiles et Sharp, 2008). Le type de matériau de la chaussée conditionne le taux d'infiltration, avec un taux plus élevé dans des chaussées en pierre concassée et en pavés que dans des chaussées en asphalte (Gilbert et Clausen, 2006). En particulier, le taux d'infiltration sur les pavés peut atteindre 30% des précipitations (Ramier, Berthier et Andrieu, 2011). Pour les sols urbains, l'implantation des arbres permet, grâce à leurs racines, de créer des macropores et des chemins préférentiels et donc restaurer l'infiltration (Bartens *et al.*, 2008).

1.2.2.4 Le transfert et la dynamique du ruissellement

En plus de leur imperméabilité, les surfaces urbaines sont caractérisées par une rugosité plus faible que les surfaces naturelles. Par conséquent, la vitesse d'écoulement est généralement plus grande, réduisant le temps de parcours (Huang *et al.*, 2008b; Paul et Meyer, 2001; Rodriguez, Andrieu et Creutin, 2003). Les réseaux de collecte d'eaux pluviales accentuent cet effet en permettant un routage plus rapide vers l'exutoire, ce qui se manifeste par des vitesses importantes dans ces canalisations (elles peuvent aller jusqu'à 2.5 m/s, voir par exemple les résultats expérimentaux de Larrarte, 2006). L'imperméabilisation des berges du cours d'eau naturel réduit la rugosité du lit, ce qui entraîne un transfert plus rapide dans le réseau hydrographique (Walsh *et al.*, 2005). De surcroît, l'imperméabilisation des versants, en plus des grandes quantités d'eau arrivant avec des vitesses importantes, a un pouvoir érosif accru et change la géomorphologie des cours d'eau naturels, qui s'ajustent par élargissement ou incision (Konrad, Booth et Burges, 2005; Navratil *et al.*, 2013; Paul et Meyer, 2001). Ceci contribue à l'instabilité du lit de la rivière (vis-à-vis de l'érosion) et réduit la variabilité spatiale de ses attributs géomorphologiques (Anim *et al.*, 2018; Vietz *et al.*, 2014).

1.2.2.5 La recharge souterraine

L'urbanisation met en place des processus directs et indirects qui contribuent selon les cas à augmenter ou à diminuer la recharge souterraine (Lee *et al.*, 2005; Lerner, 1990; Salvatore, Bronders et Batelaan, 2015).

D'une façon générale, la recharge souterraine diminue à cause de la diminution de l'infiltration.

Avec l'importance des ressources souterraines pour pourvoir la ville en eau, le pompage contribue également à l'abattement de la nappe (Foster, 2020; Foster et Chilton, 2003; Zektser, Loáiciga et Wolf, 2005). De plus, avec une remontée suffisante de la nappe pour rentrer en contact avec les conduites d'assainissement, un drainage des eaux souterraines peut se mettre en place *via* les joints (Braud *et al.*, 2013; Rodriguez *et al.*, 2020). En outre, la construction des infrastructures en milieu urbanisé peut exiger une diminution du niveau de la nappe (Kim, Lee et Sung, 2001).

En revanche, la diminution de la recharge peut être compensée par la diminution de l'évapotranspiration réelle (par réduction du couvert végétal), par l'irrigation des espaces verts urbains ou par la mise en place de techniques alternatives (voir le récapitulatif ci-dessous) permettant de ré-infiltrer les volumes ruisselés suite aux événements pluvieux (Göbel *et al.*, 2004; Hamel, Daly et Fletcher, 2013; Shuster, Gehring et Gerken, 2007). La recharge souterraine bénéficie également des fuites d'eau qui arrivent souvent dans les canalisations d'eau potable (Appleyard, 1995; Kim, Lee et Sung, 2001; Xu *et al.*, 2014; Lerner, 2002; Yang *et al.*, 1999), voire dans les réseaux d'assainissement (Wakode *et al.*, 2018). Des puits d'injection peuvent également être utilisés pour garantir une recharge artificielle (Lerner, 1990).

Les modifications qui s'opèrent dans le stock d'eau souterraine se conjuguent aux aménagements des rivières pour altérer les interactions nappe-rivière (Ellis, Mackay et Rivett, 2007). Ceci produit un impact local sur la direction d'échange et sur la zone hyporhéique où s'effectue le mélange entre les eaux de surface et les eaux souterraines (Hancock, 2002).

1.2.2.6 Récapitulatif : impact de l'urbanisation à l'échelle locale et mesures d'atténuation

Nous constatons que plusieurs études pointent du doigt les modifications causées par l'urbanisation sur la plupart (si ce n'est la totalité) des processus hydrologiques (Tableau 1.1). La modification du climat local contribue à une augmentation des précipitations. Ce changement climatique local s'accompagne d'une augmentation des températures, ce qui accroît la demande atmosphérique en eau. Ainsi, selon l'humidité du sol et la densité de la canopée végétale, l'évapotranspiration réelle peut augmenter (du fait de l'augmentation de la température et de l'irrigation du couvert végétal, ou par l'évaporation des eaux interceptées) ou diminuer (à cause de la réduction du couvert végétal). En arrivant sur les surfaces urbaines qui sont caractérisées par des taux d'imperméabilité importants, le ruissellement augmente et l'infiltration diminue. Les sols urbains favorisent le ruissellement par rapport aux sols naturels à cause de leur compaction. La diminution de la rugosité des surfaces urbaines, ainsi que le transfert rapide *via* les réseaux d'assainissement, accélèrent l'évacuation des eaux ruisselées. L'échange avec la nappe est impacté localement par les aménagements. Le niveau de la nappe est soit augmenté par la diminution de l'évapotranspiration ou par les fuites d'eau dans les canalisations, soit diminué par le drainage des conduites d'assainissement ou par le pompage.

Toutefois, il faut noter que si pour certains processus les conclusions sont plutôt convergentes, une grande variabilité caractérise l'ampleur des modifications induites par l'urbanisation, comme pour le taux d'augmentation des précipitations ou pour le taux de ruissellement sur les toitures et les chaussées. Concernant d'autres processus comme l'évapotranspiration ou la recharge souterraine, les conclusions sont plutôt contradictoires, ce qui est dû en partie à la nature double de l'impact de l'urbanisation causant des modifications dans les deux sens.

L'impact de l'urbanisation à l'échelle locale est difficilement généralisable car il dépend aussi des mesures compensatoires qui peuvent être déployées pour gérer l'eau ruisselante à la source. Désignées par « techniques alternatives » (Ahiablame, Engel et Chaubey, 2012; Eckart, McPhee et Bolisetti, 2017; Fletcher *et al.*, 2015; Hamel, Daly et Fletcher, 2013), ces mesures ont pour objectifs (i) d'améliorer la qualité des eaux drainées, et (ii) de neutraliser, hydrologiquement, la présence de surfaces imperméables drainées par un réseau d'assainissement. Ainsi, l'utilisation des techniques alternatives ambitionne « un retour des eaux ruisselantes au cycle hydrologique naturel » (Eckart, McPhee et Bolisetti, 2017). Ces techniques peuvent être classées en deux groupes (Fletcher *et al.*, 2015) :

- Techniques basées sur l'infiltration, qui permettent de restaurer l'infiltration au milieu urbain et d'augmenter la recharge souterraine. Ces techniques peuvent inclure les noues, les bassins de rétention, les chaussées poreuses et les tranchées filtrantes.
- Techniques basées sur la rétention, qui contribuent à retenir les volumes d'eau ruisselante pour atténuer le pic de débit et réduire le ruissellement en favorisant l'évapotranspiration. Parmi ces techniques, on peut citer les toitures végétalisées, les réservoirs d'eau de pluie et les bassins de rétention.

Ces techniques sont utilisées comme solutions alternatives permettant de dépasser les limites du schéma d'assainissement traditionnel en répartissant les rejets d'eau dans l'espace et dans le temps. Si plusieurs études ont montré leur efficacité (Ahiablame, Engel et Chaubey, 2012; Chui et Trinh, 2016; Eckart, McPhee et Bolisetti, 2017; Fletcher *et al.*, 2015; Versini *et al.*, 2015), leurs performances sont assez variables et impactées par plusieurs facteurs. Puisque ces techniques ciblent d'augmenter l'infiltration et l'évapotranspiration à une faible échelle (< 1 km²), leur efficacité dépend fortement de leur densité et de la localisation de leur implantation. Le rôle des caractéristiques géomorphologiques « naturelles » (type du sol et de végétation, forçages climatiques, topographie locale, etc.) et de leurs capacités infiltrantes et évaporantes est également important. De plus, les performances des techniques alternatives varient d'un événement à un autre, selon l'intensité, le volume et le délai au pic, ce qui résulte en une dépendance au régime climatique (tempéré, tropical, etc.). Enfin, ces techniques sont dimensionnées en général avec une prise en compte des enregistrements du climat passé, ce qui ne garantit pas leur résilience face aux changements climatiques futurs.

Le Tableau 1.1 dresse un récapitulatif de l'impact moyen (ou plausible) de l'urbanisation sur les processus hydrologiques à l'échelle locale, ainsi que les incertitudes à prendre en compte. On remarque qu'avec la multiplicité des sources d'incertitudes, aucun processus n'est définitivement supprimé par l'urbanisation. D'un point de vue hydrologique, la frontière qui sépare les surfaces urbaines des surfaces naturelles devient floue et est moins nette que ce que l'on pense, ce qui pousse à considérer la distinction urbain/rural uniquement comme deux extrémités d'un continuum de surfaces plus ou moins imperméables (Brabec, Schulte et Richards, [2002](#); McGrane, [2016](#)).

Tableau 1.1. Récapitulatif de l'impact de l'urbanisation sur les processus hydrologiques à l'échelle locale (ville, quartier).

Processus	Impact « moyen » de l'urbanisation	Causes	Incertitudes ou mesures compensatoires
Précipitations	Augmentation à l'aval et sur la ville	<ul style="list-style-type: none"> • Phénomène de l'îlot de chaleur urbain • Changement de rugosité • Diminution de l'humidité du sol • Émissions importantes d'aérosols 	<ul style="list-style-type: none"> • Incertitudes importantes sur les taux d'augmentation des précipitations • Difficulté dans la hiérarchisation des processus pour en définir les plus importants

Suite à la page suivante...

Tableau 1.1 suite : Récapitulatif de l'impact de l'urbanisation sur les processus hydrologiques à l'échelle locale.

Processus	Impact « moyen » de l'urbanisation	Causes	Incertitudes ou mesures compensatoires
Évapotranspiration	Diminution	<ul style="list-style-type: none"> • Réduction du couvert végétal • Diminution de l'humidité du sol par drainage rapide des eaux ruisselantes 	<ul style="list-style-type: none"> • Augmentation à travers l'évaporation des eaux interceptées par les toitures, dont le taux d'interception est variable selon la pente et l'orientation de la toiture • Évaporation significativement non nulle sur les chaussées et les routes • Augmentation de la température à cause de l'îlot de chaleur urbain, conjuguée à une disponibilité d'humidité du sol suffisante grâce à la sur-irrigation des espaces verts et aux fuites d'eau locales • Augmentation de l'évapotranspiration par l'utilisation des techniques alternatives, comme les bassins de rétention et les toitures végétalisées

Suite à la page suivante...

Tableau 1.1 suite : Récapitulatif de l'impact de l'urbanisation sur les processus hydrologiques à l'échelle locale.

Processus	Impact « moyen » de l'urbanisation	Causes	Incertitudes ou mesures compensatoires
Ruissellement et infiltration	Augmentation du ruissellement et diminution de l'infiltration	<ul style="list-style-type: none"> • Utilisation de matériaux étanches qui empêchent l'infiltration des eaux dans les toitures, les routes et les chaussées • Compactage des sols urbains 	<ul style="list-style-type: none"> • Taux de ruissellement très variables selon le type de la construction (toiture, route), son âge, le type du matériau et la saison • Infiltration non nulle sur les surfaces urbaines supposées très imperméables, à cause notamment de leur vieillissement • Impact dépendant de la connectivité des surfaces urbaines au réseau de drainage • Restauration de l'infiltration par l'utilisation des techniques alternatives, comme les tranchées drainantes, les noues, les chaussées poreuses et les bassins d'infiltration • Restauration de la capacité d'infiltration dans les sols urbains grâce à l'implantation des arbres et la création d'espaces verts

Suite à la page suivante...

Tableau 1.1 suite : Récapitulatif de l'impact de l'urbanisation sur les processus hydrologiques à l'échelle locale.

Processus	Impact « moyen » de l'urbanisation	Causes	Incertitudes ou mesures compensatoires
Dynamique de l'écoulement	Accélération	<ul style="list-style-type: none"> • Remplacement des sols naturels et végétalisés par des surfaces à faible rugosité, donc à des vitesses d'écoulement importantes • Présence de canaux de drainage artificiel • Augmentation des volumes d'eau ruisselante 	<ul style="list-style-type: none"> • Impact dépendant de la connectivité des surfaces urbaines aux réseaux de drainage artificiels • Utilisation des techniques alternatives pour atténuer le pic de débit et ralentir la propagation de crue
Recharge souterraine	Diminution	<ul style="list-style-type: none"> • Diminution de l'infiltration • Pompage • Drainage par les réseaux d'assainissement 	<ul style="list-style-type: none"> • Augmentation grâce à la restauration de l'infiltration par des techniques alternatives, aux fuites dans les réseaux artificiels et à la diminution de l'évapotranspiration • Abandon de pompage dans les villes

Fin du Tableau 1.1.

Ayant décrit les modifications des différents processus à l'échelle locale, comment se traduisent-elles dans la relation pluie-débit à l'échelle du bassin versant ?

1.2.3 Impact de l'urbanisation sur le comportement hydrologique à l'échelle du bassin versant

Dans la suite du manuscrit, l'intérêt est porté uniquement sur l'impact de l'urbanisation sur le débit observé à l'exutoire d'un bassin versant, ce qui nous permet de caractériser la réponse hydrologique à cette échelle qui peut aller de quelques km² à quelques milliers de km². Si nous supposons un état initial faiblement urbanisé, comment changent les caractéristiques de la réponse du bassin versant lorsqu'il devient urbanisé? Nous commençons par expliquer pourquoi l'appréciation de l'impact hydrologique à l'échelle locale seulement ne suffit pas pour prédire l'impact à l'échelle d'intérêt (Section 1.2.3.1). S'ensuit une revue bibliographique des changements dans le régime hydrologique dus à l'urbanisation, en focalisant sur les forts débits, sur les faibles débits, sur l'écoulement moyen et sur la dynamique de la réponse du bassin versant (Section 1.2.3.2). Cette revue est complétée par une discussion des méthodes de détection et d'attribution du changement hydrologique du fait de l'urbanisation.

1.2.3.1 Suffit-il de connaître l'impact de l'urbanisation sur les processus à l'échelle locale pour prédire l'impact sur la relation pluie-débit à l'échelle du bassin versant ?

Les divers processus hydrologiques rencontrés en milieu urbanisé sont caractérisés par une diversité d'échelles spatiale et temporelle (Figure 1.1), allant de quelques millimètres jusqu'à des dizaines de milliers de kilomètres en espace, et de quelques minutes jusqu'à des centaines d'années en temps. Si l'on regarde les processus induits par l'urbanisation (assainissement et collecte des eaux pluviales), ils sont caractérisés par de fines échelles, en particulier plus faibles que celle d'un bassin versant (supposée supérieure à 1 km²). Ceci montre que selon la taille du bassin étudié, l'effet de l'urbanisation sur un ou plusieurs processus pourrait ne pas se manifester à cette échelle. On rejoint alors l'hypothèse de Blöschl *et al.* (2007) qui stipule que l'impact du changement dépend de la taille du bassin ; plus le bassin est grand, moindre est l'effet du changement d'occupation du sol sur son comportement hydrologique.

L'impact de l'urbanisation sur les processus hydrologiques a cette particularité d'être local, alors que l'urbanisation induit également d'autres processus *diffus* dont il est difficile de se rendre compte en se limitant à l'échelle locale. Si l'on considère le bilan hydrologique pour un bassin non urbanisé et non influencé par aucune activité anthropique (comme par exemple

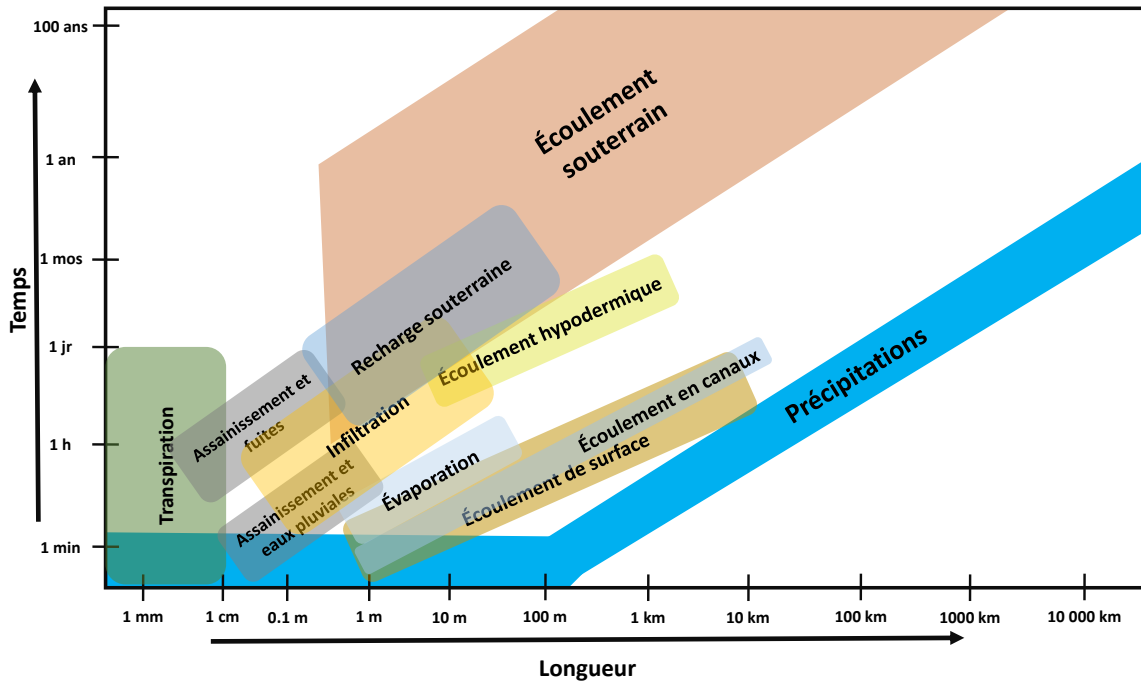


Figure 1.1. Échelle spatiale et temporelle des processus hydrologiques en milieu urbanisé. Cette figure est reprise de Salvadore, Bronders et Batelaan (2015). On définit l'échelle spatiale ou temporelle caractéristique comme la région dans l'espace ou dans le temps dans laquelle le processus prend place, où ses flux varient, ou alors la résolution spatiale ou temporelle avec laquelle il est le mieux mesuré.

les ouvrages de régulation et les barrages), l'équation peut s'écrire :

$$\frac{dS}{dt} = P - ETR - Q - Ech_{souterrain} \quad (1.2.1)$$

avec $\frac{dS}{dt}$ la variation du stock du bassin versant, matérialisé par l'humidité du sol, le contenu en eau dans le réseau hydrographique et le niveau piézométrique de la nappe en communication avec le cours d'eau. L'entrée majeure du système est P , le flux de précipitations (en unité de volume par unité de temps). Les sorties sont ETR , l'évapotranspiration réelle et Q le débit mesuré à l'exutoire. Le terme $Ech_{souterrain}$ représente l'échange avec des aquifères profonds et avec d'autres bassins versants à travers les frontières topographiques (Le Moine, 2008). Ce terme peut agir comme une entrée en cas d'apports, ou comme une sortie en cas d'exports.

D'autres flux sont à rajouter pour compléter cette équation suite à la présence d'une ville au sein du bassin versant, notamment les quantités d'eau acheminées d'autres sources externes au système, désignées sous le nom de « transferts interbassins ». Ces transferts comprennent également le pompage des eaux de nappes profondes (Burns *et al.*, 2005). Ils peuvent être très importants par rapport aux entrées *naturelles*; à titre d'exemple, les habitants de la métropole du Grand Paris consomment annuellement une quantité d'eau équivalente à la moyenne

annuelle de pluie brute sur Paris¹. Ces apports sont considérés comme externes à la ville (captage superficiel dans la Seine et la Marne, et captage souterrain dans un rayon de 80 à 150 km autour de la ville), mais ils peuvent être considérés internes au système si on se positionne à l'échelle du bassin de la Seine. Les apports externes sont donc d'autant plus importants dans l'équation du bilan que les bassins versants sont de petite taille. Pour les prendre en compte, l'équation se réécrit :

$$\frac{dS}{dt} = P + AExt - ETR - Q - Ech_{souterrain} \quad (1.2.2)$$

où $AExt$ désigne l'eau acheminée depuis l'extérieur du bassin.

Après consommation, ces quantités d'eau sont collectées dans les réseaux, et rejetées dans le cours d'eau soit directement, soit à travers une station d'épuration. Ainsi, ces effluents que l'on notera Q_{rej} devraient être intégrés dans cette équation, qui devient :

$$\begin{aligned} \frac{dS}{dt} &= P + AExt - ETR - Q - Q_{rej} - Ech_{souterrain} \\ &= P + AExt - ETR - Q_{tot} - Ech_{souterrain} \end{aligned} \quad (1.2.3)$$

Par conséquent, la réponse désormais observée à l'exutoire Q_{tot} provient de la réponse des surfaces aux événements pluviométriques, de la nappe et des rejets de la consommation des agglomérations. Comme pour les apports externes, l'importance des rejets de la consommation de l'agglomération varie selon les limites du bassin versant vis-à-vis de la localisation des rejets de la station d'épuration. Si le bassin versant est suffisamment grand pour que ces échanges se passent à l'intérieur de ses limites, les apports $AExt$ sont neutralisés par le débit rejeté Q_{rej} . On constate que la fermeture du bilan devient moins difficile quand on se positionne à l'échelle du bassin versant. Un tel positionnement permet de percevoir le bassin versant urbanisé comme montré sur la Figure 1.2.

En faveur d'un positionnement à cette échelle du bassin versant, on peut avancer un autre argument concernant le comportement même des surfaces urbaines. La faible perméabilité est une caractéristique de ces surfaces. Néanmoins, la revue de l'impact de l'urbanisation (voir la Section 1.2.2.6) sur les différents processus hydrologiques indique que le comportement

¹Le volume d'eau total facturé en 2015 s'élève jusqu'à $449 \cdot 10^6$ m³, sur une superficie de 814 km². La lame d'eau équivalente est d'environ 550 mm, à comparer avec les 600 mm qui tombent en moyenne chaque année sur la région parisienne. Voir le rapport de la Cour des Comptes de février 2018, page 446, consultable sur <https://www.ccomptes.fr/sites/default/files/2018-01/12-alimentation-eau-potable-metropole-Grand-Paris-Tome-1.pdf>.

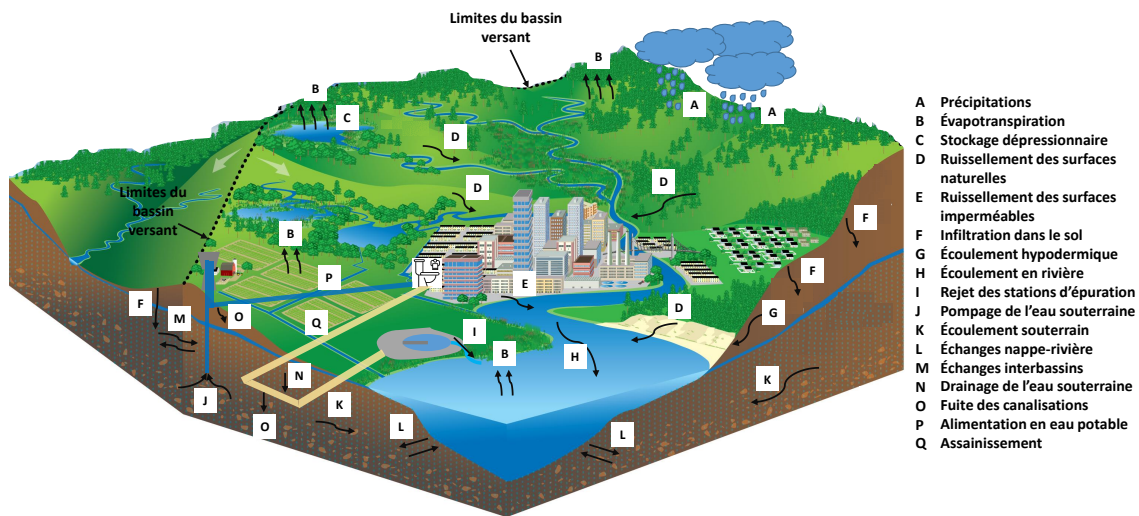


Figure 1.2. Représentation d'un bassin versant urbanisé et des principaux processus hydrologiques, traduite de Salvadore, Bronders et Batelaan (2015). L'urbanisation modifie localement le bilan hydrologique à la surface et en souterrain, et rajoute de nouveaux processus, comme l'alimentation en eau potable et l'assainissement (adaptée de <https://www.michiganseagrant.org/lessons/lessons/by-broad-concept/earth-science/exploring-watersheds/>).

des surfaces urbaines est variable (Fernández-Barrera *et al.*, 2008 ; Gilbert et Clausen, 2006 ; Redfern *et al.*, 2016) :

- dans l'espace, avec la diversité des types de surfaces urbaines qui sont caractérisées par des taux de ruissellement très variables entre une toiture, une chaussée en asphalte poreux, un pavé et un sol urbain ;
- dans le temps, avec un comportement des structures urbaines dépendant de leur âge, de la saison et du forçage climatique.

À cela se rajoutent l'organisation spatiale des surfaces imperméables et leur localisation dans le bassin versant par rapport aux surfaces rurales, ce qui s'avère d'une influence importante sur la forme de sa réponse (Grove, Harbor et Engel, 1998 ; Mejía et Moglen, 2010 ; Sheeder, Ross et Carlson, 2002 ; Shuster *et al.*, 2005 ; Zhang et Shuster, 2014). Ainsi, la fragmentation du paysage urbain peut contribuer à atténuer voire masquer l'impact de l'imperméabilité du bassin (McMahon *et al.*, 2003 ; Oudin *et al.*, 2018). En effet, même si un taux de ruissellement élevé peut s'enregistrer sur une surface imperméable, il peut être atténué par une infiltration favorisée si l'eau rencontre sur son chemin des surfaces plus perméables à l'aval avant d'atteindre le réseau hydrographique. En outre, le taux de connectivité des zones imperméables au réseau de drainage artificiel permet à une partie du bassin de ne pas être « effectivement » imperméable (Booth et Jackson, 1997 ; Boyd, Bufill et Knee, 1993 ; Lee et Heaney, 2003 ; Roy et Shuster, 2009).

En conclusion, l'appréhension de l'impact de l'urbanisation à l'échelle du bassin versant à partir d'une connaissance à l'échelle locale pose problème (comme le signale d'ailleurs la revue de McGrane, 2016). Il est nécessaire d'étudier les processus à l'échelle du bassin versant pour comprendre le lien entre l'urbanisation et le changement du comportement hydrologique à cette échelle.

1.2.3.2 Attribution du changement hydrologique à l'urbanisation à l'échelle du bassin versant

Dans la partie précédente, nous avons discuté des différences d'échelle qui peuvent impacter l'appréciation et la compréhension de l'impact de l'urbanisation sur les processus hydrologiques. Ainsi, la connaissance de l'impact de l'urbanisation à l'échelle locale (ville, quartier) seule ne permet pas de prévoir de manière fiable l'impact à l'échelle du bassin versant topographique. Pour cela, il est nécessaire de se positionner à cette échelle du bassin et relier l'urbanisation avec le changement de son comportement. Trois choix sont alors nécessaires : (1) **choisir une ou plusieurs mesures d'urbanisation**, (2) **choisir un ou plusieurs indicateurs (ou signatures) hydrologiques** en fonction des aspects visés, et (3) **choisir une méthode d'attribution** en comparant l'évolution de la mesure d'urbanisation avec celle des indicateurs hydrologiques.

Choix de la mesure d'urbanisation. Il est relativement simple de constater si un bassin est urbanisé car l'urbanisation s'accompagne d'une modification visible du paysage. Pour une étude quantitative, il faut choisir une mesure qui permette de caractériser cette urbanisation et suivre l'évolution de cette mesure. Les plus utilisées sont la densité de population (Brandes, Cavallo et Nilson, 2005 ; Chen, Chuang et Cheng, 2014 ; DeWalle *et al.*, 2000), la proportion du bassin versant occupée par les zones urbaines (Bayliss *et al.*, 2006 ; Kjeldsen, 2010 ; McMahan *et al.*, 2003 ; Miller et Grebby, 2014 ; Oudin *et al.*, 2018), ou le taux moyen d'imperméabilité totale du bassin versant (Arnold et Gibbons, 1996 ; Booth *et al.*, 2004 ; Guan, Sillanpää et Koivusalo, 2016). L'avantage de la densité de population est qu'elle est généralement disponible sur une longue période de temps. Par contre, elle n'est pas vraiment pertinente d'un point de vue hydrologique car différents schémas d'aménagement peuvent être attribués à une même densité de population (à cause, par exemple, des différences dans la hauteur de bâtiments). La proportion du bassin versant occupée par les zones urbaines est liée à l'imperméabilité totale moyenne du bassin versant : la première mesure utilise les classes d'urbanisation, alors que la deuxième mesure utilise la valeur d'imperméabilité attribuée à chacune de ces classes (voir par exemple la base de données *National Land Cover Database* (NLCD) développée aux États-Unis, où l'on définit quatre sous-classes d'urbanisation, correspondant à quatre plages d'imperméabilité : 0%-20%, 20%-50%, 50%-80% et 80%-100%, voir Homer *et al.*, 2007). L'inconvénient de ces deux mesures est leur dépendance (en pratique) à des

produits satellitaires, dont la profondeur temporelle n'est pas aussi importante que celle de la densité de population. Pour pallier ce problème, on peut recourir à une analyse de trajectoire pour extrapoler des valeurs vers le passé par recoupement avec d'autres données disponibles, ou simplement à travers des courbes empiriques (Oudin *et al.*, 2018; Verbeiren *et al.*, 2013). Ces métriques ne sont que des moments de premier ordre, c'est-à-dire une quantification de l'étendue de l'urbanisation dans le bassin versant. D'autres mesures de second ordre permettent de distinguer la connectivité des surfaces imperméables au réseau de drainage (comme le taux moyen d'imperméabilité effective, voir Ebrahimian, Wilson et Gulliver, 2016), ou la fragmentation du paysage urbain (Alberti *et al.*, 2007; Buyantuyev, Wu et Gries, 2010; Gires *et al.*, 2017; Oudin *et al.*, 2018). Puisque l'impact dépend de la juxtaposition spatiale des surfaces urbaines et naturelles au sein du bassin versant, ces métriques seraient d'une grande importance pour une meilleure discrimination entre les bassins. En revanche, il est toujours difficile de mettre en place une méthode permettant de les quantifier et de les valider d'un point de vue hydrologique, c'est-à-dire vis-à-vis de leur impact sur la réponse hydrologique.

Choix de la caractéristique de la réponse hydrologique à analyser. Le deuxième choix se fait sur l'indicateur hydrologique qui caractérise la réponse hydrologique et se trouve d'une grande importance pratique. Quatre aspects sont communément analysés (Salavati, 2015, pp. 13-18) :

- Les « hauts débits » ou débits de crue qui permettent d'analyser le comportement du bassin versant en réponse à un événement pluviométrique et qui ont des répercussions fortes au travers des inondations, de la destruction d'ouvrages et des pertes humaines. En milieu urbanisé, le risque est d'autant plus important du fait de la densité des enjeux en proximité des cours d'eau urbains tant en population qu'en patrimoine industriel et commercial (Jha, Bloch et Lamond, 2012, p. 21; Wheeler et Evans, 2009). Ce risque est généralisable à toute la ville en cas de débordement des réseaux d'évacuation et d'inondations locales. Les hauts débits jouent un rôle important dans la morphologie des cours d'eau, ce qui impacte indirectement l'habitat (Bunn et Arthington, 2002). On prend souvent le débit de pointe (total) ou le débit direct (après filtrage du débit de base) comme indicateurs (Blum *et al.*, 2020; Brandes, Cavallo et Nilson, 2005; Miller et Hess, 2017), ce qui nécessite de faire un choix de la méthode de séparation des événements. D'autres métriques se basent sur une analyse fréquentielle de toute la série des débits enregistrés, en choisissant les hautes valeurs de débits à faible probabilité de dépassement (Oudin *et al.*, 2018), ou le nombre de jours par mois pendant lesquels le débit est au-dessus d'un certain seuil (Diem, Hill et Milligan, 2018). De nombreuses études confirment que l'urbanisation accentue les débits forts (Tableau 1.2), à cause de l'augmentation du ruissellement et de la diminution de l'infiltration, accompagnées de vitesses d'écoulement importantes.

- Les « bas débits » qui permettent d'analyser la réponse du bassin versant en absence de pluie et donnent une information indirecte sur le niveau du stock souterrain connecté à la rivière. En milieu urbanisé, ils sont d'une grande importance en saison sèche, car ils permettent de diluer les rejets à cette saison quand il pleut peu. Ils jouent un rôle crucial dans la préservation de la biodiversité et des écosystèmes aquatiques (Bunn et Arthington, 2002). Comme indicateur, le débit de base (souvent interchangeable avec le débit lent), simple ou normalisé par le débit total (l'indice d'écoulement de base), est utilisé pour mesurer cet aspect de la réponse (Bhaskar, Hogan et Archfield, 2016; Braud *et al.*, 2013; Rose et Peters, 2001), ce qui exige une méthode de séparation de l'hydrogramme pour distinguer le débit lent du débit rapide. Sont également choisis les faibles quantiles de débit ou le débit minimal sur plusieurs jours (Konrad, Booth et Burges, 2005). Contrairement au débit fort, l'impact de l'urbanisation sur les bas débits est controversé (Tableau 1.2); des études montrent que l'urbanisation contribue à une diminution des bas débits (Braud *et al.*, 2013; Simmons et Reynolds, 1982), alors que d'autres constatent l'inverse (Bhaskar, Hogan et Archfield, 2016; Hollis, 1977; Sanzana *et al.*, 2019); d'autres études enfin ne notent aucun impact significatif de l'urbanisation sur cet aspect (Brandes, Cavallo et Nilson, 2005). Cette controverse est en partie liée à la conjugaison de plusieurs facteurs impactant la génération des bas débits, notamment le sol, le climat, la végétation et la topographie (Hamel, Daly et Fletcher, 2013), c'est-à-dire tous les facteurs intervenant dans la génération de la réponse du bassin versant. En outre, l'urbanisation contribue à l'augmentation du débit faible en augmentant la recharge souterraine (fuites dans les canalisations, réduction de l'évapotranspiration à cause de la réduction du couvert végétal), ou à travers les rejets des stations d'épuration en période sèche. À l'opposé, l'urbanisation contribue à la diminution du débit faible en diminuant la recharge souterraine par la diminution de l'infiltration, le pompage d'eau souterraine ou le drainage de la nappe par les réseaux d'assainissement.
- Les débits moyens permettent de caractériser le bilan d'eau sur le bassin versant et sa réponse sur le long terme. Son analyse a pour but de comprendre l'impact de l'urbanisation sur les disponibilités en eau de surface et sur la partition de la pluie entre écoulement et évapotranspiration. Pour cela, on utilise des métriques comme l'écoulement moyen annuel, simple ou normalisé par les précipitations moyennes annuelles (DeWalle *et al.*, 2000), ou la médiane des débits (Hollis, 1977). En général, on constate que le débit moyen augmente à cause de l'imperméabilisation des surfaces (Tableau 1.2) favorisant l'écoulement aux dépens de l'évapotranspiration.
- La dynamique de la réaction du bassin versant car l'anticipation du moment d'arrivée du pic de crue est cruciale pour la gestion des crues sur les cours d'eau, surtout en milieu urbanisé, vu l'ampleur du risque. L'analyse de l'impact sur la dynamique de la réponse peut se faire en analysant le délai entre les précipitations et l'hydrogramme de la réponse (Miller et Hess, 2017). D'autres études s'intéressent à la forme de l'hydrogramme uniquement (c'est-à-dire le temps de montée et de descente ou l'asymétrie de l'hydro-

gramme ; Diem, Hill et Milligan, 2018 ; Dow, 2007 ; Poff, Bledsoe et Cuhaciyan, 2006). On constate que l'urbanisation accélère la réponse du bassin versant en augmentant les vitesses de ruissellement à cause de la faible rugosité des surfaces imperméables et de la mise en place de réseaux d'assainissement.

Les conclusions à propos de l'impact de l'urbanisation sur les aspects du comportement hydrologique sont résumées dans le Tableau 1.2. La controverse notamment sur le bas débit illustre un point important : l'impact de l'urbanisation sur la réponse hydrologique n'est pas univoque (McGrane, 2016 ; O'Driscoll *et al.*, 2010), et il devrait être analysé en fonction des caractéristiques du bassin lui-même. Cela justifiera par la suite un recours à la représentation des processus hydrologiques à l'échelle du bassin (voir Brun et Band, 2000), en plus d'une étude d'un grand nombre de cas pour comprendre l'influence des caractéristiques des bassins versants (outre l'occupation du sol). Maintenant, nous allons discuter des différentes approches qui ont permis de détecter cet impact, en indiquant leurs avantages et leurs limites.

Tableau 1.2. Revue bibliographique de l'impact de l'urbanisation sur différents aspects du comportement hydrologique des bassins versants. Augmenté à partir d'Oudin *et al.* (2018).

Aspects de la réponse	Augmentation / Accélération	Diminution / Ralentissement	Impact non significatif / Pas systématique	Hypothèses avancées pour expliquer l'impact
Haut débit / débit rapide	Choi <i>et al.</i> , 2003; Huang <i>et al.</i> , 2008a; Hollis, 1977; Salavati <i>et al.</i> , 2016; Mejía <i>et al.</i> , 2015; Miller <i>et al.</i> , 2014; Petchprayoon <i>et al.</i> , 2010; Oudin <i>et al.</i> , 2018; Burns <i>et al.</i> , 2005; Diem, Hill et Milligan, 2018; Hawley et Bledsoe, 2011; Konrad et Booth, 2005; Miller et Hess, 2017; Prosdocimi, Kjeldsen et Miller, 2015; Rose et Peters, 2001; Rougé et Cai, 2014; Tetzlaff, Grottker et Leibundgut, 2005; Tong, 1990; Yang <i>et al.</i> , 2013; Yang <i>et al.</i> , 2014; Smith <i>et al.</i> , 2002; Smith <i>et al.</i> , 2013; Zhou <i>et al.</i> , 2017; Blum <i>et al.</i> , 2020; Sillanpää et Koivusalo, 2015; Sun <i>et al.</i> , 2017		Brun et Band, 2000	<ul style="list-style-type: none"> Augmentation des volumes de ruissellement avec une réduction des temps de transfert à cause de l'imperméabilisation des surfaces.

Suite à la page suivante...

Tableau 1.2 suite : Revue bibliographique de l'impact de l'urbanisation sur différents aspects du comportement hydrologique des bassins versants.

Aspects de la réponse	Augmentation / Accélération	Diminution / Ralentissement	Impact non significatif / Pas systématique	Hypothèses avancées pour expliquer l'impact
Bas débit / débit lent	Hollis, 1977; Burns <i>et al.</i> , 2005; Bhaskar, Hogan et Archfield, 2016; Diem, Hill et Milligan, 2018; Konrad et Booth, 2005; Rougé et Cai, 2014; Sanzana <i>et al.</i> , 2019	Choi <i>et al.</i> , 2003; Mejía <i>et al.</i> , 2015; Brun et Band, 2000; Kauffman <i>et al.</i> , 2009; Braud <i>et al.</i> , 2013; Klein, 1979; Rose et Peters, 2001; Simmons et Reynolds, 1982; Salavati <i>et al.</i> , 2016; Brandes, Cavallo et Nilson, 2005; Dow, 2007	Brandes, Cavallo et Nilson, 2005; Hejazi et Moglen, 2007; Lepeška <i>et al.</i> , 2020; Oudin <i>et al.</i> , 2018; Schwartz et Smith, 2014	<ul style="list-style-type: none"> • Augmentation de la recharge souterraine par la réduction de l'évapotranspiration; • Diminution de la recharge souterraine par la diminution de l'infiltration à cause de l'imperméabilisation des surfaces (c'est-à-dire, perte d'eau par ruissellement), et à cause du drainage par le réseau d'assainissement; • Diminution de la contribution des eaux souterraines à cause du pompage de la nappe proche-surface; • Augmentation des bas débits par le pompage de la nappe profonde, les fuites d'eau dans les canalisations, les débits des fosses septiques, et les effluents des stations d'épuration des eaux usées; • Pas d'impact significatif à cause de l'importance du couvert naturel dans le bassin versant.

Suite à la page suivante...

Tableau 1.2 suite : Revue bibliographique de l'impact de l'urbanisation sur différents aspects du comportement hydrologique des bassins versants.

Aspects de la réponse	Augmentation / Accélération	Diminution / Ralentissement	Impact non significatif / Pas systématique	Hypothèses avancées pour expliquer l'impact
Débit moyen	Ahn et Merwade, 2014; Petch-prayoon <i>et al.</i> , 2010; Hollis, 1977; Bhaskar <i>et al.</i> , 2015; Claessens <i>et al.</i> , 2006; DeWalle <i>et al.</i> , 2000; Rose et Peters, 2001; Putro <i>et al.</i> , 2016; Rougé et Cai, 2014; Tetzlaff, Grottker et Leibundgut, 2005; Diem, Hill et Milligan, 2018		Rose et Peters, 2001; Wang et Hejazi, 2011	<ul style="list-style-type: none"> En plus des processus qui impactent les hauts débits et les bas débits (que ce soit dans le cas d'augmentation ou de diminution), les transferts d'eau interbassins et les rejets des stations d'épuration.
Dynamique	Diem, Hill et Milligan, 2018; Soulsby, Birkel et Tetzlaff, 2014; Miller et Hess, 2017; Miller <i>et al.</i> , 2014; Sillanpää et Koivusalo, 2015; McMahon <i>et al.</i> , 2003; Sun <i>et al.</i> , 2017		Sun <i>et al.</i> , 2017	<ul style="list-style-type: none"> Augmentation des volumes de ruissellement et diminution de la rugosité de surface, en plus des vitesses importantes d'écoulement dans les réseaux de drainage artificiels.

Fin du Tableau 1.2.

Choix de la méthode de détection de l'impact de l'urbanisation. L'approche communément suivie s'appuie sur l'application de méthodes statistiques d'analyse de tendance ou de comparaison entre échantillons. Quand on dispose de longues séries de données, on décline cette analyse statistique dans un cadre temporel en analysant l'évolution de la réponse hydrologique en fonction de l'évolution de l'occupation du sol. On peut aussi diversifier les cas d'urbanisation en utilisant un échantillon de bassins versants avec différents niveaux d'urbanisation, puis chercher à relier la diversité des niveaux d'urbanisation avec la diversité des réponses hydrologiques.

L'approche temporelle compare, pour un bassin donné, l'évolution dans le temps d'une mesure de l'urbanisation avec l'évolution dans le temps du comportement hydrologique (Haase, 2009; Kalantari *et al.*, 2017; Smith *et al.*, 2002). Plus concrètement, le comportement hydrologique du bassin versant dans l'ère préurbaine est comparé avec celui de l'ère posturbaine. Cette approche exige une disponibilité de longues séries de données hydroclimatiques et d'occupation du sol, ce qui est possible en pratique pour un nombre limité de bassins versants et ne permet pas de tirer des conclusions transposables à d'autres bassins. Autre limite : il est difficile de dissocier l'impact de l'urbanisation de celui de la variabilité climatique (ou de celui d'autres modifications anthropiques) entre les deux ères préurbaine et posturbaine.

L'approche spatiale analyse le comportement hydrologique sur différents bassins versants ayant différents niveaux d'urbanisation. Ceci permet de remédier au manque d'une profondeur temporelle suffisante pour retrouver un gradient large d'urbanisation. La méthode des bassins versants appariés illustre un cas particulier où l'on compare un cas urbanisé (dans ce cas, notre bassin d'intérêt) avec un autre non urbanisé, à condition que les deux bassins aient des caractéristiques similaires par ailleurs (Bonneau *et al.*, 2018; Prosdocimi, Kjeldsen et Miller, 2015). Cette similarité pourrait être obtenue en utilisant des bassins versants proches dans l'espace, ce qui induit généralement des caractéristiques climatiques et géopédologiques similaires. En revanche, il est difficile de trouver un bassin versant fortement urbanisé à proximité d'un bassin faiblement urbanisé tout en garantissant une similarité dans les autres caractéristiques, surtout en termes morphologiques (taille, topographie).

Ce qui complique souvent l'exercice d'attribution est l'existence de différences dans le temps et dans l'espace, autres que l'urbanisation, qui peuvent avoir un effet confondant et biaiser l'interprétation. Dans ce cas-là, il est difficile de mettre en évidence l'effet causal de l'urbanisation. Une méthode statistique a été récemment proposée par Blum *et al.* (2020) pour mieux cerner cet effet causal. Cette méthode couramment utilisée en économétrie s'appuie sur une régression sur données de panel (caractérisées par une variabilité dans l'espace et dans le temps) qui permet d'éliminer les biais liés aux différences spatiales entre les bassins (par exemple, topographie) et temporels (d'une période à une autre). Le point fort de cette approche est de prendre en compte les différents facteurs grâce à un graphe causal, permettant notamment de décortiquer ces facteurs en des effets liés à des différences purement

temporelles, purement spatiales (facteurs régionaux), et croisées (facteurs spatiotemporels). Si cette méthode nouvelle permet d'améliorer l'attribution, elle n'enlève pas l'intérêt de l'application des modèles hydrologiques pour exploiter les connaissances sur les processus hydrologiques.

Le principe dans un modèle hydrologique est que les paramètres synthétisent les caractéristiques physiques du bassin versant et ne varient pas (en théorie) avec les forçages climatiques, ce qui permet de limiter le problème de la variabilité climatique rencontré dans l'approche temporelle (McIntyre *et al.*, 2014). De plus, l'utilisation d'un modèle hydrologique permet de faire le lien entre l'impact à l'échelle locale et la simulation à l'échelle du bassin versant. Pour la détection et la quantification de l'impact de l'urbanisation, le modèle peut être utilisé de différentes manières (Salavati, 2015, p. 26; Salavati *et al.*, 2016) :

1. Utiliser des modèles à base physique ou conceptuels dont les équations permettent de décrire les processus à l'échelle locale, sur des mailles ou sur des sous-entités homogènes du bassin d'intérêt, et dont les paramètres sont explicitement liés aux mesures de l'urbanisation (Guan, Sillanpää et Koivusalo, 2015; Branger *et al.*, 2013; Niehoff, Fritsch et Bronstert, 2002; Zhang et Shuster, 2014; Zhou *et al.*, 2013). Ces sous-entités sont ensuite liées au réseau de drainage qui permet de combiner leurs effets et les traduire à l'échelle du bassin versant. Plusieurs arguments sont en faveur de l'utilisation de cette approche, notamment la possibilité d'estimer les paramètres directement sans avoir recours au calage numérique. De plus, cette approche permet d'étudier l'effet causal de chaque facteur indépendamment des autres. En revanche, il est difficile de spécifier le jeu de paramètres auquel la réponse (et donc l'impact) est la plus sensible (Redfern *et al.*, 2016). En pratique, le manque de données exigées pour déterminer *a priori* les paramètres amène en général à effectuer un calage.
2. Utiliser un modèle conceptuel à l'échelle du bassin versant en le calant sur la période préurbaine pour obtenir un jeu de paramètres correspondant à un comportement d'avant urbanisation du bassin, et simuler avec ce jeu le comportement pendant la période posturbaine, ce qui permet de comparer l'écoulement simulé (qui est supposé être la réponse du bassin en état préurbain) avec celui observé (qui est la réponse actuelle du bassin en état urbanisé). L'analyse statistique des résidus que sont les différences entre l'écoulement simulé et l'écoulement observé facilite l'appréciation et la quantification du changement dû à l'urbanisation (Salavati *et al.*, 2016; Seibert et Meerveld, 2016). Comme exigence, le modèle devrait reproduire de manière satisfaisante l'écoulement observé sur la période préurbaine.
3. Utiliser un modèle conceptuel à l'échelle du bassin versant en calant ses paramètres sur des fenêtres temporelles glissantes pour comparer leur évolution dans le temps avec l'évolution de l'urbanisation du bassin versant (Miller *et al.*, 2014; Pathiraja *et al.*, 2018;

Sillanpää et Koivusalo, 2015). En particulier, on peut se limiter à comparer le jeu de paramètres obtenus sur une période préurbaine avec celui d'une période posturbaine. Une exigence supplémentaire par rapport au cas précédent : le modèle devrait reproduire l'écoulement pour toute la période, y compris les situations urbanisées. De plus, les paramètres devraient avoir un lien plus ou moins explicite avec l'urbanisation pour faciliter leur interprétation.

4. Utiliser un modèle à l'échelle du bassin versant et comparer ses paramètres calés sur des bassins fortement urbanisés avec ceux des bassins faiblement urbanisés, ce qui revient à l'application d'une approche spatiale en s'aidant du modèle (Zégre *et al.*, 2010). Ceci permet éventuellement de régionaliser les paramètres du modèle pour obtenir une relation avec les caractéristiques de l'urbanisation du bassin versant. Les performances du modèle utilisé devraient alors être satisfaisantes pour l'ensemble des bassins étudiés.

Un des avantages de l'utilisation d'un modèle hydrologique par rapport à une approche statistique ordinaire (c'est-à-dire sans utilisation de modèle hydrologique) est la possibilité d'obtenir des scénarios hydrologiques à partir des scénarios d'urbanisation, *via* la variation des paramètres du modèle. Ceci exige une structure à la fois liée aux caractéristiques de l'urbanisation et capable de reproduire de manière satisfaisante les écoulements observés sur les bassins urbanisés. À la lumière de ces exigences, nous allons faire un aperçu général sur les outils de modélisation existants pour les bassins versants urbanisés.

1.3 Aperçu des approches de modélisation des bassins versants urbanisés

1.3.1 Notions générales et classifications classiques

Nous entendons par modèle hydrologique une représentation simplifiée des processus hydrologiques qui permette d'expliquer la transformation du signal de précipitations en signal de débit observé à la sortie du système modélisé (une entité spatiale, dans notre cas un bassin versant limité par les lignes topographiques de partage des eaux, voir Figure 1.2), sur une période de temps donnée. Souvent, le terme désigne l'ensemble des équations qui interviennent dans la transformation pluie-débit. On peut décrire un modèle comme une fonction

$$f(FC(t), S(t) | \theta) = Q_{sim}(t) \quad (1.3.1)$$

où f désigne la structure du modèle, c'est-à-dire l'ensemble des équations qui le constituent. $FC(t)$ est l'ensemble des forçages climatiques, qui sont les entrées du modèle. $S(t)$ représente les variables d'état du modèle, un ensemble de variables internes qui peuvent être conçues comme les états dynamiques des différents compartiments du bassin versant à l'instant t (par exemple, le niveau d'un modèle-réservoir peut être lié au niveau piézométrique ou au niveau de l'humidité du sol), et qui font le lien entre les forçages et la réponse. θ est le vecteur des paramètres du modèle, censés traduire/être liés aux caractéristiques du bassin versant (par exemple, la porosité ou la topographie). $Q_{sim}(t)$ désigne le débit simulé par le modèle.

Le modèle communique l'information des forçages aux variables d'état, puis des variables d'état à la sortie, *via* des flux internes. Cette communication est conditionnée par les paramètres du modèle; par exemple, si l'on considère un réservoir représentant le sol, avec un paramètre θ_1 représentant la capacité maximale du réservoir (qui peut être associée à la réserve utile du sol), et un niveau dynamique $HS(t)$ représentant le niveau d'humidité du sol, alors la proportion des précipitations (forçage) qui alimente ce réservoir devrait être régie par le paramètre θ_1 , ainsi que le niveau dynamique $HS(t)$ (état du modèle, du bassin, au moment du forçage). En théorie, les paramètres devraient être indépendants des forçages et du temps, et n'être liés qu'aux caractéristiques physiques spatiales du bassin versant. Ils peuvent être soit estimés à partir de mesures, soit calés par optimisation numérique sur une période donnée de manière à avoir un débit simulé le plus proche possible du débit observé. Une revue synthétique des modèles pluie-débit en hydrologie peut être retrouvée dans Singh et Woolhiser (2002).

Les modèles sont utilisés pour différents objectifs : prévoir des crues et des sécheresses, compléter les séries manquantes de débit, estimer des débits de projet pour dimensionner les ouvrages, ou simuler l'impact du climat et du changement d'occupation du sol sur la réponse hydrologique (Hrachowitz et Clark, 2017). Les modèles ne servent pas uniquement à des fins opérationnelles; ils encapsulent un jeu d'hypothèses, et le test du modèle sert à valider ou réfuter ce jeu (ou au moins à en tracer les limites), ce qui permet d'avancer nos connaissances sur le fonctionnement hydrologique du bassin versant (Blöschl, 2017).

Les modèles en hydrologie sont très variés et aucun consensus n'existe par rapport à la supériorité d'un modèle sur un autre, vu le manque de compréhension suffisante des différents processus (en partie à cause d'un manque de mesures des processus aux échelles d'intérêt) et la divergence en matière d'application et de choix des approches de modélisation (Weiler et Beven, 2015). Néanmoins, on peut discerner des traits communs qui permettent de classer ces modèles (Ficchi, 2017, pp. 22-29) :

1. Selon le pas de temps de fonctionnement du modèle : en des modèles au pas de temps infra-horaire (quelques minutes), horaire, infra-journalier, journalier, mensuel, annuel, ou interannuel.

2. Selon la représentation spatiale du bassin versant : dans un modèle global (*lumped* en anglais), on représente le bassin versant comme une entité homogène spatialement, et les forçages sont spatialement moyennés sur cette entité. Les modèles distribués ont pour ambition de représenter explicitement l'hétérogénéité spatiale des processus hydrologiques en rajoutant un degré de variabilité spatiale au sein du bassin versant avec des entités plus petites. Dans tous les cas, il y a une résolution maximale au-delà de laquelle l'hétérogénéité spatiale est masquée.
3. Selon l'approche de représentation des processus : quand le modèle se sert d'équations développées à petite échelle pour représenter les processus hydrologiques (équations de Darcy/Richards, Green-Ampt, Saint-Venant, etc.), on parle alors d'un modèle à base physique. Sinon, quand le modèle utilise des équations permettant de représenter de façon simplifiée les traits de comportement observés d'un bassin versant, comme les hydrogrammes unitaires ou les réservoirs, il s'agit alors d'un modèle conceptuel (ou modèle « boîte grise »). Un dernier cas des modèles s'appuie sur des relations statistiques entre les entrées et les sorties, sans possibilité de distinguer des processus hydrologiques au sein de leurs structures : on parle alors de modèles « boîte noire ».

Nous allons voir que les modèles utilisés pour la modélisation des bassins versants urbanisés obéissent aux mêmes classifications (Fletcher, Andrieu et Hamel, 2013; Salvadore, Bronders et Batelaan, 2015). Dans la prochaine section, nous ferons un aperçu général des modèles utilisés et nous discuterons de leurs limites vis-à-vis de nos objectifs.

1.3.2 Panorama des modèles hydrologiques en milieu urbanisé

Les objectifs et les échelles de modélisation du milieu urbanisé ont évolué dans le temps (Carré et Deutsch, 2015, p. 46). Si l'on écarte le dimensionnement des réseaux d'eau potable et d'assainissement, la modélisation pluie-débit s'appuyait en général sur la méthode rationnelle et se focalisait sur l'estimation du ruissellement pour dimensionner les réseaux d'évacuation des eaux pluviales et se protéger contre les inondations. Avec la prise de conscience des enjeux environnementaux à partir des années 1970, l'utilisation des techniques alternatives pour gérer le ruissellement à la source et rétablir le régime naturel a créé de nouveaux objectifs de modélisation. Des modèles ont alors été utilisés pour dimensionner ces techniques, les représenter et vérifier leur efficacité, et ce à partir des années 1980 (voir les revues de Zoppou, 2001; Elliott et Trowsdale, 2007). Vers les années 1990-2000, et en continuité avec la gestion à la source, on commença à mettre en place des approches intégrées qui visaient à concevoir le cycle de l'eau de manière plus complète à l'échelle de la ville, en prenant en compte l'alimentation en eau potable, l'assainissement et l'écoulement souterrain, en plus d'une inclusion des différentes parties prenantes dans la gestion de l'eau en milieu urbain, et des différents

aspects environnementaux du métabolisme urbain (Bach *et al.*, 2014; Fletcher, Andrieu et Hamel, 2013; Fletcher *et al.*, 2015; Hutchins *et al.*, 2017). Cet appel à revoir le mouvement de l'eau de manière intégrée imposa d'élargir le périmètre des processus hydrologiques qui devaient être pris en compte en milieu urbain, alors qu'on se limitait jusqu'à présent à la représentation du ruissellement et du drainage ou à considérer chaque processus à part. Sous des contraintes environnementales (impact de l'urbanisation, changement climatique), conjuguées désormais à des contraintes de planification (objectifs du développement durable), l'externalisation des besoins des citoyens au-delà de leur ville se conçoit comme une rupture de l'équilibre écologique local (Carré et Deutsch, 2015, p. 40), ce qui contraint à considérer le cycle hydrologique à l'échelle du bassin hydrographique pour sécuriser les approvisionnements dans le contexte du changement actuel et gérer au mieux les rejets des stations d'épuration. Cela s'est traduit en modélisation par un appel à une prise en compte des zones rurales/naturelles entourant les villes (Braud, Fletcher et Andrieu, 2013), ce qui donna lieu à des outils de modélisation pour les bassins caractérisés par une mosaïque de couverts naturel/urbain.

Ces objectifs de modélisation en milieu urbanisé sont toujours poursuivis, comme on peut le constater à partir du Tableau 1.3, où on liste un échantillon de 45 applications de modèles conçus pour les bassins urbanisés. L'objectif de cette revue est de donner un aperçu des approches existantes pour représenter les milieux urbanisés, même si les objectifs de modélisation sont fondamentalement différents d'une application à une autre. Les sections suivantes proposent une discussion sur la base de cette synthèse.

Tableau 1.3. Revue bibliographique des modèles hydrologiques utilisés pour les bassins urbanisés. Cette revue est largement appuyée sur la synthèse de Salvatore, Bronders et Batelaan (2015). Par rapport à cette revue, nous avons classifié les modèles (autant que possible) selon le pas de temps et la représentation spatiale, et nous avons ajouté le nombre de paramètres (pour donner une indication sur le niveau de complexité du modèle et sa demande potentielle en données), ainsi que l'objectif de modélisation. Le nombre de paramètres est parfois difficile à estimer des publications et il est indiqué ici pour les études le citant explicitement. Il est plutôt sous-estimé pour les approches distribuées et semi-distribuées, et devrait être multiplié par le nombre des sous-entités spatiales servant à représenter le bassin. Pour les approches à base physique, les paramètres sont souvent contraints de manière expertisée par le modélisateur, alors qu'ils sont estimés par un calage automatique pour les approches conceptuelles. Pour les modèles semi-distribués, ÉHU signifie élément hydrologique urbain (bâtiment, bloc de bâtiments, quartier, rue, etc.). MNT signifie Modèle Numérique du Terrain.

Modèle / Application	Conceptuel ou à base physique?	Événementiel ou continu?	Nombre de bassins	Taille des bassins (km ²)	Représentation spatiale	Objectif	Nombre de paramètres
Distribué au pas de temps horaire ou infra-horaire (19)							
Modèle d'Aronica et Lanza (2005)	À base physique	Événementiel	1	< 1	Distribué	Inondation	?
SURF (Bellal, Sillen et Zech, 1996; Rodriguez, Andrieu et Zech, 2000; Zech <i>et al.</i> , 1994)	À base physique	Continu	3	< 1	Distribué	Simulation de la réponse d'un bassin urbanisé, impact des aménagements urbains	7 – 12
GUFIM (Chen, Hill et Urbano, 2009)	Conceptuel / à base physique	Événementiel	1	~1	Distribué	Inondation	?

Suite à la page suivante...

Tableau 1.3 suite : Revue bibliographique des modèles hydrologiques utilisés pour les bassins versants urbanisés.

Modèle / Application	Conceptuel ou à base physique?	Événementiel ou continu?	Nombre de bassins	Taille des bassins (km ²)	Représentation spatiale	Objectif	Nombre de paramètres
SURFEX (Hamdi, Termolina et Baguis, 2011; Lemonu, Masson et Berthier, 2007)	À base physique	Continu	2	< 1 – 150	Distribué	Couplage sol-atmosphère, climat en milieu urbain	15
U-McIUH (Gironás <i>et al.</i> , 2009)	Conceptuel	Événementiel	2	1 – 10	Distribué	Amélioration de la représentation du routage en milieu urbain	> 18
TOPURBAN (Valeo et Moïn, 2000)	Conceptuel	Continu	1	~10	Distribué	Impact de la résolution spatiale sur la réponse hydrologique d'un bassin urbanisé	?
GwNeu (Göbel <i>et al.</i> , 2004)	À base physique	Événementiel	1	~10	Distribué	Impact des techniques alternatives sur la recharge souterraine	?

Suite à la page suivante...

Tableau 1.3 suite : Revue bibliographique des modèles hydrologiques utilisés pour les bassins versants urbanisés.

Modèle / Application	Conceptuel ou à base physique?	Événementiel ou continu?	Nombre de bassins	Taille des bassins (km ²)	Représentation spatiale	Objectif	Nombre de paramètres
HYDRUS_2D (Göbel <i>et al.</i> , 2004)	À base physique	Événementiel	1	~10	Distribué	Impact des techniques alternatives sur la recharge souterraine	?
SPRING (Göbel <i>et al.</i> , 2004)	À base physique	Événementiel	1	~10	Distribué	Impact des techniques alternatives sur la recharge souterraine	?
GSSHA + SUPERLINK (Ogden <i>et al.</i> , 2011)	À base physique	Continu	1	> 10	Distribué	Sensibilité de la réponse d'un bassin versant aux forçages climatiques	> 24
DHSVM (Cuo <i>et al.</i> , 2008)	À base physique	Continu	1	> 20	Distribué	Bilan hydrologique et énergétique des bassins versants à topographie complexe	?
WEP (Jia <i>et al.</i> , 2001)	À base physique	Continu	1	~25	Distribué	Bilan d'énergie en milieu urbanisé	8

Suite à la page suivante...

Tableau 1.3 suite : Revue bibliographique des modèles hydrologiques utilisés pour les bassins versants urbanisés.

Modèle / Application	Conceptuel ou à base physique?	Événementiel ou continu?	Nombre de bassins	Taille des bassins (km ²)	Représentation spatiale	Objectif	Nombre de paramètres
DORS (Zhou <i>et al.</i> , 2010)	À base physique	Continu	2	25 – 100	Distribué / Semi-distribué ÉHU	Simulation de la réponse des bassins versants urbanisés	?
ISBA-TOPMODEL (Furu-sho, Chancibault et Andrieu, 2013)	Conceptuel / à base physique	Continu	1	~30	Distribué	Représentation du cycle de l'eau pour les bassins péri-urbains	5
WetSpa (Berezowski <i>et al.</i> , 2012; Chormanski <i>et al.</i> , 2008; Verbeiren <i>et al.</i> , 2013)	À base physique	Continu	4	30 – 150	Distribué	Impact de l'estimation de l'imperméabilité sur les pics de débit, impact des différentes classes d'occupation du sol	> 9
MERCEDES (Lhomme, Bouvier et Perrin, 2004)	Conceptuel	Événementiel	1	> 50	Distribué	Apport des MNT dans la simulation de la réponse d'un bassin versant urbanisé	?

Suite à la page suivante...

Tableau 1.3 suite : Revue bibliographique des modèles hydrologiques utilisés pour les bassins versants urbanisés.

Modèle / Application	Conceptuel ou à base physique?	Événementiel ou continu?	Nombre de bassins	Taille des bassins (km ²)	Représentation spatiale	Objectif	Nombre de paramètres
Modèle de Mejía et Mørgen (2010)	Conceptuel / à base physique	Événementiel	1	> 100	Distribué	Impact de la distribution spatiale de l'imperméabilité sur la réponse hydrologique	5
MIKE SHE / MIKE 11 (Im et al., 2008; Vrebos et al., 2014)	À base physique	Continu	2	> 200	Distribué	Simulation de la réponse d'un bassin versant urbanisé	32
TAC (Ott et Uhlenbrook, 2004)	À base physique	Continu	1	> 250	Distribué	Impact du changement d'occupation du sol	?
Semi-distribué au pas de temps horaire ou infra-horaire (14)							
Modèle de Xiao et al. (2007)	À base physique	Continu	1	< 1	Semi-distribué ÉHU	Impact des techniques alternatives à l'échelle résidentielle	?
Modèle de Berthier, Andrieu et Creutin (2004)	Conceptuel / à base physique	Continu	1	< 1	Semi-distribué ÉHU	Rôle du sol en milieu urbain	12

Suite à la page suivante...

Tableau 1.3 suite : Revue bibliographique des modèles hydrologiques utilisés pour les bassins versants urbanisés.

Modèle / Application	Conceptuel ou à base physique?	Événementiel ou continu?	Nombre de bassins	Taille des bassins (km ²)	Représentation spatiale	Objectif	Nombre de paramètres
URBS (Rodriguez, Andrieu et Creutin, 2003; Rodriguez, Andrieu et Morena, 2008)	À base physique	Continu	2	<1	Semi-distribué ÉHU	Simulation de la réponse des bassins versants urbanisés	11
SURFEX-TEB (Stavropoulos-Laffaille <i>et al.</i> , 2018)	À base physique	Continu	2	<1	Semi-distribué	Couplage sol-atmosphère	> 9
IUHM (Cantone et Schmidt, 2011)	À base physique	Événementiel	1	<1	Semi-distribué	Représentation du routage en milieu urbain	?
CAT (Miller <i>et al.</i> , 2014)	À base physique	Événementiel	2	~5	Semi-distribué	Impact de l'urbanisation sur la réponse hydrologique d'un bassin péri-urbain	11
PUMMA (Fuamba <i>et al.</i> , 2019; Jankowfsky <i>et al.</i> , 2014)	À base physique	Continu	2	>1	Semi-distribué ÉHU	Représentation intégrale des processus hydrologiques dans les bassins versants péri-urbains	32

Suite à la page suivante...

Tableau 1.3 suite : Revue bibliographique des modèles hydrologiques utilisés pour les bassins versants urbanisés.

Modèle / Application	Conceptuel ou à base physique?	Événementiel ou continu?	Nombre de bassins	Taille des bassins (km ²)	Représentation spatiale	Objectif	Nombre de paramètres
SWMM (Petrucci et Bonhomme, 2014)	À base physique	Événementiel	1	> 1	Semi-distribué	Complexité des modèles et stratégies de calage	> 15
Modèle d'Aronica et Canarozzo (2000)	Conceptuel	Événementiel	1	< 1	Semi-distribué perméable / imperméable	Effet de la variabilité spatiale de pluie sur la réponse hydrologique	6
RORB (Selvalingam, Liang et Manoharan, 1987)	Conceptuel	Événementiel	1	> 1	Semi-distribué perméable / imperméable	Simulation de la réponse d'un bassin versant urbanisé	> 8
SUES (Grimmond et Oke, 1991 ; Mitchell <i>et al.</i> , 2008)	À base physique	Continu?	2	< 1-30	Semi-distribué	Estimation de l'évapotranspiration en milieu urbanisé	> 30?
MUSIC (Dotto <i>et al.</i> , 2011 ; Hamel et Fletcher, 2014)	Conceptuel	Continu	6	< 1-40	Semi-distribué perméable / imperméable	Impact des techniques alternatives	5-16
CANOE (Lhomme, Bouvier et Perrin, 2004)	Conceptuel / à base physique	Événementiel	1	> 50	Semi-distribué	Apport des MNT dans la simulation de la réponse d'un bassin urbanisé	3?

Suite à la page suivante...

Tableau 1.3 suite : Revue bibliographique des modèles hydrologiques utilisés pour les bassins versants urbanisés.

Modèle / Application	Conceptuel ou à base physique?	Événementiel ou continu?	Nombre de bassins	Taille des bassins (km ²)	Représentation spatiale	Objectif	Nombre de paramètres
CAREDas (Thielen et Creutin, 1997)	À base physique	Événementiel	1	> 100	Semi-distribué	Couplage atmosphère en période de crue	?
Global au pas de temps horaire ou infra-horaire (2)							
KAREN (Dotto et al., 2011; Kleidorfer et al., 2009)	Conceptuel	Continu	5	< 1 avec un bassin > 1	Global	Impact de l'incertitude des données sur les paramètres du modèle, couplage avec les modèles de qualité	4
Modèle de Sun et al. (2017)	Conceptuel	Événementiel	2	< 1	Global	Impact de l'urbanisation sur la réponse hydrologique	4
Distribué au pas de temps journalier (1)							
SMDR (Easton et al., 2007)	À base physique	Continu	1	> 1	Distribué	Simulation du ruissellement et de l'humidité du sol sur une aire imperméabilisée	> 8

Suite à la page suivante...

Tableau 1.3 suite : Revue bibliographique des modèles hydrologiques utilisés pour les bassins versants urbanisés.

Modèle / Application	Conceptuel ou à base physique?	Événementiel ou continu?	Nombre de bassins	Taille des bassins (km ²)	Représentation spatiale	Objectif	Nombre de paramètres
Semi-distribué au pas de temps journalier (6)							
Aquacycle, UVQ, NEIMO (Mitchell, Mein et McMahon, 2001; Mitchell et Diaper, 2005; Morris et al., 2007)	Conceptuel	Continu	1	30	Semi-distribué ÉHU	Représentation du cycle de l'eau en milieu urbanisé	28
AVSWAT (Dixon et Earls, 2012; Franczyk et Chang, 2009)	Conceptuel / à base physique	Continu	2	200-850	Semi-distribué ÉHU	Impact du changement climatique et d'occupation du sol sur la réponse hydrologique	?
HBV + SOBEK (Te Linde, Aerts et Kwadijk, 2010)	Conceptuel / à base physique	Continu	1	> 105	Semi-distribué	Impact de l'urbanisation sur le risque d'inondation	?
SIMHYD (Singh, Maheshwari et Malano, 2009)	Conceptuel	Continu	1	> 500	Semi-distribué perméable / imperméable + des flux hydrologiques d'origine anthropique	Représentation du cycle de l'eau pour les bassins péri-urbains	21

Suite à la page suivante...

Tableau 1.3 suite : Revue bibliographique des modèles hydrologiques utilisés pour les bassins versants urbanisés.

Modèle / Application	Conceptuel ou à base physique?	Événementiel ou continu?	Nombre de bassins	Taille des bassins (km ²)	Représentation spatiale	Objectif	Nombre de paramètres
WEAP-MODFLOW (Sanzana <i>et al.</i> , 2019)	À base physique	Continu	1	> 130	Semi-distribué / distribué	Impact des fuites d'eau et de l'irrigation sur le régime hydrologique d'un bassin urbanisé	15
J2000 (Branger <i>et al.</i> , 2013)	Conceptuel / à base physique	Continu	1	150	Semi-distribué ÉHU	Impact de l'utilisation de différentes bases de données spatiales sur l'estimation des composantes du bilan hydrologique d'un bassin péri-urbain	> 20
Global au pas de temps journalier (3)							
Basins (Brun et Band, 2000)	Conceptuel	Continu	1	> 150	Global	Impact du changement d'occupation du sol	?
ReFH (Faulkner et Barber, 2009; Kjeldsen, 2010; Kjeldsen, Miller et Packman, 2013)	Conceptuel	Événementiel	7	40-250	Global	Dimensionnement, simulation de la réponse des bassins versants urbanisés	4

Suite à la page suivante...

Tableau 1.3 suite : Revue bibliographique des modèles hydrologiques utilisés pour les bassins versants urbanisés.

Modèle / Application	Conceptuel ou à base physique?	Événementiel ou continu?	Nombre de bassins	Taille des bassins (km ²)	Représentation spatiale	Objectif	Nombre de paramètres
SCS + Cascade de Nash (Huang <i>et al.</i> , 2008b)	Conceptuel	Événementiel	1	> 200	Global	Impact de l'imperméabilité sur le comportement hydrologique	3

Fin du Tableau 1.3.

1.3.2.1 Représentation des spécificités de l'urbanisation dans la structure du modèle

Les objectifs de gestion de l'eau en milieu urbain ont généralement une portée locale, ce qui a engendré un focus à faible échelle temporelle et spatiale sur les processus hydrologiques (Berne *et al.*, 2004; Cristiano, Veldhuis et Giesen, 2017). De plus, le paysage urbain est considéré très hétérogène spatialement à cause de la diversité des aménagements urbains (Gires *et al.*, 2017). La forte hétérogénéité des bassins versants urbanisés matérialisée par la complexité du paysage a conduit à une représentation détaillée des processus dans l'espace et dans le temps pour 33/45 applications de modèles listées dans le Tableau 1.3. Les bassins versants étudiés sont généralement d'une faible taille (ne dépassant pas une dizaine de km²), induisant des temps de transfert à l'exutoire courts. Cette dynamique rapide de la réponse est accentuée par des vitesses importantes induites par la faible rugosité des surfaces imperméabilisées et l'importante densité de drainage. Pour mieux reproduire cette dynamique, l'utilisation du pas de temps horaire et infra-horaire devient alors une exigence. En outre, face à l'hétérogénéité du paysage urbain, les auteurs ont privilégié les approches (i) distribuée, (ii) semi-distribuée avec une distinction des objets modélisés (par exemple, la représentation basée sur l'élément hydrologique urbain), ou (iii) semi-distribuée avec une prise en compte d'un premier degré de différence perméable/imperméable, qui revient généralement à convertir toute l'eau tombant sur la partie imperméable en ruissellement. Enfin, la majorité de ces modèles utilisent des équations à base physique, ce qui fournit un lien direct entre les propriétés physiques du bassin (y compris celles des surfaces urbaines) et la structure et/ou les valeurs des paramètres du modèle.

À l'autre extrémité, les modèles conceptuels globaux offrent un nombre de paramètres réduit par rapport aux modèles à représentation détaillée. Ceci s'accompagne de l'utilisation d'outils conceptuels où le lien entre la structure du modèle et les caractéristiques d'urbanisation est généralement implicite, ce qui limite l'application de tels modèles pour les études d'impact hydrologique de scénarios d'urbanisation. Une exception est illustrée par la méthode du Curve Number dont le paramètre *CN* peut être lié à l'occupation du sol par des relations empiriques mais qui restent à ce jour peu validées (Ogden *et al.*, 2017). Enfin, on remarque que les modèles journaliers sont appliqués sur des bassins de taille relativement importante, allant de 1 jusqu'à 10⁵ km².

1.3.2.2 Limitation de la complexité de l'outil

Si les modèles à représentation explicite (distribués, à base physique) offrent un lien explicite entre les spécificités de l'urbanisation et la réponse hydrologique du bassin versant, ils sont généralement lourdement paramétrés voire surparamétrés : le nombre de leurs paramètres dépasse la quantité disponible de données nécessaires à les contraindre (Jakeman et Hornber-

ger, 1993). En soi, cela n'est pas étonnant, car ce n'est que le résultat trivial d'une volonté de représenter la complexité paysagère des bassins urbanisés. Ce choix est motivé par le besoin de simulations distribuées notamment pour les modèles de transport des contaminants ou la gestion des réseaux en milieu urbain (Petrucci et Bonhomme, 2014); mais il est également fondé sur l'argument que la prédiction de l'impact de l'occupation des sols ne pourrait se faire qu'à partir des modèles distribués à base physique (Beven, 2002; Bronstert, Niehoff et Bürger, 2002; Niehoff, Fritsch et Bronstert, 2002; Wijesekara *et al.*, 2012). La surparamétrisation accentue le problème d'équifinalité et de compensation entre les paramètres (Ebel et Loague, 2006; Beven, 2006), ce qui met en question la robustesse du modèle. Pour limiter ce problème, des contraintes sur les paramètres devraient être imposées à partir d'un large volume de données (comme par exemple les banques de données urbaines, voir Rodriguez, Andrieu et Creutin, 2003; Berthier, Andrieu et Rodriguez, 1999), dont l'acquisition nécessite un déploiement instrumental important et possible pour un nombre limité de cas.

Comme mentionné précédemment, l'utilisation des modèles globaux réduit fortement le nombre de paramètres à estimer. Ces paramètres représentent une information agrégée sur le bassin versant, où le lien avec l'urbanisation est plus diffus et moins explicite, ce qui rend leur application pour l'étude des scénarios difficile (voir la sous-section précédente).

1.3.2.3 Test intensif des modèles

Pour être robuste et transposable (temporellement et spatialement), un modèle hydrologique dédié aux bassins versants urbanisés doit prouver qu'il peut reproduire de manière satisfaisante l'écoulement observé sur un grand nombre de bassins versants et pour différentes situations d'urbanisation, ce qui le rendrait crédible pour une prédiction de l'impact d'une urbanisation future. Or, on remarque que tous les exemples d'application montrés dans le Tableau 1.3 ont concerné des échantillons faibles de bassins versants, en majorité un bassin par application (Figure 1.3). Pour être utile pour différentes applications, le modèle doit être capable de simuler une large gamme de débit, que ce soit à l'échelle de l'événement pluviométrique, mais également sur des temps plus longs. Les applications sont partagées entre la simulation de la réponse du bassin en continu (27/45) et sur quelques événements (18/45), prouvant que les processus de subsurface qui interviennent notamment dans la réponse du bassin en dehors des événements sont également ciblés dans la plupart des modèles en milieu urbanisé. Pour les modèles à représentation explicite, l'utilisation de faibles échantillons peut être imputée à la demande importante de données pour estimer/contraindre leurs paramètres, en particulier les caractéristiques des éléments hydrologiques urbains et les paramètres géométriques des réseaux artificiels d'acheminement d'eau. En revanche, la question d'accès aux données est beaucoup moins fondamentale pour les modèles relativement simples (conceptuels, globaux), pour lesquels la paramétrisation est beaucoup moins lourde. La plupart des grands échantillons

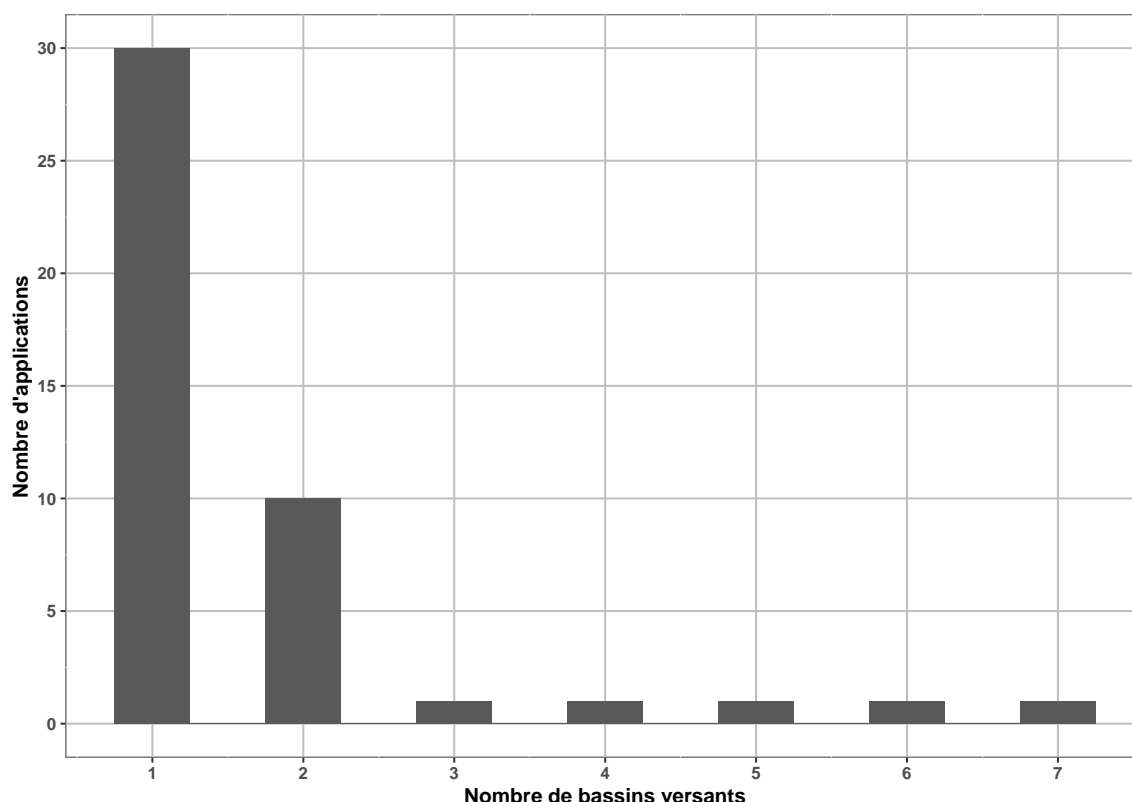


Figure 1.3. Distribution du nombre des bassins versants utilisés pour les applications des 45 modèles listées dans le Tableau 1.3.

actuels comportent plutôt des bassins versants non urbanisés (Gupta *et al.*, 2014), ce qui reflète l'aspect *monographique* de la modélisation des bassins versants urbanisés.

1.3.3 Vers une approche hybride

Notre objectif est de chercher une structure de modèle qui permette une étude d'impact des scénarios de l'urbanisation sur le comportement hydrologique des bassins versants. Cette structure recherchée doit donc satisfaire les conditions suivantes : **(1) comprendre un lien entre la structure/les paramètres du modèle et les caractéristiques de l'urbanisation** pour rendre possible une traduction des scénarios d'urbanisation en scénarios hydrologiques ; **(2) avoir un équilibre entre le niveau de paramétrisation et le volume de données disponibles pour un grand nombre de bassins versants** ; **(3) obtenir des performances satisfaisantes sur un large échantillon de bassins versants**, ce qui prouverait la fiabilité des prédictions du modèle. À partir de la revue précédente (Section 1.3.2), nous constatons le manque d'une structure à la fois simple, présentant un lien direct entre les propriétés urbaines du bassin et sa réponse hydrologique, et testée sur un échantillon suffisamment large de bassins versants présentant une diversité de situations d'urbanisation. Comme résumé sur la Figure 1.4, deux options sont envisageables : **(1)** partir d'une structure de modèle où le lien est déjà présent et

la simplifier pour ne garder que les composantes suffisantes pour une bonne reproduction de la relation pluie-débit d'un bassin versant urbanisé avec des données nécessaires *minimales* (typiquement les forçages climatiques et une classification de l'usage du sol), ou (2) partir d'une structure simple, et la complexifier pour améliorer sa prise en compte de l'urbanisation. Dans les deux cas, le développement de la structure du modèle doit s'effectuer sur un large échantillon de bassins versants afin de garantir une certaine transposabilité spatiale et temporelle (c'est-à-dire la « robustesse » du modèle).

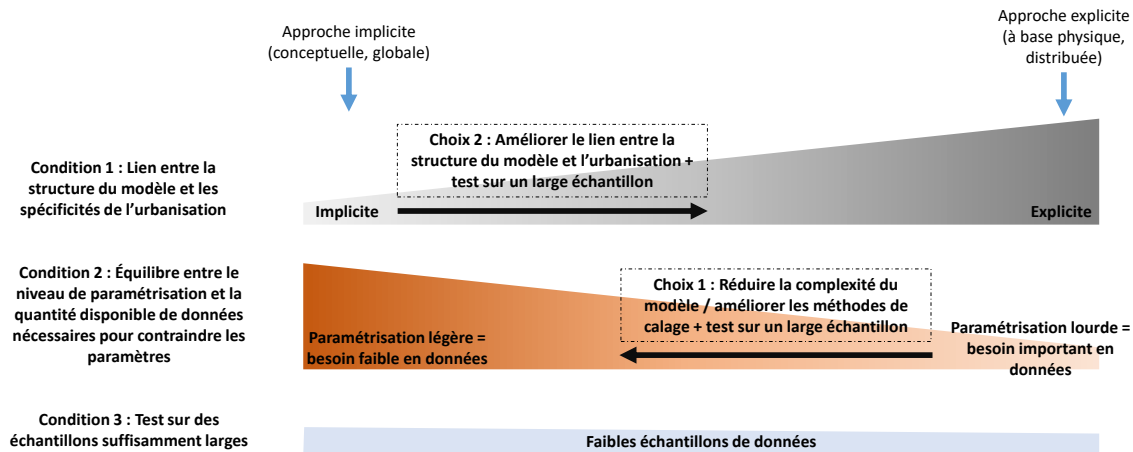


Figure 1.4. Synthèse des approches actuelles de modélisation en milieu urbanisé. On illustre le compromis entre la simplicité de l'outil (c'est-à-dire les exigences en données nécessaires pour contraindre les paramètres) et le lien entre sa structure et les spécificités urbaines du bassin versant.

1.3.3.1 Choix 1 : réduire la complexité d'un modèle à représentation explicite

Petrucci et Bonhomme (2014) ont cherché une stratégie pour contraindre les paramètres d'un modèle semi-distribué par l'utilisation de données non hydrologiques (issue de données géospatiales, notamment la géométrie des versants), en testant différents scénarios de calage-validation correspondant à des niveaux de complexité différents sur un bassin versant de 2.3 km² de surface. Ils ont conclu qu'il est possible de diminuer le niveau de détail dans la représentation spatiale tout en obtenant des performances égales voire meilleures. Par ailleurs, Awol, Coulibaly et Tolson (2018) ont testé plusieurs approches de calage d'un modèle distribué (SWMM) en utilisant 14 événements sur un bassin de 911 km² situé à Ontario. Leur objectif était de développer une méthode permettant de réduire le nombre de paramètres à caler en passant par une analyse de sensibilité, une analyse de la variabilité spatiotemporelle des paramètres et une combinaison de plusieurs algorithmes de calage. Cette étude a illustré le besoin de méthodes de calage robustes avec des choix judicieux sur les fonctions objectif pour pallier l'incertitude importante dans le modèle distribué.

Ces deux exemples sont une illustration du choix de réduire le niveau de complexité du modèle en gardant un lien entre sa structure et les propriétés physiques du bassin. Mais partir de ce choix sur un large échantillon de bassins versants n'est pas possible.

1.3.3.2 Choix 2 : améliorer la représentation des spécificités de l'urbanisation dans la structure du modèle

Il existe plusieurs manières d'améliorer la prise en compte des spécificités du milieu urbain dans les modèles initialement non dédiés aux bassins urbanisés. Pour les modèles adoptant déjà une représentation explicite du milieu urbain, l'objectif est de trouver une manière de représenter la mosaïque de surfaces rurales et urbaines, comme c'est exigé pour les bassins péri-urbains (Braud, Fletcher et Andrieu, 2013). Valeo et Moin (2000) ont modifié la formulation du TOPMODEL (Beven et Kirkby, 1979) pour permettre la représentation à la fois des surfaces urbaines (imperméables) et rurales (perméables) dans leur modèle distribué. Ils ont rajouté une hypothèse principale stipulant que les cellules imperméables ne reçoivent pas de recharge et convertissent toute l'eau qui tombe en ruissellement. Furusho, Chancibault et Andrieu (2013) ont amélioré la prise en compte de l'imperméabilité, du drainage des surfaces imperméables par le réseau artificiel et de l'échange entre le réseau artificiel et la nappe souterraine dans le modèle distribué ISBA-TOPMODEL, initialement dédié aux bassins versants ruraux.

Les exemples qui partent d'une structure de modèle à représentation implicite (conceptuelle, globale) et relativement simple nous sont plus intéressants car ils permettent des tests sur de larges échantillons.

Singh, Maheshwari et Malano (2009) ont complexifié la structure du modèle conceptuel SIMHYD pour prendre en compte les surfaces imperméables et les flux d'origine anthropique (alimentation en eau potable, assainissement, pompage des eaux souterraines). Ils ont testé leur modèle sur un bassin versant situé en Australie d'une taille de 625 km² pour une période de 15 ans au pas de temps journalier. Ils ont montré que l'inclusion des flux anthropiques a amélioré la simulation du bilan hydrologique mensuel sur le bassin. En revanche, leur approche n'a pas suivi une démarche de complexification pas-à-pas (c'est-à-dire justifiée par une amélioration des performances du modèle à *chaque* modification), ce qui conduit à une structure finale lourdement paramétrée.

Kjeldsen, Miller et Packman (2013) ont relié les paramètres du modèle conceptuel et événementiel ReFH à une mesure du couvert urbain. Ils ont testé sur sept bassins versants en Grande Bretagne trois scénarios de complexification graduelle pour la prise en compte des spécificités de l'urbanisation : d'abord dans la composante de ruissellement, ensuite dans les composantes du ruissellement et du routage, et finalement dans les composantes du ruissellement, du routage et du débit de base. Ainsi, ils ont illustré une approche de complexification graduelle

justifiée par les performances (à travers une méthode de calage-validation), par laquelle ils ont réussi à améliorer à la fois la prise en compte des spécificités de l'urbanisation et les performances par rapport au modèle du départ. Leur objectif s'est limité à améliorer la reproduction des hydrogrammes observés uniquement en période de crue, ce qui est moins ambitieux qu'une reproduction d'une large gamme de débits observés à l'exutoire du bassin urbanisé.

Un dernier exemple est donné par Hamel et Fletcher (2014) qui ont complexifié une structure de modèle pour mieux reproduire la réponse d'un bassin versant non urbain en testant trois configurations avec différents niveaux de complexité. Leur objectif était de mieux représenter les processus de subsurface, puis de regarder l'impact des techniques alternatives dans l'atténuation de l'impact de différents scénarios d'urbanisation. Ils ont analysé le comportement hydrologique du bassin étudié *via* différentes signatures hydrologiques (par exemple, les coefficients de ruissellement) avant de développer leur approche de complexification pas-à-pas. Cette analyse de comportement a facilité la définition des modifications à tester.

Nous avons fait le choix de partir d'une structure de modèle relativement simple et de complexifier pas-à-pas cette structure de manière à (1) améliorer le lien entre la structure du modèle de départ et les caractéristiques urbaines, et (2) améliorer les performances du modèle sur les bassins urbanisés. Pour obtenir une structure robuste, nous avons choisi de ne pas focaliser sur un seul bassin versant mais d'utiliser un large échantillon de bassins versants et de regarder l'amélioration pas-à-pas et de manière statistique.

1.4 Synthèse

Les modèles hydrologiques sont d'un grand potentiel pour prédire l'impact de l'urbanisation sur le comportement hydrologique, en bénéficiant de la connaissance des mécanismes de génération de la réponse hydrologique. Ces mécanismes échappent aux approches purement statistiques qui, quant à elles, se limitent à la détection.

Pour relier de manière fiable des scénarios d'urbanisation aux scénarios hydrologiques, nous avons déterminé trois conditions nécessaires pour une structure de modèle hydrologique : (1) avoir un lien avec les propriétés décrivant le niveau d'urbanisation du bassin versant, (2) garantir un équilibre entre le niveau de paramétrisation (qui reflète la demande en données) et le volume de données disponibles, et (3) montrer une bonne reproduction de la réponse hydrologique pour une large gamme de situations d'urbanisation à travers un grand échantillon de bassins versants.

La revue des modèles hydrologiques actuels a montré que :

- Si les modèles à base physique, distribués ou semi-distribués, permettent un lien direct entre les propriétés physiques et les processus hydrologiques du bassin versant urbanisé, ils souffrent de structures lourdement paramétrées nécessitant **(1)** un grand volume de données souvent difficiles à acquérir (description fine des réseaux de drainage artificiel et du paysage urbain/rural, données sur l'alimentation en eau potable et les rejets des stations d'épuration, données hydroclimatiques à haute résolution, voir Braud *et al.*, 2010; Braud, Fletcher et Andrieu, 2013; Jankowsky *et al.*, 2013; Rodriguez, Bocher et Chancibault, 2013) et **(2)** une démarche de détermination des paramètres souvent expertisée.
- Si les modèles conceptuels, globaux ou semi-distribués, sont simples à appliquer, ils souffrent de l'absence d'un lien explicite entre leur structure et les caractéristiques de l'urbanisation.
- Finalement, l'application des modèles adaptés aux bassins versants urbanisés est majoritairement restreinte à un faible nombre de bassins versants, ce qui ne permet pas de juger de leur robustesse, et en particulier de leur transposabilité spatiale et temporelle.

Se présentent alors deux choix méthodologiques de croisement des deux approches de modélisation (explicite et implicite) afin de bénéficier des avantages de chacune. Le premier est de partir d'une structure de modèle à représentation explicite et de la simplifier. Cela nous paraît plus difficile que de partir d'une structure simple et d'améliorer sa prise en compte des caractéristiques urbaines. Pour pallier le troisième défaut sur l'évaluation de la robustesse, nous avons choisi de développer et de tester le modèle sur un large échantillon de bassins versants.

CHAPITRE 2

ÉCHANTILLON DES BASSINS VERSANTS

Sommaire

2.1	Introduction	65
2.2	Données hydroclimatiques	67
2.3	Sélection des bassins versants	77
2.4	Caractéristiques d'occupation du sol et d'urbanisation	89
2.5	Caractéristiques hydroclimatiques de l'échantillon	100
2.6	Caractéristiques morphologiques de l'échantillon	111
2.7	Caractéristiques géopédologiques de l'échantillon	118
2.8	Synthèse	126

Acronymes

BD Base de données. 112

CLC CORINE Land Cover. 90

CLMS Copernicus Land Monitoring Service. 90

COMEPHORE Combinaison en vue de la Meilleure Estimation de la Précipitation HOraire.
71, 77, 79

ETP Évapotranspiration potentielle. 72, 74, 75, 76, 77, 78, 86, 87, 88, 102, 104, 106, 108, 109, 128

FAO Food and Agriculture Organization. 118

GAGES Geospatial Attributes of Gages for Evaluating Streamflow. 68, 77, 78, 79, 92, 113

GLHYMPS GLobal HYdrogeology MaPS. 118

GMTED2010 Global Multi-resolution Terrain Elevation Data 2010. 111

HDMA Hydrologic Derivatives for Modeling and Analysis. 111, 112

HPD Hourly Precipitation Data. 69

HWSD Harmonized World Soil Database. 118

HydroSHEDS Hydrological data and maps based on SHuttle Elevation Derivatives at multiple
Scales. 111

IGN Institut Géographique National. 112

IIASA International Institute for Applied Systems Analysis. 118

ISRIC International Soil Reference and Information Centre. 118

ISS-CAS Institute of Soil Science - Chinese Academy of Sciences. 118

JRC Joint Research Centre of the European Commission. 118

NCEP National Centers for Environmental Predictions. 70, 77

NDVI Normalized Difference Vegetation Index. 91

NHD National Hydrography Dataset. 113

NLCD National Land Cover Database. 68, 89, 90, 91, 101

NOAA/NWS National Ocean and Atmospheric Administration/National Weather Service. 69

ORNL DAAC Oak Ridge National Laboratory Distributed Active Archive Center. 70

SAFRAN Système d'Analyse Fournissant des Renseignements Atmosphériques à la Neige. 72, 73, 77, 79

Sandre Service d'administration national des données et référentiels. 68

SCHAPI Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations. 68

SRTM Shuttle Radar Topography Mission. 111

US United States. 73

USGS United States Geological Survey. 68, 77, 88, 113

UTC Coordinated Universal Time. 68, 70, 74, 75

2.1 Introduction

La revue bibliographique du chapitre précédent a montré que l'utilisation d'un large échantillon de bassins versants urbanisés n'est pas une pratique très répandue. Pourtant, une controverse est toujours d'actualité sur l'impact de l'urbanisation sur les caractéristiques hydrologiques d'un bassin versant et cette controverse ne peut pas être tranchée sans avoir recours à un travail intensif de collecte de données. Un large échantillon permet de mieux apprendre sur le fonctionnement hydrologique des bassins versants urbanisés, en offrant la possibilité de tirer des conclusions générales et transposables à condition de disposer d'une grande diversité de régimes hydroclimatiques (Gupta *et al.*, 2014), conjuguée à une richesse en termes de situations d'urbanisation.

Les atouts du développement de modèles sur un large échantillon sont multiples. En comparant plusieurs modèles sur un large échantillon de données, on cherche une structure qui fonctionne statistiquement le mieux pour plusieurs cas, ce qui permet d'éviter un « sur-ajustement » de la structure où l'on se contenterait d'imiter les idiosyncrasies d'un ou de quelques bassins versants choisis (Kirchner, 2006). L'utilisation d'un échantillon restreint empêche de tester la robustesse de la structure et conditionne la crédibilité du modèle à la qualité des données utilisées pour son développement. Enfin, plus l'échantillon est grand, plus le risque de défaillance du modèle augmente car on le confronte à des comportements

hydrologiques très variés (Andréassian *et al.*, 2009). Ceci facilite l'identification des sources de cette défaillance pour chercher à y remédier.

En partant de ces considérations, nous avons adopté une approche statistique basée sur un large échantillon de données, à l'instar de plusieurs thèses de développement et de test de modèles (par exemple, Ficchi, 2017; Le Moine, 2008). Nous avons poursuivi cette approche dans le cadre de l'étude des bassins versants urbanisés, initiée par les travaux de thèse de Salavati (2015).

Dans le présent chapitre, nous décrivons la procédure de préparation et de sélection d'un large échantillon de bassins versants situés en France et aux États-Unis. Le choix des États-Unis est motivé par la disponibilité de plusieurs bases de données en accès libre, en plus de l'offre d'un grand nombre de bassins versants avec une forte densité du couvert urbain. L'extension de l'échantillon par l'inclusion des bassins situés en France part d'une ambition à moyen/long terme qui permette de tester si les différences en matière d'aménagement urbain, de structuration et de fragmentation du paysage urbain entre les deux rives de l'Atlantique Nord auraient également un impact sur le comportement hydrologique des bassins versants. La motivation à court terme consiste à enrichir la base de données par des bassins avec des spécificités non rencontrées aux États-Unis. Cet enrichissement donnera l'opportunité au modèle de dépasser les problèmes liés aux hétérogénéités des bases de données. L'étude se porte sur la période 1997-2017 car les cas d'urbanisation les plus intéressants sont les plus récents; en outre, la caractérisation de l'occupation du sol sur les deux continents est possible grâce aux bases de données spatiales mises à disposition à partir des années 1990 (1990 pour l'Europe et 1992 pour les États-Unis).

Le présent chapitre est structuré comme suit :

1. Nous commençons par les données hydroclimatiques nécessaires pour caler et tester le modèle (Section 2.2). Nous décrivons les sources et les méthodes de traitement des séries de débit, de précipitation et d'évapotranspiration potentielle, au pas de temps horaire. Pour chacune des séries, nous évoquons d'abord des bassins situés aux États-Unis puis ceux situés en France.
2. Ensuite, nous détaillons les critères de sélection de ces bassins versants pour arriver à l'échantillon final (Section 2.3). Nous avons adopté trois critères : présence limitée de la neige, impact limité des ouvrages de régulation et de rétention, et disponibilité suffisante des séries hydroclimatiques. L'échantillon sélectionné fait l'objet d'une première caractérisation hydrologique sur la base d'un grand nombre d'événements préparés à partir des séries pluie-débit, puis d'une analyse du bilan hydrologique interannuel. Nous finissons cette section par une discussion sur la qualité des données et sur les problèmes inhérents au tri effectué.

3. Puisque nous nous intéressons à l'urbanisation, nous consacrons la Section 2.4 à la caractérisation de l'occupation du sol. En particulier, nous détaillons les mesures de l'urbanisation utilisées pour ces travaux et les méthodes d'agrégation spatiale. À l'aide de ces mesures, nous analysons l'évolution temporelle de l'urbanisation des bassins versants sélectionnés et la diversité de l'échantillon en matière de situations d'urbanisation.
4. Enfin, nous complétons le présent chapitre par les caractéristiques hydroclimatiques (Section 2.5), morphologiques (Section 2.6) et géopédologiques (Section 2.7) de notre échantillon. Dans chaque section, nous décrivons **(i)** les indicateurs utilisés, **(ii)** leur répartition dans l'espace, **(iii)** les spécificités des bassins situés en France par rapport à ceux situés aux États-Unis, et **(iv)** les spécificités des bassins fortement urbanisés par rapport à l'échantillon total.

Nous finissons par une synthèse des caractéristiques de l'échantillon et la définition des sous-échantillons utilisés dans chacun des prochains chapitres.

2.2 Données hydroclimatiques

Un modèle hydrologique décrit de manière simplifiée la relation pluie-débit à l'échelle du bassin versant. Il est donc nécessaire de disposer des séries de pluie pour alimenter le modèle choisi, et de débit pour juger de ses performances. S'agissant d'une composante importante du bilan hydrologique, l'évapotranspiration devrait être prise en compte comme entrée du modèle; puisque nous nous basons sur une approche de modélisation simple, nous représentons ce processus à l'aide du concept de l'*évapotranspiration potentielle*. En tout, trois séries de données hydroclimatiques sont préparées : le débit, la précipitation et l'évapotranspiration potentielle.

Il faut maintenant spécifier la résolution spatiale et temporelle à laquelle nous préparons ces données. Dans le précédent chapitre, nous avons fait le choix d'étudier le comportement hydrologique et la manière dont il est impacté par l'urbanisation à l'échelle du bassin versant. Par conséquent, nous avons adopté une approche de modélisation globale, selon laquelle les données de forçage climatique sont agrégées (moyennées) à l'échelle d'intérêt (c'est-à-dire l'échelle du bassin versant).

Notre choix de la résolution temporelle a été dicté par **(i)** le besoin d'un pas de temps fin pour capturer suffisamment la dynamique de la réponse hydrologique (limite supérieure), et **(ii)** la disponibilité des données de forçage climatique et de débit (limite inférieure). Un compromis entre ces deux exigences a entraîné le choix du pas de temps horaire. À la fin de ce chapitre, nous critiquons ce choix à partir de l'échantillon qui a été sélectionné et caractérisé.

2.2.1 Données de débit

2.2.1.1 Aux États-Unis

Pour les bassins versants nord-américains, nous avons d'abord téléchargé les débits instantanés dans un environnement de développement en R (R Core Team, 2019) en utilisant le package `dataRetrieval` (Cicco *et al.*, 2018) pour les stations de jaugeage répertoriées dans la base de données GAGES-II (Falcone, 2011), maintenue par l'USGS. Le téléchargement a été effectué en deux phases : la première s'est déroulée du 22 décembre 2017 au 17 janvier 2018, puis la deuxième du 17 novembre 2018 au 22 novembre 2018². Nous avons ciblé la période la plus récente du 01/01/1988 au 31/12/2017 ; le choix de ne pas aller avant 1988 est dicté par le plus ancien produit de la base de données NLCD pour l'occupation du sol qui date de 1992. Les débits téléchargés sont enregistrés à différents pas de temps (par rapport à la référence du temps universel UTC) en ft^3/s . Pour préparer les séries au pas de temps horaire, nous sommes passés par deux étapes : (1) interpolation linéaire pour calculer les valeurs de débit instantané à chaque heure ronde (par exemple, 13:00, 14:00, etc.), à condition que la durée maximale entre deux mesures utilisées pour l'interpolation ne dépasse pas 24 heures. Autrement, la valeur instantanée à l'heure ronde est considérée comme lacune ; (2) calcul du volume écoulé entre l'heure ronde $h - 1$ et l'heure ronde h pour donner le débit horaire moyen à l'heure h , comme montré sur la Figure 2.1. Cette intégration est faite à l'aide de la méthode des trapèzes.

2.2.1.2 En France

Les données de débit pour les bassins versants situés en France ont été extraites à partir de la base de données Banque HYDRO (Leleu *et al.*, 2014) qui comporte des données hydrométriques pour plus que 4000 stations (Bourgin *et al.*, 2010). Le travail d'extraction de ces séries de débit a été effectué entre le 04 juin 2018 et le 17 décembre 2018³. Pour télécharger les données, on a eu besoin d'un identifiant de chaque station défini par la nomenclature Sandre⁴. Une partie des données de débit a pu être téléchargée à l'aide de l'application HYDRO2 du SCHAPI. La majeure partie a été téléchargée de manière automatisée à partir de la plateforme www.hydro.eaufrance.fr en se basant sur des fonctions du package `RSelenium` (Harrison et Kim, 2020) dans un environnement de développement en R (R Core Team, 2019). Nous avons

²Au début de la thèse, nous n'étions intéressés qu'aux bassins fortement urbanisés pour lesquels les données ont été téléchargées pendant la première phase. Le besoin d'un échantillon témoin de bassins non urbanisés a engendré une deuxième campagne de collecte de données fin 2018.

³Au départ, l'étape de collecte de données a concerné uniquement les bassins fortement urbanisés pour lesquels le téléchargement des données a duré du 04 juin 2018 au 06 juin 2018 (161 stations). Ayant besoin de bassins versants témoins non urbanisés, nous avons repris le téléchargement à partir du 19 novembre 2018 jusqu'au 17 décembre 2018.

⁴Plus de spécifications sur cette nomenclature sont données sur : <http://www.sandre.eaufrance.fr/definition/HYD/2.1/CdSiteHydro>.

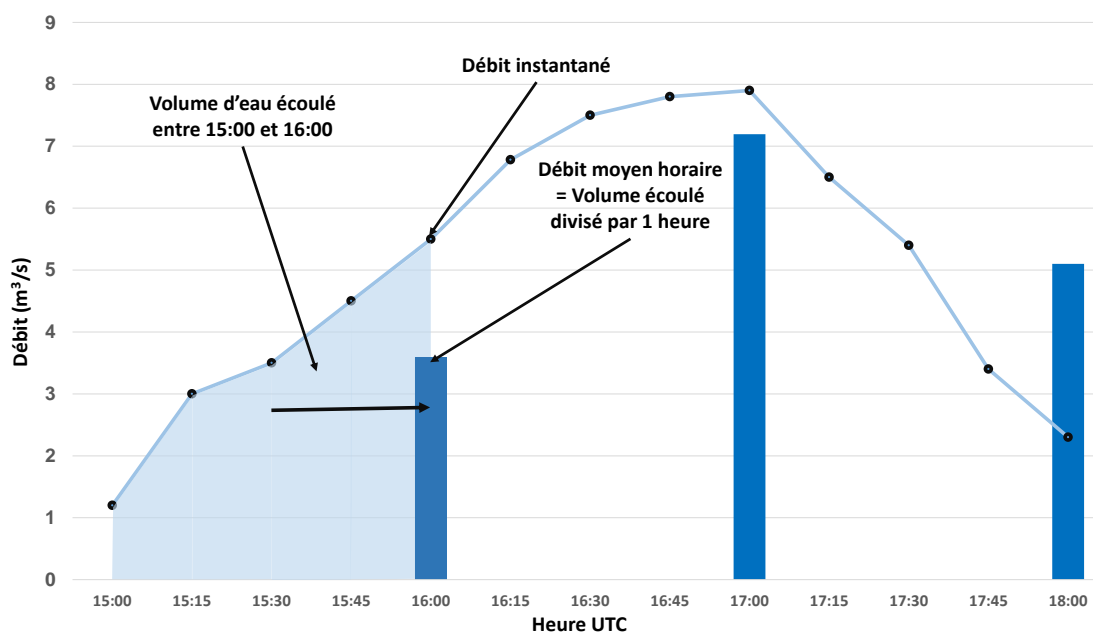


Figure 2.1. Calcul du débit moyen horaire à l'aide des mesures de débit instantanées. Si les mesures aux heures rondes ne sont pas disponibles, on les détermine grâce à une interpolation linéaire. Ensuite, le débit moyen horaire correspondant à l'heure h (ici 16:00) est déterminé en calculant le volume d'eau écoulé entre $h - 1$ et h (ici 15:00 et 16:00). Ce volume est donné par l'aire sous la courbe (approximative de l'intégrale) du débit instantané estimée à l'aide de la méthode des trapèzes.

téléchargé les débits horaires issus d'un traitement QTFIX qui fournit les débits moyens en l/s ou en m^3/s correspondant à un pas de temps fixe (multiple de l'heure et sous-multiple de 24 heures). Ce traitement QTFIX passe par la conversion des mesures hauteur-temps en séries temporelles de débit grâce aux courbes de tarage. Nous avons utilisé ces séries telles quelles, sans traitement supplémentaire.

2.2.2 Données de précipitation

2.2.2.1 Aux États-Unis

Nous avons commencé par collecter les données de précipitation à partir des stations pluviométriques recensées dans la base de données HPD du NOAA/NWS. Ce jeu de données a été comparé à un jeu préparé à partir de la réanalyse de Livneh *et al.* (2015), disponible à une résolution de $1/16^\circ$ (~ 6 km) au pas de temps journalier. Cette comparaison a montré que le jeu de données des stations pluviométriques est d'une qualité insuffisante, liée en partie à une faible densité des stations.

Nous nous sommes alors tournés vers le jeu de données radar du NCEP Stage IV⁵, issu d'une réanalyse combinant les données des pluviomètres et les mesures radar pour produire des lames d'eau au pas de temps horaire (échelle UTC) sur les États-Unis maillées à une résolution de 4 km (Lin et Mitchell, 2005). Ce jeu est disponible à partir de 2002. Nous l'avons téléchargé directement pour chaque bassin grâce au package en R `geoknife` (Read *et al.*, 2015), qui calcule les lames d'eau horaires moyennes pour chaque bassin. Le poids de chaque maille est fonction de l'aire de son intersection avec le polygone du bassin versant.

Dans notre cas, si la présence de lacunes est acceptable pour les données de débit, elle ne l'est pas pour les forçages du modèle. Or, nous avons constaté que les données de radar sur certains bassins versants nord-américains contiennent des lacunes sur plusieurs heures. Pour combler ces lacunes, nous avons fait appel au produit Daymet de l'ORNL DAAC (Thornton *et al.*, 2016), contenant un jeu de précipitations disponible sur la période 1980-2017 au pas de temps journalier sur une résolution de 1 km (sans lacunes). Grâce à la méthode des polygones de Thiessen, nous avons déterminé les lames d'eau journalières Daymet pour un échantillon de 1188 bassins versants. Nous avons estimé le rapport de la précipitation moyenne interannuelle issue de Daymet sur la moyenne de précipitation interannuelle issue de NCEP Stage IV. La médiane est de 1.1, ce qui indique que les deux jeux de données sont comparables avec une surestimation du volume précipité par Daymet par rapport à NCEP Stage IV (écart interquartile : 1.05-1.15).

Le jeu de données Daymet nous a donc servi comme jeu de référence pour combler les lacunes horaires des données de NCEP Stage IV. Ce comblement de lacunes est parti d'un principe simple⁶ :

1. Si la fenêtre lacunaire s'étend sur moins de 6 heures, une interpolation linéaire est effectuée.
2. Sinon, si elle s'étend sur moins de 24 heures, on compare le volume total de précipitations enregistré jusqu'alors par la série horaire avec celui de Daymet. Si ce dernier est dépassé, on attribue à toute la fenêtre lacunaire une valeur de zéro. Sinon, le volume d'eau manquant pour atteindre le volume enregistré par Daymet est distribué uniformément sur les lacunes.
3. Sinon, quand la fenêtre lacunaire dépasse 24 heures, on la comble en distribuant uniformément les lames d'eau journalières de Daymet (c'est-à-dire en divisant la lame d'eau journalière par 24).

Pour limiter l'effet des heures comblées, nous avons pris en considération le nombre total

⁵Voir : <https://data.eol.ucar.edu/dataset/21.093>.

⁶Que nous reconnaissons très discutable.

des heures comblées pour chaque bassin dans la procédure de sélection des bassins versants, comme nous allons le voir plus tard.

2.2.2.2 En France

Pour les bassins situés en France, nous avons utilisé la réanalyse COMEPHORE de Météo France (Tabary *et al.*, 2013), issue d'une fusion des mesures radar et des mesures de précipitation au sol. Une description de la procédure de préparation de cette réanalyse est donnée par Lobligeois (2014, pp. 60-66). Ce jeu s'étend sur la période 1997-2015 et contient les lames d'eau horaires sur toute la France métropolitaine à une résolution de 1 km. Contrairement aux données nord-américaines, cette réanalyse ne contient pas de lacunes. L'agrégation spatiale à l'échelle du bassin versant a été effectuée par la méthode des polygones de Thiessen, ce qui revient à pondérer la lame d'eau enregistrée sur chaque maille par l'aire de son intersection avec le polygone du bassin versant. Un exemple est illustré sur la Figure 2.2.

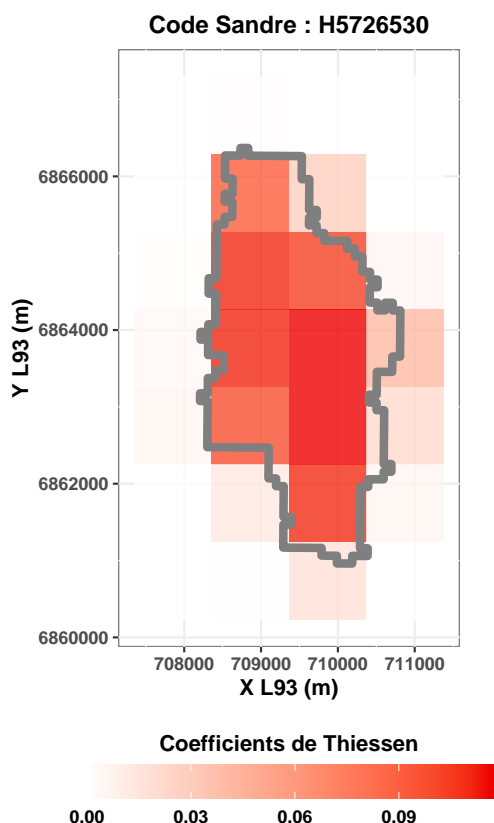


Figure 2.2. Coefficients de Thiessen pour le bassin versant drainé par Le ru de Choqueuse à Doue, d'une taille de 8.75 km². Les couleurs indiquent le poids attribué à chaque maille COMEPHORE pour calculer la moyenne de précipitation pour ce bassin. Le système de projection utilisé est Lambert 93.

2.2.3 Données d'évapotranspiration potentielle

Pour obtenir les données d'évapotranspiration potentielle (ETP) au pas de temps horaire, nous avons d'abord préparé les séries de température au pas de temps journalier à l'échelle du bassin versant. Moyennant une formule basée sur la température journalière (Oudin *et al.*, 2005), nous avons obtenu l'ETP au pas de temps journalier. Enfin, nous avons désagrégé l'ETP journalière pour obtenir une série au pas de temps horaire et ce, grâce à la variabilité temporelle du rayonnement extraterrestre au pas de temps horaire.

2.2.3.1 Séries de température au pas de temps journalier

Pour les bassins versants situés aux États-Unis, nous avons extrait les données de températures journalières minimale et maximale à partir du produit Daymet (Thornton *et al.*, 2016). Comme pour la précipitation, ces données sont disponibles au pas de temps journalier à une résolution de 1 km et couvrent la période 1980-2017. L'agrégation spatiale a été effectuée à l'aide des polygones de Thiessen (Figure 2.3b). La moyenne journalière a été considérée comme la moyenne arithmétique des températures journalières maximale et minimale à l'échelle du bassin versant.

Pour les bassins situés en France, nous avons extrait les données de température à partir du produit SAFRAN de Météo France (Vidal *et al.*, 2010). Ce jeu de données est disponible au pas de temps horaire à partir de 1958, à une résolution de 8 km. Nous avons obtenu la température moyenne à l'échelle du bassin versant par une agrégation spatiale moyennant les polygones de Thiessen (Figure 2.3a), puis nous avons déterminé la température journalière à l'aide d'une moyenne arithmétique des valeurs horaires.

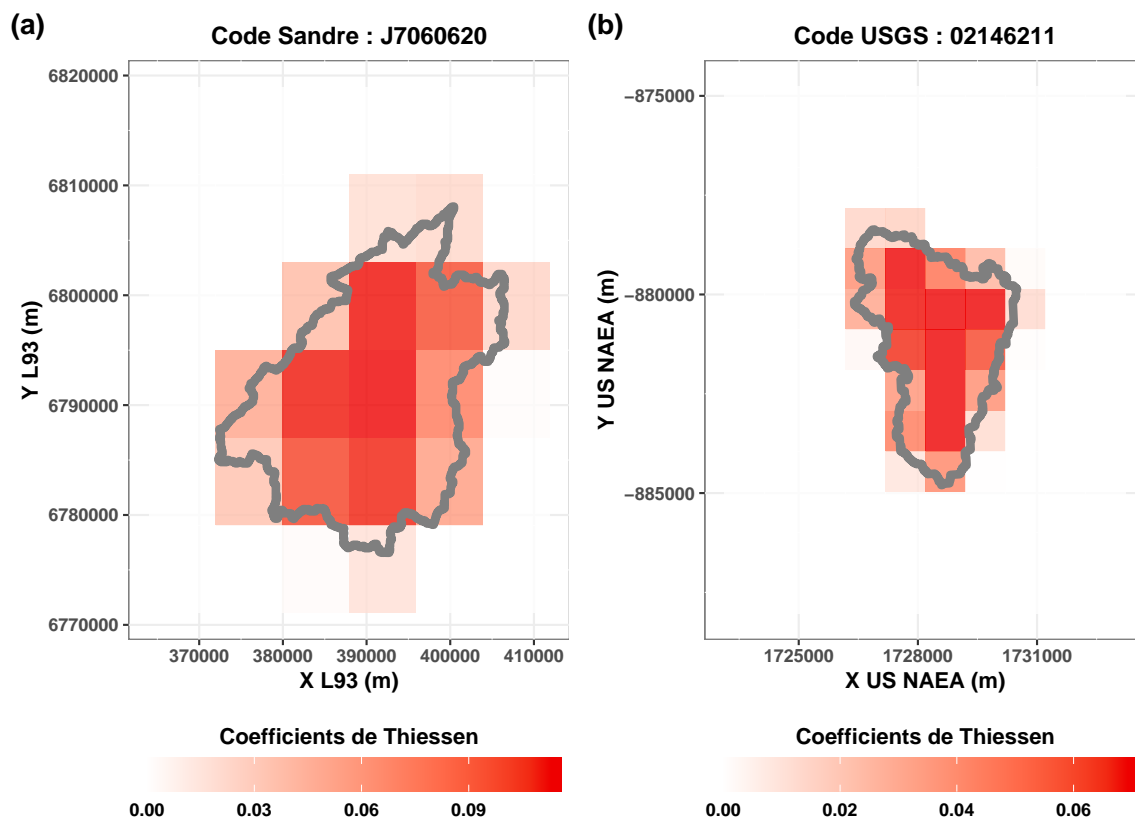


Figure 2.3. Coefficients de Thiessen choisis comme poids pour déterminer la température moyenne à l'échelle du bassin versant. **(a)** Exemple d'un cas situé en France (La Vilaine à Châteaubourg, 565 km²), où les mailles SAFRAN sont d'une résolution de 8 km. La projection utilisée est Lambert 93. **(b)** Exemple d'un cas situé aux États-Unis (Irwin Creek à Statesville Ave, Charlotte, Caroline du Nord, 14.8 km²), où les mailles Daymet sont à une résolution de 1 km. La projection utilisée est US National Atlas Equal Area (NAEA).

2.2.3.2 De la température journalière à l'ETP journalière

Pour déterminer l'ETP au pas de temps journalier, nous avons utilisé la formule proposée par Oudin *et al.* (2005), au vu de sa simplicité et de son adéquation pour les modèles pluie-débit à l'échelle du bassin versant. L'ETP journalière $ETP(J)$ en mm/j est calculée selon l'équation :

$$ETP(J) = \begin{cases} \frac{1000 \cdot Re(\varphi, J)}{\rho \cdot L_E(T_{mean})} \cdot \frac{T_{mean}(J) + 5}{T_0} & \text{si } T_{mean}(J) > -5^\circ C \\ 0 & \text{sinon.} \end{cases} \quad (2.2.1)$$

où $Re(\varphi, J)$ est le rayonnement extraterrestre (en MJ/m²/jr), fonction de la latitude φ du centroïde du bassin versant (en rad) et du jour julien J , $\rho = 998.29$ kg/m³ est la masse volumique de l'eau, $L_E(T_{mean})$ est la chaleur latente de vaporisation de l'eau (en MJ/kg), $T_{mean}(J)$ est la température moyenne journalière de l'air sur le bassin versant (en °C), et $T_0 = 100$ °C.

Pour un bassin versant donné, $ETP(J)$ n'est donc fonction que de la température journalière, les autres variables pouvant être calculées suivant (Allen *et al.*, 1998) :

$$L_E = 2.501 - 0.002361 \cdot T_{mean} \quad (2.2.2)$$

$$Re = \frac{24 \cdot 60}{\pi} \cdot G_{sc} \cdot d_r \cdot (\omega_s \cdot \sin(\delta) \cdot \sin(\varphi) + \sin(\omega_s) \cdot \cos(\delta) \cdot \cos(\varphi)) \quad (2.2.3)$$

$$G_{sc} = 0.082 \text{ MJ} \cdot \text{m}^{-2} \cdot \text{min}^{-1} \quad (2.2.4)$$

$$d_r = 1 + 0.033 \cdot \cos\left(2\pi \cdot \frac{J}{365}\right) \quad (2.2.5)$$

$$\delta = 0.409 \cdot \sin\left(2\pi \cdot \frac{J}{365} - 1.39\right) \quad (2.2.6)$$

$$\omega_s = \arccos(-\tan(\varphi) \cdot \tan(\delta)) \quad (2.2.7)$$

où G_{sc} est la constante solaire (en MJ/m²/min), d_r est la distance relative Terre-Soleil, δ est la déclinaison solaire (en rad), et ω_s l'angle horaire au coucher de Soleil (en rad).

2.2.3.3 De l'ETP journalière à l'ETP horaire

Des approximations sont utilisées pour désagréger l'ETP au pas de temps horaire à partir du cumul journalier. On peut citer l'approximation par une parabole comme l'ont appliquée Lobligois (2014) et Ficchi (2017). Nous sommes partis du même principe en utilisant la distribution du rayonnement extraterrestre horaire (Allen *et al.*, 1998) qui tient compte de la variabilité intra-annuelle du rayonnement et du décalage horaire par rapport à la zone UTC. Cependant, nous estimons que ces corrections auraient peu d'impact sur la performance du modèle dans

la reproduction de la relation pluie-débit, comme indiqué par Oudin *et al.* (2004).

D'abord, pour une heure h (entre 1 et 24) du jour J fixé, nous déterminons :

$$\omega(h, J) = \frac{\pi}{12} \cdot \left(h - 12.5 + \frac{12}{\pi} \cdot (\lambda - \lambda_{UTC}) + S_c(J) \right) \quad (2.2.8)$$

$$\omega_1(h, J) = \omega(h, J) - \frac{\pi}{24} \quad (2.2.9)$$

$$\omega_2(h, J) = \omega(h, J) + \frac{\pi}{24} \quad (2.2.10)$$

$$S_c(J) = 0.1645 \cdot \sin(2b) - 0.1255 \cdot \cos(b) - 0.025 \cdot \sin(b) \quad (2.2.11)$$

$$b = 2\pi \cdot \frac{J - 81}{364} \quad (2.2.12)$$

où ω , ω_1 et ω_2 sont les angles solaires horaires au milieu, au début et à la fin de l'intervalle horaire, λ est la longitude du centroïde du bassin versant (en rad), et λ_{UTC} est la longitude du centre de la zone UTC, prise égale à 0 rad. Finalement, $S_c(J)$ est la correction saisonnière pour l'heure solaire, dépendante uniquement du jour julien J .

Ensuite, pour chacune des heures h de 1 à 24, nous déterminons les coefficients $\beta_{h,J}$ comme suit :

$$\beta_{h,J} = \begin{cases} 0 & \text{si } |\omega(h, J)| > \omega_s \\ (\omega_2 - \omega_1) \cdot \sin(\varphi) \cdot \sin(\delta) + (\sin(\omega_2) - \sin(\omega_1)) \cdot \cos(\varphi) \cdot \cos(\delta) & \text{sinon.} \end{cases} \quad (2.2.13)$$

où ω_s l'angle horaire au coucher de Soleil estimé à partir de l'Équation 2.2.7.

Finalement, $ETP_h(J)$ la valeur de l'ETP (en mm/h) à l'heure h UTC du jour julien J est estimée par :

$$ETP_h(J) = \frac{\beta_{h,J}}{\sum_{\substack{h=1 \\ h \in J}}^{24} \beta_{h,J}} \cdot ETP(J) \quad (2.2.14)$$

où $ETP(J)$ est la valeur de l'ETP au jour J (en mm/j).

Pour illustration, nous montrons sur la Figure 2.4, pour deux périodes différentes de l'année, l'évolution de l'ETP horaire pour deux bassins versants ayant des longitudes différentes.

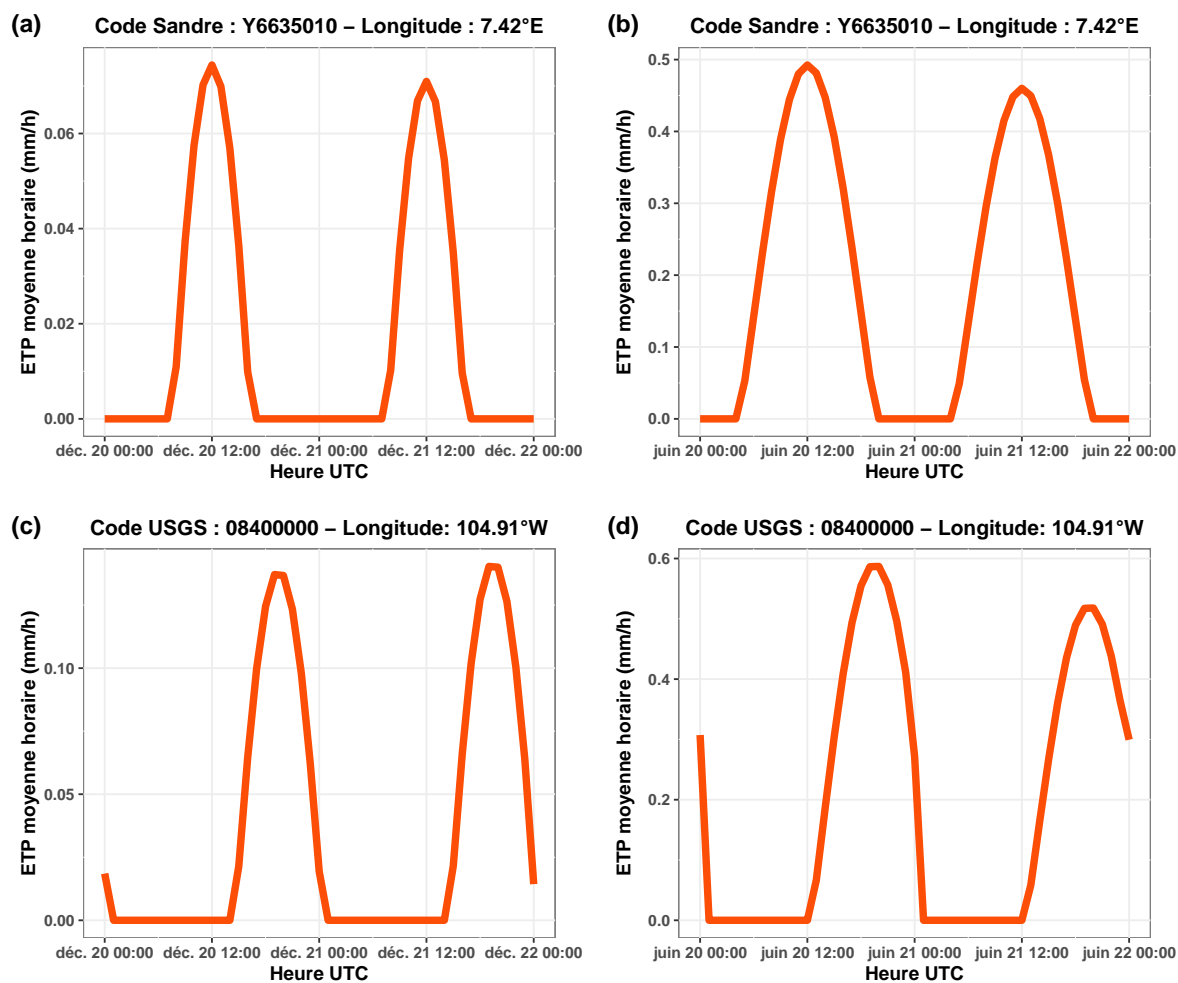


Figure 2.4. Évolution de l'ETP moyenne horaire pour deux bassins versants : (a) et (b) pour un bassin situé en France, à une longitude de 7.42°E (La Bévéra à Sospel), (c) et (d) pour un bassin versant situé aux États-Unis, à une longitude de 104.91°W (Fourmile Draw near Lakewood, Nouveau-Mexique). (a) et (c) au solstice d'hiver, et (b) et (d) au solstice d'été pour l'année 2007.

Tableau 2.1. Récapitulatif des sources de données hydroclimatiques et de leurs périodes de disponibilité pour les bassins situés en France et aux États-Unis.

Pays	Données	Source de données	Résolution spatiale	Période de disponibilité	Période effective
États-Unis	Débit	USGS GAGES-II (Falcone, 2011)	–	1988-2017	2002-2017
	Précipitation	NCEP Stage IV (Lin et Mitchell, 2005) complées à l'aide de Daymet (Thornton <i>et al.</i> , 2016)	4 km pour NCEP Stage IV, 1 km pour Daymet	2002-2017	
	ETP	Daymet (Thornton <i>et al.</i> , 2016)	1 km	1980-2017	
France	Débit	Banque HYDRO (Leleu <i>et al.</i> , 2014)	–	1985-2017	1997-2012
	Précipitation	COMEPHORE (Tabary <i>et al.</i> , 2013)	1 km	1997-2012 ^a	
	ETP	SAFRAN (Vidal <i>et al.</i> , 2010)	8 km	1958-2019	

^a Les données pour 2013-2015 ont été acquises tardivement par rapport à l'étape de préparation de données.

2.2.4 Récapitulatif des sources et des disponibilités de données hydroclimatiques

L'échantillon de données hydroclimatiques est donc construit à partir d'une multitude de sources, avec des différences de résolution spatiale entre la France et les États-Unis. Le Tableau 2.1 dresse un récapitulatif de la provenance de chacune des données avec sa période de disponibilité. Nous remarquons que les séries s'étendent entre 1997 et 2017 pour les deux pays, avec une longueur maximale de 16 ans. Les précipitations et l'ETP étant en général sans lacune, le taux de lacunes dans les données de débit sera un critère principal pour la sélection des bassins versants.

2.3 Sélection des bassins versants

Dans cette section, nous allons détailler la procédure qui nous a permis de sélectionner notre échantillon de 2225 bassins versants à partir d'un échantillon initial de 12768 bassins versants (3701 bassins français et 9067 bassins nord-américains). Nous commençons par

énoncer les critères de sélection, puis nous expliquons quels bassins n'ont pas été retenus. Une caractérisation du bilan et du comportement hydrologique sera montrée pour ces 2225 bassins versants, suivie par une discussion sur le tri des données.

2.3.1 Critères de sélection des bassins versants

Pour sélectionner les bassins versants qui nous intéressent, nous nous sommes appuyés sur trois critères.

Nous avons écarté les bassins versants ayant une forte influence de la neige. Bien que cela rajouterait à la diversité de l'échantillon, en ayant des cas situés dans des régions froides (Sillanpää et Koivusalo, 2015), la prise en compte des processus de fonte et de routage de la neige constitue une dimension de complexité supplémentaire que nous laissons pour de futurs travaux. Nous avons donc choisi les bassins versants pour lesquels la fraction en moyenne de précipitations sous forme solide ne dépassait pas 10%. Cette fraction a été fournie par la base de données GAGES-II pour les bassins nord-américains (Falcone, 2011), et a été considérée comme le pourcentage de précipitations sous forme de neige sur la période 1901-2000. Pour les bassins versants français, cette fraction constitue une sortie du modèle CemaNeige (Valéry, 2010; Valéry, Andréassian et Perrin, 2014), dont les estimations pour chaque bassin versant ont été fournies par Bourgin *et al.* (2010).

Nous avons également enlevé les bassins versants ayant une influence importante des barrages. Même si la construction des ouvrages de stockage fait partie des aménagements anthropiques qui impactent le fonctionnement hydrologique des bassins versants (FitzHugh et Vogel, 2011), la prise en compte de cet impact dans les modèles hydrologiques est une question qui devrait être traitée en soi. Pour nous, l'inclusion des bassins versants avec une influence importante des barrages-réservoirs aurait exigé un choix supplémentaire d'une technique de naturalisation de leur réponse avant de les étudier. Pour quantifier l'impact des barrages-réservoirs, pour chaque bassin, nous avons converti le volume total stockable par les barrages présents à l'amont de son exutoire en une lame d'eau moyenne interannuelle (volume divisé par la surface du bassin versant). Ensuite, nous avons rapporté cette lame d'eau à la lame d'eau observée à l'exutoire durant la période de disponibilité des données de débit. Enfin, nous avons écarté les bassins pour lesquels ce rapport dépassait 5%. Les volumes des barrages pour chaque bassin sont fournis par la base de données GAGES-II pour les bassins nord-américains (Falcone, 2011). Les volumes stockés par les barrages français ont été recensés par Payan (2007).

Le troisième critère concerne la disponibilité de données hydroclimatiques. Nous avons exigé un minimum de 8 ans de données de débit, de précipitation et d'ETP. Aucune lacune de précipitations ou d'ETP n'est permise. De plus, le taux de lacunes en données horaires de débit

ne doit pas dépasser 30% de la longueur totale de la série. Le choix de 8 ans comme minimum se base sur l'exigence d'une période suffisamment longue pour obtenir par calage un jeu de paramètres stable et robuste (350 jours tirés aléatoirement à partir d'une série de débit plus longue, comme recommandé par Perrin *et al.*, 2007 ; 5 ans recommandés par Merz, Parajka et Blöschl, 2009). Pour les bassins nord-américains, le problème se pose également au niveau des données de précipitation, dont une partie a été comblée. Nous avons alors fait le choix d'écarter les bassins versants ayant plus que 700 heures comblées ou n'ayant pas satisfait les exigences en débit. En France, des bassins versants frontaliers ont été écartés car les données SAFRAN ou COMEPHORE ne les couvraient pas intégralement.

Des différences de temps de préparation des données en France en comparaison avec les États-Unis ont impliqué des différences dans l'enchaînement des étapes de la sélection des bassins versants entre les deux pays. En revanche, tous les bassins versants sélectionnés respectent les critères mentionnés ci-dessus. Dans la suite, nous détaillons les étapes de sélection pour chacun des sous-échantillons.

2.3.2 Étapes de sélection

Pour les bassins versants nord-américains, nous sommes partis d'un échantillon de 9067 bassins versants dont les contours sont fournis par la base de données GAGES-II (Falcone, 2011), qui répertorie 9322 stations hydrométriques aux États-Unis. Les critères ont été appliqués de la manière suivante :

- Nous avons écarté **5886/9067** bassins versants dont l'influence de la neige était importante (c'est-à-dire dépassant 10% du volume total de précipitations), pour ne garder que **3181** bassins ;
- Nous avons écarté **1439/3181** bassins versants dont l'influence des ouvrages de stockage était importante (c'est-à-dire dépassant 5% de l'écoulement moyen annuel), pour ne garder que **1742** bassins ;
- Nous avons écarté **452/1742** bassins versants dont les séries de débit étaient soit absentes soit trop courtes (c'est-à-dire plus courtes que $0.7 \cdot 8 = 5.6$ ans) entre 1988 et 2017, ou contenaient des valeurs négatives (influence marine), pour ne garder que **1290** bassins ;
- Nous avons écarté **102/1290** bassins versants dont les séries de débits avaient un taux de lacunes supérieur à 30% entre 2002 et 2017, pour ne garder que **1188** bassins ;
- Les données de précipitations ont pu être téléchargées et comblées pour les **1188** bassins versants restants. Pour limiter l'effet des lames d'eau comblées, nous avons

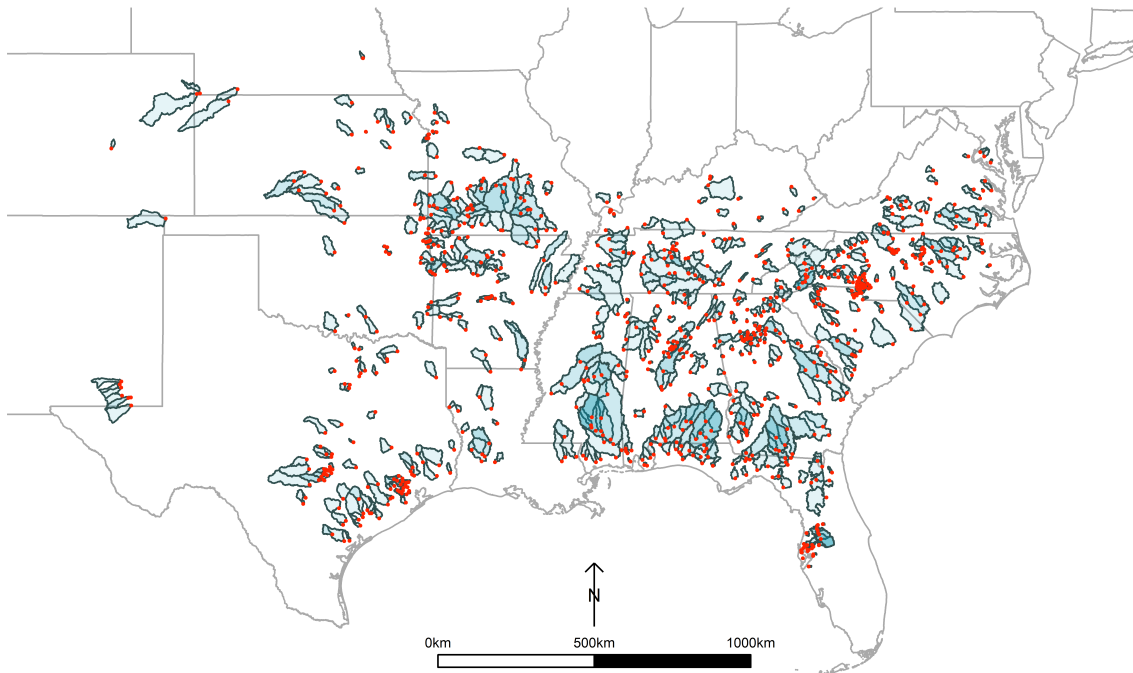


Figure 2.5. Contours (en bleu) et localisations des exutoires (points rouges) des 870 bassins versants nord-américains.

écarté **318/1188** bassins versants pour lesquels le nombre d'heures lacunaires des précipitations dépassait 700 heures, pour ne garder que **870** bassins.

Au final, 870 bassins versants situés aux États-Unis ont été retenus. La plupart de ces bassins ont des longueurs de séries de 16 ans (677/870). Le taux de disponibilité en débit varie entre 70.2% et 99.9%, avec une médiane de 95.4% (écart interquartile : 87.1%-98.3%). Le pourcentage de lacunes des précipitations comblées était faible, et ne dépassait pas 0.8% de la longueur totale de la période. La taille des bassins versants va de 1.1 km² jusqu'à 21290 km², avec une médiane de 323 km². Ils se situent principalement aux États-Unis du sud-est, une région qui a connu un développement urbain très important au cours du dernier siècle par rapport au reste des États-Unis (O'Driscoll *et al.*, 2010). Les contours des bassins versants nord-américains et leurs exutoires sont montrés sur la Figure 2.5.

Pour les bassins versants situés en France, nous sommes partis de la base de données constituée par Bourgin *et al.* (2010), qui contient les contours de 3701 bassins versants. Nous avons procédé à l'application des critères de la manière suivante :

- Nous avons écarté **1004/3701** bassins versants n'ayant pas des séries de débit suffisamment longues entre 1985 et 2017, en se basant uniquement sur leurs dates de mise en/hors service, pour ne garder que **2697** bassins ;
- Nous avons écarté **553/2697** bassins versants dont l'impact de la neige a été jugé important (c'est-à-dire dépassant 10% du volume de précipitations total), pour ne

garder que **2144** bassins ;

- Nous avons lancé le téléchargement de données de débit à partir de la banque HYDRO pour **2144** bassins versants. Ceci a permis de vérifier la disponibilité des débits entre 1985 et 2017, ce qui a permis d'écartier **198/2144** bassins versants, pour ne garder que **1946** bassins ;
- Les données de débit ont été préparées pour les **1946** bassins versants restants, parmi lesquels **387/1946** n'ont pas satisfait le critère de disponibilité de débit entre 1997 et 2012. Nous n'avons donc gardé que **1559** bassins ;
- Les débits étant préparés, l'influence des barrages a pu être estimée. Nous avons alors écarté **164/1559** bassins versants ayant une influence supérieure à 5%, pour ne garder que **1395** bassins ;
- Enfin, nous avons écarté **40/1395** bassins versants qui présentaient des lacunes dans leurs séries de température, pour ne garder que **1355** bassins.

Au final, 1355 bassins versants situés en France ont été retenus, dont 971 ont 16 ans de données au pas de temps horaire. Le taux de disponibilité en débit varie entre 70.6% et 100%, avec une médiane de 97.6% (écart interquartile : 86.0%-99.8%). La taille des bassins versants va de 0.48 km² jusqu'à 110000 km², avec une médiane de 183 km². Ils couvrent la majeure partie du territoire métropolitain (hors zones de montagne). Les contours des bassins versants français et leurs exutoires sont montrés sur la Figure 2.6.

Ce large échantillon de 2225 bassins versants français et nord-américains couvre des régimes climatiques assez variés. La plupart des bassins versants nord-américains se situent dans une région caractérisée par un climat tempéré, humide subtropical (type Cfa dans la classification Köppen-Geiger, voir Beck *et al.*, 2018). Quelques-uns se situent dans les états du Texas et du Nouveau-Mexique au sud, et d'autres drainent une partie des états du Colorado et du Kansas au nord. Ces régions centrales sont caractérisées par un climat semi-aride chaud ou froid (type BSh ou BSk). En France, la majeure partie des bassins versants sont caractérisés par un climat tempéré océanique (Cfb), alors qu'une partie non négligeable se situe dans un climat méditerranéen (Csa). Nous parlerons plus en détail des caractéristiques hydroclimatiques de l'échantillon dans la Section 2.5. Pour le moment, nous exploitons les séries hydroclimatiques préparées pour regarder les caractéristiques de la réponse hydrologique des bassins à travers un échantillon des événements sélectionnés et un calcul du bilan hydrologique.

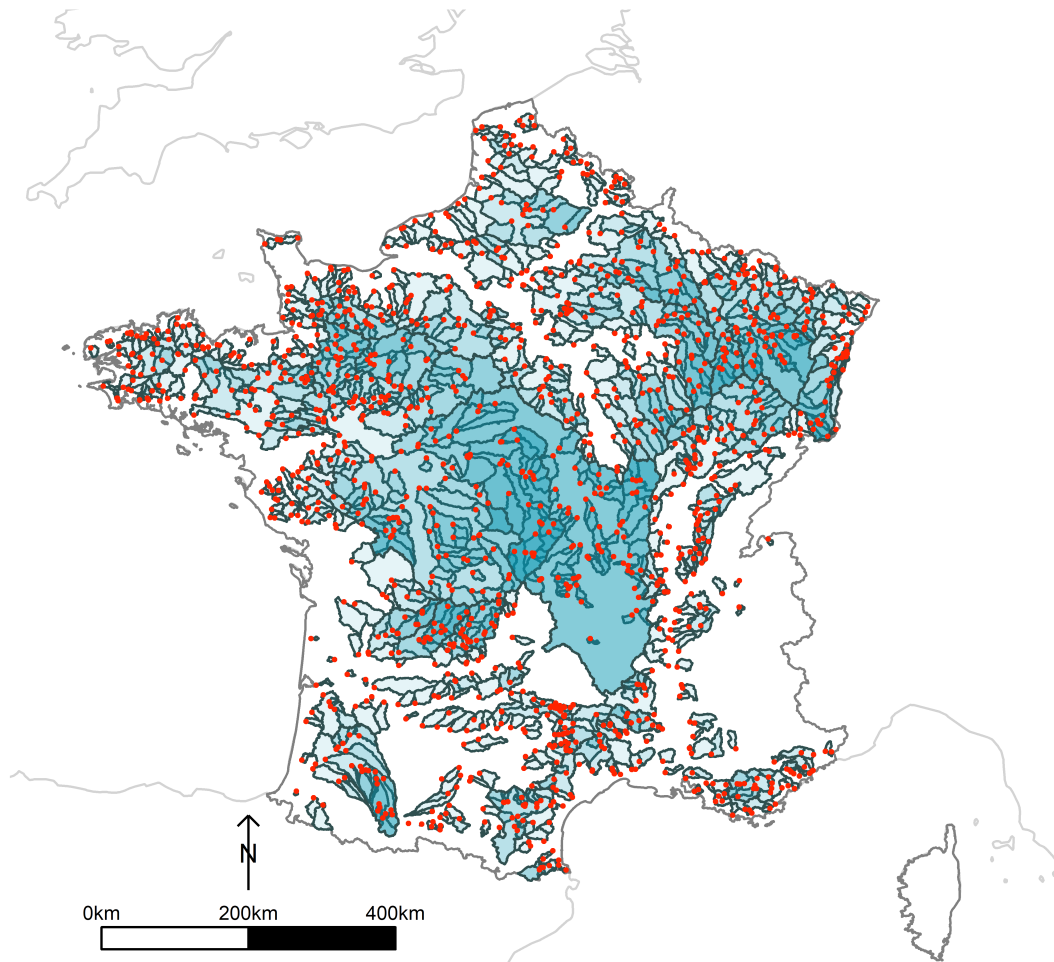


Figure 2.6. Contours (en bleu) et localisations des exutoires (points rouges) des 1355 bassins versants français.

2.3.3 Aperçu des caractéristiques événementielles et des bilans hydrologiques des bassins versants sélectionnés

2.3.3.1 Caractérisation de la réponse hydrologique à partir des événements sélectionnés

Les événements sont des périodes très riches en informations sur le comportement hydrologique d'un bassin versant. Quand il y a une crue après une précipitation, c'est l'opportunité de voir le bassin « dévoiler son caractère », avec une réponse mise en évidence par l'événement pluviométrique. Mais délimiter les événements est un travail qui n'obéit à aucune méthode consensuelle.

La taille de l'échantillon nous a amenés à adopter une méthode automatique, avec plusieurs contraintes à satisfaire par les événements, basée sur la méthode de Lobligeois (2014) mais adaptée pour permettre de cibler les événements dont la réponse prenait origine dans les écoulements de surface, et d'écarter les événements dont la contribution de la nappe était importante. La méthode est détaillée dans l'Annexe A. Ici, nous ne donnons que les grandes lignes. La première étape consiste en une séparation verticale pour extraire le débit de surface du débit total, itérativement à l'aide d'un filtre numérique (Eckhardt, 2005). Cette séparation de débit de base (que nous détaillons dans la Section 2.5 pour l'indice de l'écoulement de base) permet d'estimer le débit direct (la différence entre le débit total et celui de base), que l'on considérera comme la réponse du bassin par l'écoulement de surface à un événement de précipitation. La deuxième étape consiste en une séparation horizontale : une fois la série temporelle de débit direct préparée, on l'examine itérativement pour sélectionner à chaque fois le pic le plus important. L'événement correspondant étant choisi, le pic est enlevé de la série et on reprend la recherche d'un nouveau pic. En fonction du pic, des seuils de coupure sont estimés pour saisir le début et la fin de l'événement au voisinage du pic. Ce voisinage est défini par des fenêtres temporelles qui sont estimées pour chaque bassin à l'aide d'une part des corrélations entre le débit direct et les précipitations, et d'autre part de l'autocorrélation du débit direct. Cette méthode a généré 179866 événements, dont 70227 événements ont eu lieu sur les bassins nord-américains et 109639 sur les bassins situés en France.

Pour caractériser chaque événement, nous avons choisi les métriques suivantes :

1. La durée de l'événement (en heures), qui est le délai entre le début et la fin de l'événement. Plus la durée est grande, plus l'événement est long.
2. L'asymétrie de l'hydrogramme de l'événement ($-$), déterminée comme le rapport entre la durée de la montée de l'hydrogramme de surface (différence entre le moment du pic de surface et le début de l'événement) et la durée totale de l'événement. L'asymétrie varie entre 0 et 1, les valeurs proches de zéro étant observées pour les événements dont la

montée est très rapide.

3. Le délai de la réponse (en heures), pris comme le délai entre le centre de gravité du hyétogramme et le centre de gravité de l'hydrogramme du débit direct. Il s'agit d'une caractéristique de la dynamique de la réponse du bassin versant. De faibles délais caractérisent les bassins à réaction rapide.
4. En plus du délai de la réponse, nous avons estimé un indicateur similaire mais basé uniquement sur le débit direct, à savoir le temps caractéristique événementiel (en heures). Il est défini comme le rapport entre le volume d'eau ruisselée en surface (en mm) sur le pic de débit de surface (en mm/h). Ce temps a été utilisé par Tarasova *et al.* (2018a) pour caractériser la rapidité de la réponse. Nous l'utilisons comme alternative au délai de la réponse.
5. Le volume total d'eau précipitée (en mm), défini par la somme des lames d'eau précipitées entre le début et la fin de l'événement.
6. L'intensité moyenne de précipitation (en mm/h), prise comme le rapport du volume total d'eau précipitée sur la durée totale de l'événement. De grandes intensités sont observables pour des événements orageux.
7. La précipitation maximale horaire (en mm/h), prise comme la lame d'eau maximale précipitée sur une durée d'une heure pendant l'événement. Elle caractérise l'ampleur de l'événement.
8. Le coefficient de ruissellement (-), calculé comme le rapport du volume d'eau ruisselée à partir de la surface (c'est-à-dire le volume du débit direct) sur le volume d'eau précipitée. Un important coefficient de ruissellement est associé à une conversion importante des précipitations en ruissellement de surface.

Les distributions de chacune de ces caractéristiques sont montrées sur la Figure 2.7.

Nous remarquons que les bassins versants contribuent non uniformément à l'échantillon des événements, mais la contribution maximale par bassin reste négligeable par rapport à la taille de l'échantillon. Les événements durent en moyenne un peu plus que 2 jours, et nous observons que les événements aux bassins versants nord-américains sont plutôt courts par rapport aux événements ayant lieu aux bassins français, même si la taille médiane des bassins nord-américains dépasse la taille médiane des bassins français. Ceci est confirmé par le délai de la réponse du bassin versant : en médiane, les bassins mettent 15 heures, en France 17 heures contre uniquement 9 heures aux États-Unis. Au passage, nous remarquons que la plupart des événements ont un délai de la réponse plus court qu'un jour (74% des événements). La majorité des événements ont un délai supérieur strictement à une heure (97% des événements). Le temps caractéristique événementiel, calculé à partir du volume ruisselé et du pic de débit

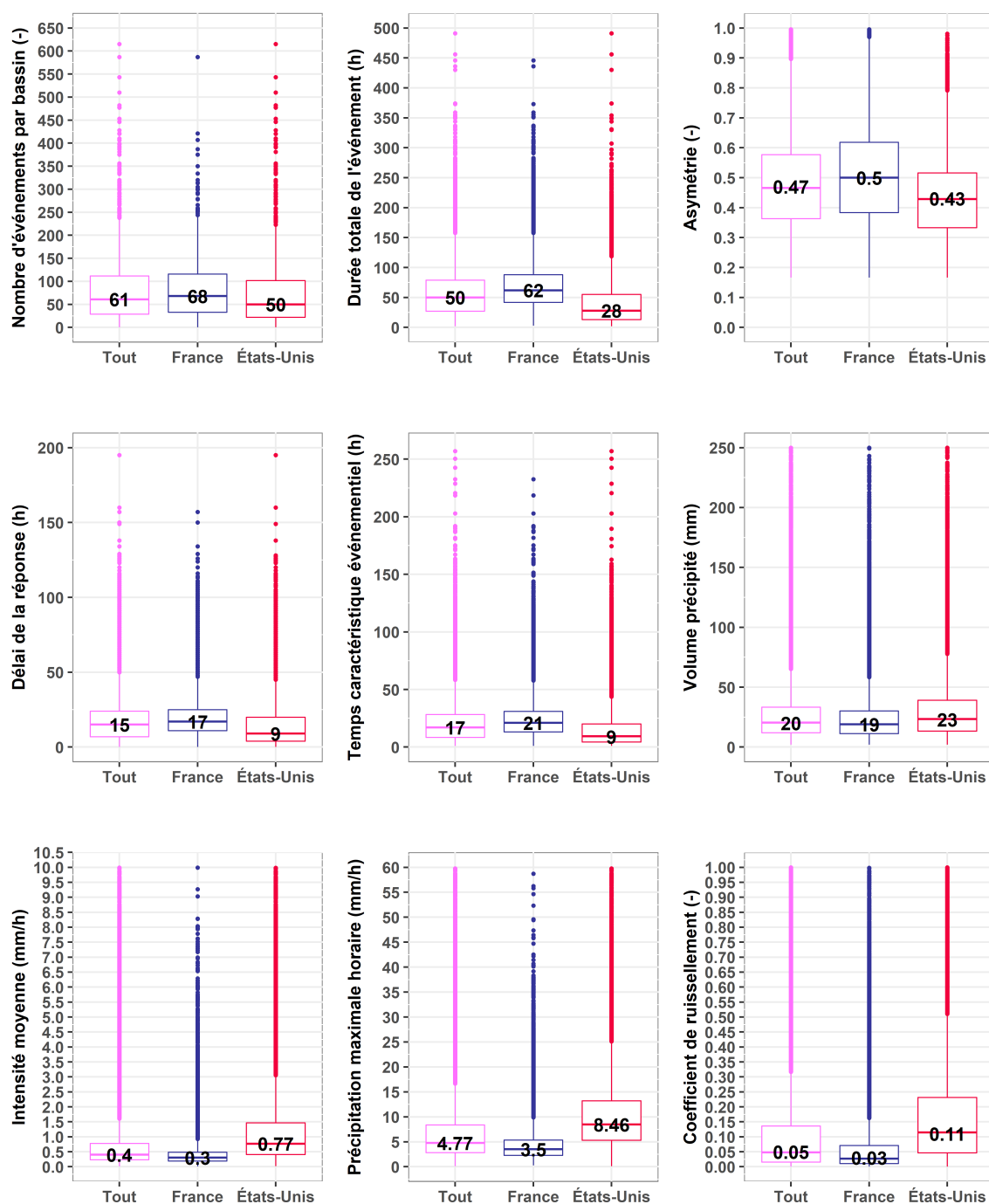


Figure 2.7. Distributions des caractéristiques des événements extraits des séries pluie-débit des 2225 bassins versants. Pour chacune des caractéristiques, les distributions sont montrées pour l'échantillon total (en magenta), français (en bleu) et nord-américain (en rouge). Les chiffres indiquent les valeurs médianes. Pour quelques caractéristiques (délai de la réponse, volume précipité, intensité moyenne et précipitation maximale horaire), les limites du graphique sont tronquées pour faciliter la visualisation, enlevant au plus 0.5% du nombre total des événements.

de surface, affiche des valeurs similaires à celles du délai de la réponse : 17 heures en médiane pour tout l'échantillon, 21 heures en France et 9 heures aux États-Unis. Les événements pluviométriques ont engendré des volumes précipités allant jusqu'à 19-23 mm en médiane, selon des intensités moyennes de 0.3-0.77 mm/h en médiane, avec des événements plus intenses aux États-Unis qu'en France. Ceci est confirmé par les précipitations maximales horaires, dont la médiane est de 8.46 mm/h aux États-Unis contre 3.5 mm/h en France. 40% des événements aux États-Unis et 5% des événements en France ont des précipitations maximales horaires qui dépassent 10 mm/h, limite au-dessus de laquelle on peut considérer qu'il s'agit d'un événement à forte précipitation. Les coefficients de ruissellement sont très faibles en médiane en France par rapport aux bassins nord-américains, ce qui est attribué en partie aux différences en termes d'intensité de précipitation.

Nous concluons donc que la majeure partie des événements sélectionnés montrent que nos bassins versants doivent être étudiés avec un pas de temps horaire. Nous constatons également des différences importantes entre les échantillons des événements français et nord-américains, ce qui contribue à la richesse de l'échantillon total.

2.3.3.2 Bilan hydrologique

Pour caractériser la diversité du régime hydrologique des bassins versants sélectionnés, nous montrons sur la Figure 2.8, pour les 2225 bassins versants, le rendement moyen du bassin versant (c'est-à-dire le rapport du débit moyen interannuel sur la précipitation moyenne interannuelle Q_m/P_m) en fonction de l'indice d'humidité (rapport de la précipitation moyenne interannuelle sur l'ETP moyenne interannuelle P_m/ETP_m).

On constate une diversité importante des régimes hydroclimatiques rencontrés dans l'échantillon. La majeure partie des bassins versants (**2008/2225**) satisfait l'inégalité $P_m - ETP_m < Q_m < P_m$, indiquant des volumes de débit plus faibles que les précipitations et au-dessus du volume minimal disponible pour l'écoulement (c'est-à-dire $P_m - ETP_m$). Notons la présence de quelques bassins versants arides situés principalement aux États-Unis, dont les indices d'humidité sont faibles (< 1) et les réponses négligeables (c'est-à-dire $Q_m \sim 0$). En revanche, une partie non négligeable des bassins versants français (**150**) et nord-américains (**14**) ont des débits moyens plus faibles que la différence $P_m - ETP_m$ (correspondant aux points au-dessous de la courbe en traits jaunes, Figure 2.8), ce qui pourrait être une indication de la présence de pertes importantes vers l'extérieur du système, autrement que par évapotranspiration. Par ailleurs, une faible proportion de l'échantillon a des sorties plus importantes que les entrées (**3** bassins versants nord-américains et **50** bassins situés en France). Ceci pourrait être en partie lié aux apports des réservoirs souterrains, en particulier la présence de sources karstiques (par exemple, La Sorgue à Fontaine-de-Vaucluse V6155010/V6155020, la source karstique la plus importante en France), ou la présence des agglomérations à l'amont du bassin versant

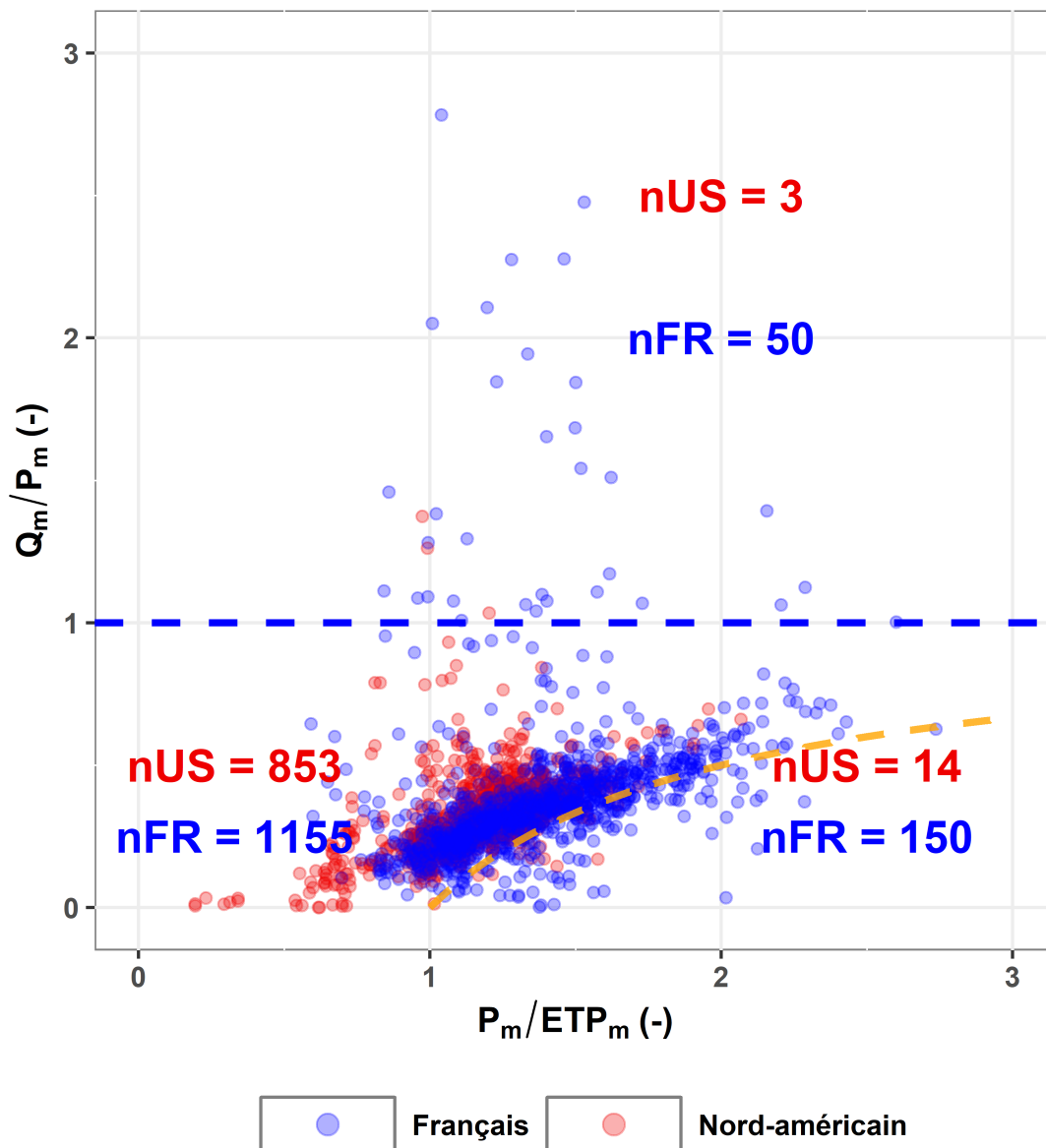


Figure 2.8. Bilan hydrologique interannuel des 2225 bassins versants sélectionnés, où Q_m/P_m indique le rendement moyen, et P_m/ETP_m l'indice d'humidité. Q_m , P_m et ETP_m sont les moyennes interannuelles de débit, de précipitation et d'ETP. Le trait pointillé bleu indique la limite $Q_m = P_m$, ce qui équivaut à une conversion complète de la précipitation en débit. Le trait pointillé jaune indique $Q_m = P_m - ETP_m$. Les points se trouvant au-dessus du trait pointillé bleu ont un débit plus important que la précipitation (c'est-à-dire, $Q_m > P_m$). Les points se situant au-dessous du trait pointillé jaune ont un débit plus faible que le minimum disponible (c'est-à-dire, $Q_m < P_m - ETP_m$). Le reste des points satisfont $P_m - ETP_m < Q_m < P_m$. En bleu sont indiqués les bassins versants situés en France, alors que les bassins versants situés aux États-Unis sont indiqués en rouge. nUS est le nombre de bassins versants nord-américains et nFR est le nombre de bassins versants français. Noter que 17 bassins versants ont eu un Q_m/P_m supérieur à 3 et ne sont pas représentés sur la figure.

(USGS 02310525 Weeki Wachee near Brooksville, Floride, ou USGS 08116400 Dry Creek near Rosenberg, Texas). Doit-on éliminer ces bassins dont le bilan hydrologique est suspect ?

2.3.4 Est-il nécessaire d'effectuer un tri supplémentaire dans les données ?

Quelques bassins versants présentent des cas de données suspectes, dont une cinquantaine avait des volumes de débit plus importants que les volumes de précipitation (Figure 2.8). De plus, l'examen visuel des séries de précipitation-débit a révélé l'existence de 69 bassins versants avec quelques problèmes d'interpolation dans les séries de débit ou avec une évolution relativement lente de l'hydrogramme (3.1% de l'échantillon), 23 bassins versants avec une réponse fortement bruitée par des variations à haute fréquence (1%), et 93 bassins versants avec une réponse bruitée par des variations à haute fréquence mais très faibles par rapport au signal moyen (4.2%). Au total, l'examen visuel des séries a montré qu'environ 185 cas (± 10 bassins versants) présentaient des inconsistances présumées sur une partie plus ou moins importante de leurs séries de débit.

Les erreurs qui entachent la qualité des données sont de plusieurs types : erreurs d'interpolation dans les séries de débit, incertitudes dans les courbes de tarage, dysfonctionnement des stations de mesure souvent au lendemain d'une crue destructrice, problèmes de délimitation du bassin versant, représentativité des données de précipitation, choix de la formule de l'ETP... Les bilans douteux peuvent aussi provenir de l'influence anthropique, comme le rajout d'une entrée autre que la précipitation (apports externes se manifestant par des rejets de stations d'épuration), ou par l'effet des ouvrages de régulation. Dans les travaux de développement de modèles, la qualité des données a toujours posé problème, même pour des travaux utilisant un échantillon relativement large. Par exemple, Lobligois (2014) a analysé l'impact de la résolution spatiale des données de précipitation sur la représentation des processus hydrologiques par un modèle pluie-débit ; la qualité des données d'entrée était donc d'une importance cruciale pour juger de la sensibilité des modèles à la résolution spatiale des forçages. En outre, Ficchi (2017) a rajouté des critères de qualité dans la sélection de l'échantillon qu'il a utilisé afin de développer une structure de modèle fonctionnant à différents pas de temps. Là encore, l'objectif impose un examen de la qualité des données et un rejet de celles dont la représentativité temporelle est mise en doute. En revanche, Le Moine (2008) a estimé que l'effet de la présence de données suspectes est atténué par la taille de l'échantillon, et ne devrait pas poser problème quand on travaille dans un cadre de comparaison de modèles.

Pour notre cas, nous rejoignons Le Moine (2008) dans la mesure où l'utilisation d'un large échantillon permet de considérer comme aléatoires les erreurs dans les données. De plus, le cadre comparatif que nous mettons en place fait que la présence de données suspectes n'est

pas si décisive dans le choix du modèle optimal, en partant du postulat qu'il n'y a pas meilleur modèle que les données elles-mêmes. Alors si celles-ci sont mauvaises, tout modèle échouerait à les représenter. Un obstacle vis-à-vis du jugement de la qualité des séries hydrologiques obtenues est lié à l'absence d'un critère de qualité objectif qui nous aiguillerait dans la sélection et le rejet des bassins versants à données erronées, surtout quand il s'agit de bassins sous une influence anthropique importante, comme dans le cas de l'urbanisation. Par exemple, si l'on tente d'écartier ceux qui ont un rendement supérieur à 1, on risque d'exclure des bassins urbanisés dont le régime est impacté et dont l'inclusion serait bénéfique au modèle que nous cherchons. Enfin, la dépendance de la qualité du modèle à la qualité des données est atténuée par l'utilisation d'un large échantillon. Ce problème se pose donc plus concrètement pour un travail se limitant à représenter le fonctionnement d'un ou de quelques bassins versants uniquement. Nous estimons même qu'on *devrait* inclure, en tant que benchmark, des données suspectes ou irréalistes dans le développement des modèles; au cas où le modèle réussirait à les reproduire, on serait face à des cas où le modèle fonctionne bien pour les mauvaises raisons (Kirchner, 2006), ce qui complète l'exercice de diagnostic du modèle.

Nous n'appliquerons donc pas de critère de sélection supplémentaire, sauf quand un objectif précis l'exigera (par exemple, ne choisir que les bassins fortement urbanisés). Nous allons dans la suite caractériser l'échantillon des 2225 bassins versants, et décrire leurs spécificités, en particulier l'occupation du sol et le niveau d'urbanisation.

2.4 Caractéristiques d'occupation du sol et d'urbanisation

Après avoir détaillé les sources de données hydroclimatiques pour le développement des modèles, puis les étapes de sélection des bassins, nous regardons dans cette section l'évolution de l'occupation du sol des bassins sélectionnés. Nous commençons par expliciter les métriques utilisées. Ensuite, nous regardons les différences d'occupation du sol entre les bassins français et nord-américains. Enfin, nous montrons l'évolution sur la période d'étude de l'urbanisation des bassins versants sélectionnés.

2.4.1 Mesures utilisées

2.4.1.1 Bases de données

Pour caractériser l'occupation du sol aux États-Unis, nous avons utilisé les bases de données NLCD disponibles pour les années 2001, 2006, 2011 et 2016, qui décrivent l'occupation du sol par 21 classes avec une résolution de 30 m. De plus, NLCD fournit des cartes d'imperméabilité

pour les années 2001, 2006, 2011 et 2016, avec la même résolution que les cartes d'occupation du sol. Cette imperméabilité traduit le pourcentage du pixel qui est imperméable à l'eau (*sealed*), et varie entre 0 et 1. Comme nous l'avons mentionné dans le Chapitre 1, les valeurs d'imperméabilité sont fonction de la classe d'occupation du sol (Homer *et al.*, 2007), avec des valeurs plus importantes pour les classes d'urbanisation, comme décrit par le Tableau 2.2.

Tableau 2.2. Classes d'occupation du sol dans la base de données NLCD, d'après Homer *et al.* (2007).

Classe d'occupation du sol	Signification	Valeur d'imperméabilité totale (-)
11	<i>Open Water</i> – Plan d'eau	0
12	<i>Perennial Ice/Snow</i> – Couvert neigeux	0
21	<i>Developed Open Space</i> – Urbanisé espaces ouverts	[0, 0.2]
22	<i>Developed Low</i> – Urbanisé faible densité	[0.2, 0.5]
23	<i>Developed Medium</i> – Urbanisé moyenne densité	[0.5, 0.8]
24	<i>Developed High</i> – Urbanisé forte densité	[0.8, 1]
31	<i>Barren Land</i> – Terre dénudée	0
32	<i>Unconsolidated Shore</i> – Sols avec très peu de végétation	0
41	<i>Deciduous Forest</i> – Forêt décidue	0
42	<i>Evergreen Forest</i> – Forêt sempervirente	0
43	<i>Mixed Forest</i> – Forêt mixte	0
51	<i>Dwarf Scrub</i> – Arbuste nain	0
52	<i>Shrub/Scrub</i> – Arbuste	0
71	<i>Grassland/Herbaceous</i> – Prairie/Herbacée	0
72	<i>Sedge/Herbaceous</i> – Cariçaie/Herbacée	0
73	<i>Lichens</i> – Lichens	0
74	<i>Moss</i> – Mousse	0
81	<i>Pasture/Hay</i> – Pâturage	0
82	<i>Cultivated Crops</i> – Cultures	0
90	<i>Woody Wetlands</i> – Zones humides boisées	0
95	<i>Emergent Herbaceous Wetlands</i> – Zones humides herbacées émergentes	0

En France, nous avons utilisé le produit CLC du service Copernicus pour la surveillance des sols (CLMS). Il est disponible pour les années 1990, 2000, 2006 et 2012, et caractérise l'occupation du sol à une résolution de 20 m en utilisant 44 sous-classes (Tableau 2.3⁷). L'information sur l'imperméabilité est fournie par les couches de densité d'imperméabilité (*Imperviousness Density Layers*), à une résolution de 20 m, disponible pour les années 2006, 2009, 2012 et 2015.

⁷Cette classification peut être retrouvée dans : <https://land.copernicus.eu/user-corner/technical-library/corine-land-cover-nomenclature-guidelines/html>.

Ces couches contiennent le taux d'imperméabilité sur chaque pixel, déterminé en se basant sur l'indice NDVI (Langanke *et al.*, 2016). Ainsi, les valeurs d'imperméabilité ne sont pas spécifiées pour chaque classe d'occupation du sol (comme c'est le cas pour NLCD).

Tableau 2.3. Classes et sous-classes d'occupation du sol pour le produit CORINE Land Cover. Les sous-classes sont détaillées uniquement pour les classes qui nous intéressent par la suite.

Type	Classes	Sous-classes
1 <i>Artificial surfaces</i> – Surfaces artificielles	1.1 <i>Urban fabric</i> – Tissu urbain	1.1.1 <i>Continuous urban fabric</i> – Tissu urbain continu 1.1.2 <i>Discontinuous urban fabric</i> – Tissu urbain discontinu
	1.2 <i>Industrial, commercial and transport units</i> – Unités industrielles, commerciales et de transport	1.2.1 <i>Industrial or commercial units</i> - Unités industrielles ou commerciales 1.2.2 <i>Road and rail networks and associated lands</i> - Routes, réseau ferroviaire et terres associées 1.2.3 <i>Port Areas</i> - Aires portuaires 1.2.4 <i>Airports</i> - Aéroports
	1.3 <i>Mine, dump and construction sites</i> – Mines, décharges et sites de construction	1.3.1, 1.3.2 et 1.3.3
	1.4 <i>Artificial, non-agricultural vegetated areas</i> – Aires artificielles, non agricoles et végétalisées	1.4.1 et 1.4.2
2 <i>Agricultural areas</i> – Surfaces agricoles	2.1 <i>Arable land</i> – Terres arables	2.1.1, 2.1.2 et 2.1.3
	2.2 <i>Permanent crops</i> – Cultures	2.2.1, 2.2.2 et 2.2.3
	2.3 <i>Pastures</i> – Pâturage	2.3.1
	2.4 <i>Heterogeneous agricultural areas</i> – Aires agricoles hétérogènes	2.4.1, 2.4.2, 2.4.3 et 2.4.4
3 <i>Forest and seminatural areas</i> – Forêt et surfaces semi-naturelles	3.1 <i>Forest</i> – Forêt	3.1.1 <i>Broad-leaved forest</i> – Forêt de feuillus 3.1.2 <i>Coniferous forest</i> – Forêt de conifères 3.1.3 <i>Mixed forest</i> – Forêt mixte
	3.2 <i>Shrub and/or herbaceous vegetation associations</i> – Arbuste et/ou herbacé	3.2.1, 3.2.2, 3.2.3 et 3.2.4

Suite à la page suivante...

Tableau 2.3 suite : Classes et sous-classes d'occupation du sol pour le produit CORINE Land Cover.

Type	Classes	Sous-classes
	3.3 <i>Open spaces with little or no vegetation</i> – Espaces ouverts avec ou sans végétation	3.3.1, 3.3.2, 3.3.3, 3.3.4 et 3.3.5
4 <i>Wetlands</i> – Zones humides	4.1 <i>Inland wetlands</i> – Zones humides intérieures	4.1.1 et 4.1.2
	4.2 <i>Coastal wetlands</i> – Zones humides côtières	4.2.1, 4.2.2 et 4.2.3
5 <i>Water bodies</i> – Plans d'eau	5.1 <i>Inland waters</i> – Eaux intérieures	5.1.1 <i>Water courses</i> – Cours d'eau 5.1.2 <i>Water bodies</i> – Plans d'eau
	5.2 <i>Marine waters</i> – Eaux marines	5.2.1, 5.2.2 et 5.2.3

Fin du Tableau 2.3.

2.4.1.2 Métriques utilisées

Dans ces travaux, nous avons fait le choix d'adopter une caractérisation dynamique de l'occupation du sol sur un pas de temps annuel. Ce choix permet de tenir compte des bassins versants qui ont pu connaître une évolution importante de leur occupation du sol au cours de la période de disponibilité des données hydroclimatiques. Nous avons focalisé sur trois aspects : le couvert urbain qui est la préoccupation centrale de cette thèse, le couvert forestier et les plans d'eau comme variables importantes de point de vue de l'évapotranspiration. Nous avons choisi 6 métriques principales : 4 pour caractériser le couvert urbain, 1 pour caractériser le couvert forestier, et 1 pour caractériser les plans d'eau.

Pour le couvert urbain, nous définissons :

1. La proportion du bassin versant occupée par les classes d'urbanisation, notée *CPD* (pour *Catchment Percent Developed*, héritée de la nomenclature de la base de données GAGES-II). Les classes choisies pour le cas des bassins nord-américains sont 21, 22, 23 et 24 (voir Tableau 2.2). Pour les bassins situés en France, le choix était moins facile car la grande famille des surfaces artificialisées comporte d'autres classes qui ne sont pas forcément liées à l'urbanisation (Tableau 2.3). Au final, nous avons choisi les classes 1.1 (*Urban fabric*) et 1.2 (*Industrial, commercial and transport units*). Nous avons décompté, pour chaque bassin, le nombre de pixels occupés par ces classes. Ce nombre a été divisé par le nombre total des pixels du bassin pour donner *CPD*. Cette métrique varie de 0

pour un bassin ne comptant aucune classe d'urbanisation, jusqu'à 1 pour un bassin complètement urbanisé.

2. Le taux moyen d'imperméabilité totale du bassin versant, noté TIA (pour *Total ImperVIOUS Area*). Ce taux a été estimé par une moyenne arithmétique des valeurs d'imperméabilité totale extraites à partir des pixels du bassin versant. Comme CPD , il varie entre 0 pour un bassin ne comptant aucune classe d'urbanisation, et 1 pour un bassin dont les pixels sont tous complètement imperméabilisés.
3. La fraction des pixels ayant une valeur d'imperméabilité totale qui dépasse 0.8, noté $IMP80$. Cette fraction donne une indication sur le taux de surfaces fortement imperméables du bassin versant. Elle varie entre 0 pour un bassin ne comptant aucun pixel avec une imperméabilité supérieure à 0.8, et 1 pour un bassin dont l'imperméabilité minimale de ses pixels est supérieure ou égale à 0.8.
4. L'irrégularité de l'imperméabilité, notée $IrIMP$. Elle est calculée à partir de l'équation suivante :

$$IrIMP = \frac{\sqrt{\sum_i IMP_i^2}}{\sum_i IMP_i} = \frac{\sqrt{\sum_i IMP_i^2}}{N_{BV} \cdot TIA}$$

où IMP_i est le taux d'imperméabilité totale du $i^{\text{ème}}$ pixel appartenant au bassin versant, et N_{BV} est le nombre total des pixels du bassin versant. Cette mesure permet de rendre compte de la variabilité spatiale de l'imperméabilité; on pourrait avoir deux bassins avec le même taux moyen d'imperméabilité totale TIA , mais avec deux distributions différentes, ce qui engendrera deux valeurs de $IrIMP$ différentes. Pour un même TIA , plus $IrIMP$ est importante, plus la distribution est dispersée, donc plus le couvert imperméable est hétérogène spatialement.

En plus des mesures du couvert urbain, nous avons déterminé deux autres mesures d'occupation du sol :

1. La fraction du bassin versant occupée par la forêt, notée $fFOR$. Elle est calculée en prenant le nombre des pixels du bassin versant occupés par les classes de forêts. Les classes choisies pour les bassins nord-américains sont 41, 42 et 43 (voir Tableau 2.2). Pour les bassins français, nous avons choisi les pixels occupés par une des sous-classes de la classe 3.1 (3.1.1, 3.1.2 et 3.1.3, voir Tableau 2.3). Cette fraction varie entre 0 pour un bassin ne comptant aucune classe de forêt, et 1 pour un bassin complètement forestier.
2. La fraction du bassin versant occupée par un plan d'eau, notée fW . Comme pour $fFOR$, fW est déterminée en divisant le nombre des pixels occupés par les classes concernées sur le nombre total des pixels du bassin versant. Pour les bassins versants

nord-américains, nous avons choisi les deux classes 11 (*Open water*) et 12 (*Perennial ice/snow*). Pour les bassins situés en France, nous avons choisi la classe 5.1 avec ses sous-classes (5.1.1 et 5.1.2). Noter que même si les pixels appartenant à ces classes sont hydrologiquement imperméables, ils sont attribués d'une imperméabilité nulle du fait qu'ils ne sont pas concernés par l'urbanisation.

Ces 6 métriques sont calculées sur chaque année de disponibilité des bases de données spatiales et sont considérées variables au pas de temps annuel. Or, plusieurs années de données hydroclimatiques se trouvent sans caractérisation d'occupation du sol. Nous détaillons dans la prochaine partie la méthode d'estimation de chacune des 6 mesures pour les années manquantes.

2.4.1.3 Caractérisation de l'occupation du sol pour les années lacunaires

Pour les bassins versants situés aux États-Unis, le comblement des années manquantes a été entrepris de manière empirique et simple. Rappelons que les données hydroclimatiques s'étendent de 2002 à 2017. L'occupation du sol ainsi que l'imperméabilité sont disponibles pour 2001, 2006, 2011 et 2016. Ainsi, de 2002 à 2016, une interpolation linéaire entre les valeurs disponibles a permis de combler les lacunes. Pour l'année 2017, nous avons effectué une extrapolation linéaire en prenant comme gradient d'évolution la médiane des trois gradients calculés sur 2001-2006, 2006-2011 et 2011-2016. Ce gradient médian a permis d'estimer les métriques en 2017 à partir de la situation en 2016.

En revanche, l'exercice a été moins facile pour les bassins situés en France. Les données hydroclimatiques couvrent la période 1997-2012. Pour les métriques caractérisant l'occupation du sol (*CPD*, *fFOR* et *fW*), les données sont disponibles en 1990, 2000, 2006 et 2012, ce qui a facilité le comblement de lacunes moyennant une interpolation linéaire. Par contre, l'information sur l'imperméabilité n'est disponible qu'à partir de 2006. Si l'interpolation a servi pour combler les lacunes en *TIA*, *IMP80* et *IrIMP* entre 2006 et 2012, le problème se pose quant à l'estimation de l'imperméabilité sur la période 1997-2005. Pour cela, nous avons exploité l'information sur *CPD*, disponible par interpolation de 1990 à 2005, pour déterminer les valeurs manquantes en *TIA*, *IMP80* et *IrIMP*.

Pour ce faire, des formules de passage de *CPD* aux métriques d'imperméabilité devaient être

estimées. Nous avons choisi *a priori* les relations empiriques suivantes :

$$TIA = a \cdot \left(1 - \sqrt{1 - CPD}\right) \quad (2.4.1)$$

$$IMP80 = b \cdot CPD \quad (2.4.2)$$

$$IrIMP = \frac{c}{CPD^d} \quad (2.4.3)$$

où a , b , c et d sont les paramètres à caler par régression linéaire. Le choix de ces relations a été orienté par une simple visualisation de la variation des métriques d'imperméabilité en fonction de CPD (Figure 2.9). Pour avoir une estimation robuste des paramètres, nous avons inclus également les bassins versants nord-américains. Le nombre de points { CPD , métrique d'imperméabilité} disponibles sur une année est de 4 pour un bassin situé aux États-Unis (correspondant aux années 2001, 2006, 2011 et 2016). Nous n'avons calculé ces métriques que pour 1261 bassins nord-américains (incluant les 870 bassins versants sélectionnés). Par ailleurs, le nombre de points { CPD , métrique d'imperméabilité} disponibles sur une année n'est que de 2 pour les bassins français (correspondant aux années 2006 et 2012). Ces deux métriques ont été calculées pour les 3701 bassins français. Au total, cela a fourni un nombre de 12446 couples { CPD , métrique d'imperméabilité} pour pouvoir estimer les relations empiriques. Les relations obtenues pour chaque pays et pour chaque métrique sont montrées sur la Figure 2.9.

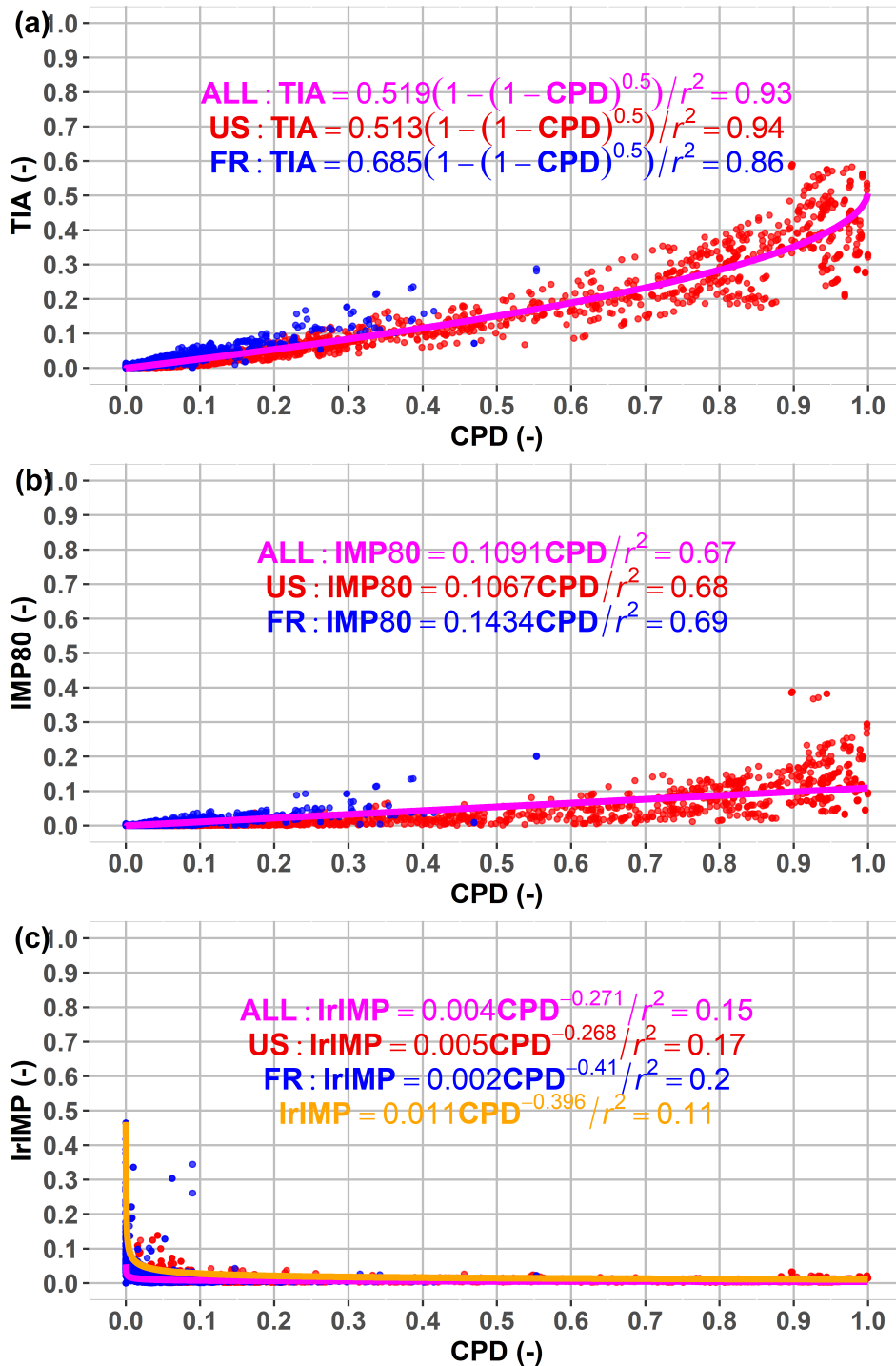


Figure 2.9. Relations empiriques entre CPD et (a) TIA , (b) $IMP80$ et (c) $IrIMP$. Les points rouges indiquent ceux qui proviennent de l'échantillon nord-américain (5044 points). Les points en bleu indiquent ceux qui proviennent de l'échantillon français (7402 points). Les relations sont indiquées pour l'échantillon total (ALL), nord-américain (US) et français (FR). r^2 est le coefficient de détermination. La courbe en magenta représente la relation estimée à partir de l'échantillon total. Pour $IrIMP$, la courbe en orange correspond à la relation adoptée pour cette métrique, son r^2 étant pris comme le carré de la corrélation entre les valeurs observées et estimées par la relation.

Nous remarquons une forte corrélation entre TIA et CPD , une corrélation importante entre $IMP80$ et CPD mais une faible corrélation entre $IrIMP$ et CPD . La forte corrélation entre TIA et CPD est confirmée par d'autres études (Bhaskar *et al.*, 2020; Oudin *et al.*, 2018). L'inclusion de l'échantillon nord-américain a permis de couvrir toute la plage des situations d'urbanisation, en rajoutant des cas avec des valeurs de CPD très fortes. Ceci s'est traduit par de faibles différences entre les paramètres estimés pour l'échantillon total et le sous-échantillon nord-américain, suggérant un poids important des points provenant des bassins versants nord-américains. Si on reproduit relativement bien le comportement au voisinage de CPD proche de 1 pour la relation $TIA = f(CPD)$, ceci est moins le cas pour $IMP80$ et l'est beaucoup moins pour $IrIMP$, pour laquelle aucune des relations obtenues ne permet de bien suivre la variabilité du nuage des points (Figures 2.9b et 2.9c, courbes en magenta). Plus tard, nous allons voir que cela n'est pas si préjudiciable pour nos travaux, pour lesquels TIA se révèle d'une plus grande importance par rapport à $IrIMP$ et $IMP80$. Pour $IrIMP$, nous avons remplacé la courbe issue de la régression par une courbe alternative (Figure 2.9c, courbe en orange), dont l'équation obtenue par un ajustement manuel des paramètres n'améliore pas pour autant la corrélation mais suit visuellement mieux la variabilité des points, surtout au voisinage de CPD proche de 0. Pour TIA et $IMP80$, nous avons gardé les relations estimées à partir de l'échantillon total (Figures 2.9a et 2.9b, courbes en magenta).

Cet exercice a montré que les relations peuvent être considérées homogènes entre les deux pays, malgré les différences méthodologiques d'élaboration de ces bases de données et malgré les différences en termes de niveau d'urbanisation entre les deux rives de l'Atlantique Nord. Nous allons donc regarder ces différences plus en détail pour l'ensemble des métriques d'occupation du sol.

2.4.2 Variabilité spatiale des cas d'urbanisation et différences d'occupation du sol entre la France et les États-Unis

Après avoir comblé les lacunes, les moyennes de chacune des métriques d'occupation du sol au long de la période de disponibilité de données hydroclimatiques ont été estimées pour chaque bassin versant. Les distributions de chacune des 6 métriques sont montrées sur la Figure 2.10.

L'échantillon des bassins versants préparé couvre toute la plage des situations d'urbanisation, avec des valeurs de CPD qui vont de 0 jusqu'à 1. En revanche, la distribution est très asymétrique; la plupart des cas sont ruraux ou faiblement urbanisés, ce qui est reflété par une médiane de CPD à 0.038. La distribution de TIA suit celle de CPD avec une médiane de 0.009 (moins de 1%); en revanche, les moyennes de TIA à l'échelle de chaque bassin versant ne dépassent pas 0.6. Le nombre de bassins versants ayant une valeur de TIA supérieure à 5%,

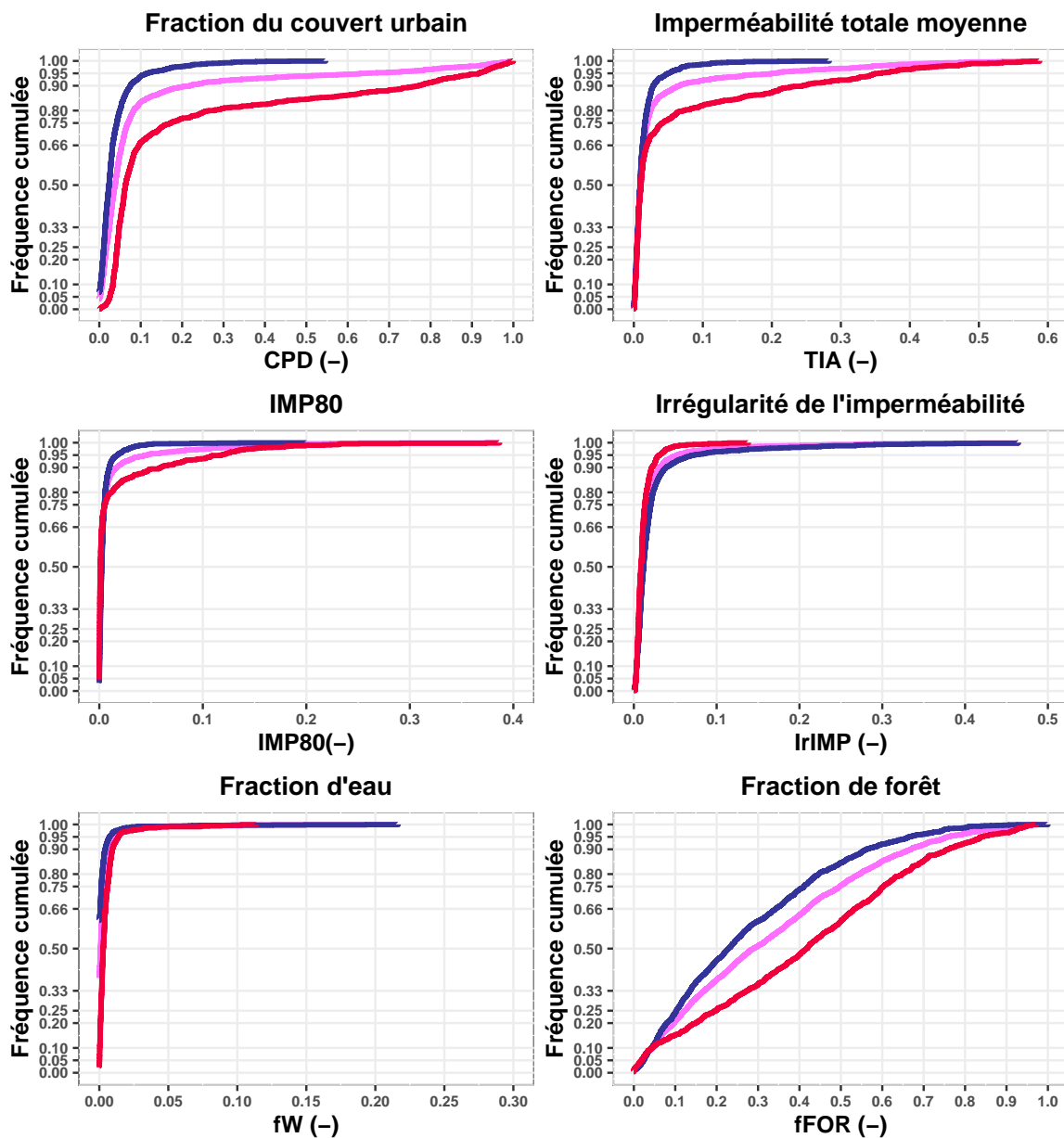


Figure 2.10. Courbes cumulées des 6 métriques utilisées pour la description de l'occupation du sol. Pour rappel, *IMP80* désigne la fraction de l'aire du bassin versant dont l'imperméabilité dépasse 80%. Les courbes en bleu, en magenta et en rouge représentent respectivement les courbes de l'échantillon français, total et nord-américain.

10% et 20% est respectivement 273, 175 et 114, ce que l'on peut considérer très important par rapport aux tailles des échantillons des bassins versants fortement urbanisés reportés dans la littérature. *IMP80* et *IrIMP* indiquent une plus faible variabilité, avec des distributions plus asymétriques et des valeurs dans la plupart des cas proches de zéro. La fraction d'eau *fW* couvre également des valeurs majoritairement nulles. Par ailleurs, nous notons une présence importante des bassins versants forestiers, manifestée par *fFOR* qui s'étale de 0 jusqu'à 1 avec une médiane de 0.29.

En regardant les différences France/États-Unis, nous notons que la grande majorité des bassins versants fortement urbanisés se situent aux États-Unis, comme indiqué par une plage de *CPD* plus large (entre 0 et 1 pour les bassins nord-américains contre 0 et 0.55 en France), d'un écart-type plus important (0.28 aux États-Unis contre 0.05 en France). Les médianes respectives sont 0.063 pour les bassins nord-américains contre 0.023 pour les bassins français. Ceci est confirmé par la distribution de *TIA*, qui ne dépasse pas 0.28 pour les bassins situés en France alors qu'elle atteint 0.59 aux États-Unis. Le nombre de bassins français comparé au nombre de bassins nord-américains est 3 contre 111 lorsque *TIA* dépasse 20%, 19 contre 156 lorsque *TIA* dépasse 10%, et 68 contre 205 lorsque *TIA* dépasse 5%. Si la tendance de *IMP80* suit celle de *TIA* et *CPD*, *IrIMP* montre une plus grande diversité sur les bassins versants français par rapport aux bassins versants nord-américains, ce qui peut être lié aux différences de résolution spatiale entre les bases de données d'occupation du sol. La fraction d'eau montre qu'elle est légèrement plus importante dans l'échantillon nord-américain que dans l'échantillon français. Même s'ils affichent les plus gros taux d'urbanisation, les bassins nord-américains ont des fractions de couvert forestier plus importantes que celles rencontrées dans les bassins français, avec une médiane de 0.42 contre 0.23 en France.

Ces différences entre les bassins nord-américains et français peuvent être attribuées aux différences de taille des deux pays, ce qui induit statistiquement une plus grande diversité des cas illustrés par les bassins nord-américains. Elles reflètent aussi des différences d'aménagement qui font que le couvert urbain est plus étalé aux États-Unis qu'en France⁸. Cette comparaison montre en particulier la richesse en termes de situations d'urbanisation (diversité spatiale) de l'échantillon préparé, ce qui permettra de regarder la question de l'impact de l'urbanisation sur le comportement hydrologique pour un gradient d'imperméabilité riche.

On se demande donc si ce gradient spatial (c'est-à-dire d'un bassin versant à l'autre) s'accompagne d'un gradient temporel de l'urbanisation (c'est-à-dire une évolution de l'urbanisation pour un seul bassin versant au cours du temps).

⁸Ces différences d'urbanisation entre la France et les États-Unis ne seront pas abordées en détail dans cette thèse.

2.4.3 Évolution temporelle de l'urbanisation de chacun des bassins versants sélectionnés

Dans cette partie, nous nous intéressons uniquement à l'évolution de la fraction du bassin versant occupée par les classes d'urbanisation, mesurée par *CPD*. Pour rappel, celle-ci a été déterminée sur les bassins français pour les années 1990, 2000, 2006 et 2012. Sur les bassins situés aux États-Unis, elle a été déterminée pour les années 2001, 2006, 2011 et 2016. Sur la Figure 2.11, nous montrons l'évolution de la distribution de *CPD* pour l'échantillon français et nord-américain pour chacune des 4 années. De plus, nous montrons les distributions des gradients d'urbanisation déterminés pour chaque bassin versant.

Nous remarquons que les courbes cumulées de *CPD* sont très serrées (Figure 2.11a et 2.11c), ce qui indique une faible évolution temporelle de l'urbanisation à l'échelle de tout l'échantillon. Ceci est en partie confirmé par les distributions de l'évolution du couvert urbain (Figures 2.11b et 2.11d), où l'on constate que la conversion des surfaces non urbaines en surfaces urbaines a concerné quelques cas uniquement. Les gradients d'urbanisation sont plus faibles en France qu'aux États-Unis : une conversion maximale de +20% de la taille du bassin en surface urbanisée entre 1990 et 2012 (22 ans) contre une conversion maximale d'environ +35% sur la période 2001-2016 (15 ans) pour les bassins nord-américains. En France, uniquement 3 bassins versants ont eu une différence absolue de *CPD* supérieure à +10% entre 1990 et 2012, alors que 45 bassins versants nord-américains ont eu une conversion de plus que +10% de leur surface en couvert urbain de 2001 à 2016.

On peut donc conclure que l'étendue de la période étudiée n'a permis d'avoir un gradient d'urbanisation suffisamment important que sur quelques bassins versants. Néanmoins, le nombre de ces bassins (48 pour un accroissement du taux des surfaces urbanisées ΔCPD supérieur à +10% sur 15-22 ans) reste important par rapport aux tailles des échantillons de bassins versants urbanisés généralement étudiés. De toute manière, cela n'exclue pas la richesse de l'échantillon d'un point de vue de l'occupation du sol de par son offre diversifiée de situations d'urbanisation dans l'espace. La richesse de cet échantillon en conditions hydroclimatiques est explorée dans la prochaine section.

2.5 Caractéristiques hydroclimatiques de l'échantillon

Nous avons collecté un échantillon de bassins versants qui offre une large panoplie de situations d'urbanisation. Un des points forts de l'utilisation d'un large échantillon est la diversité de régimes climatiques et hydrologiques, ce qui est bénéfique pour le développement et le test de modèles pluie-débit. La diversité géographique de notre échantillon suggère une diversité

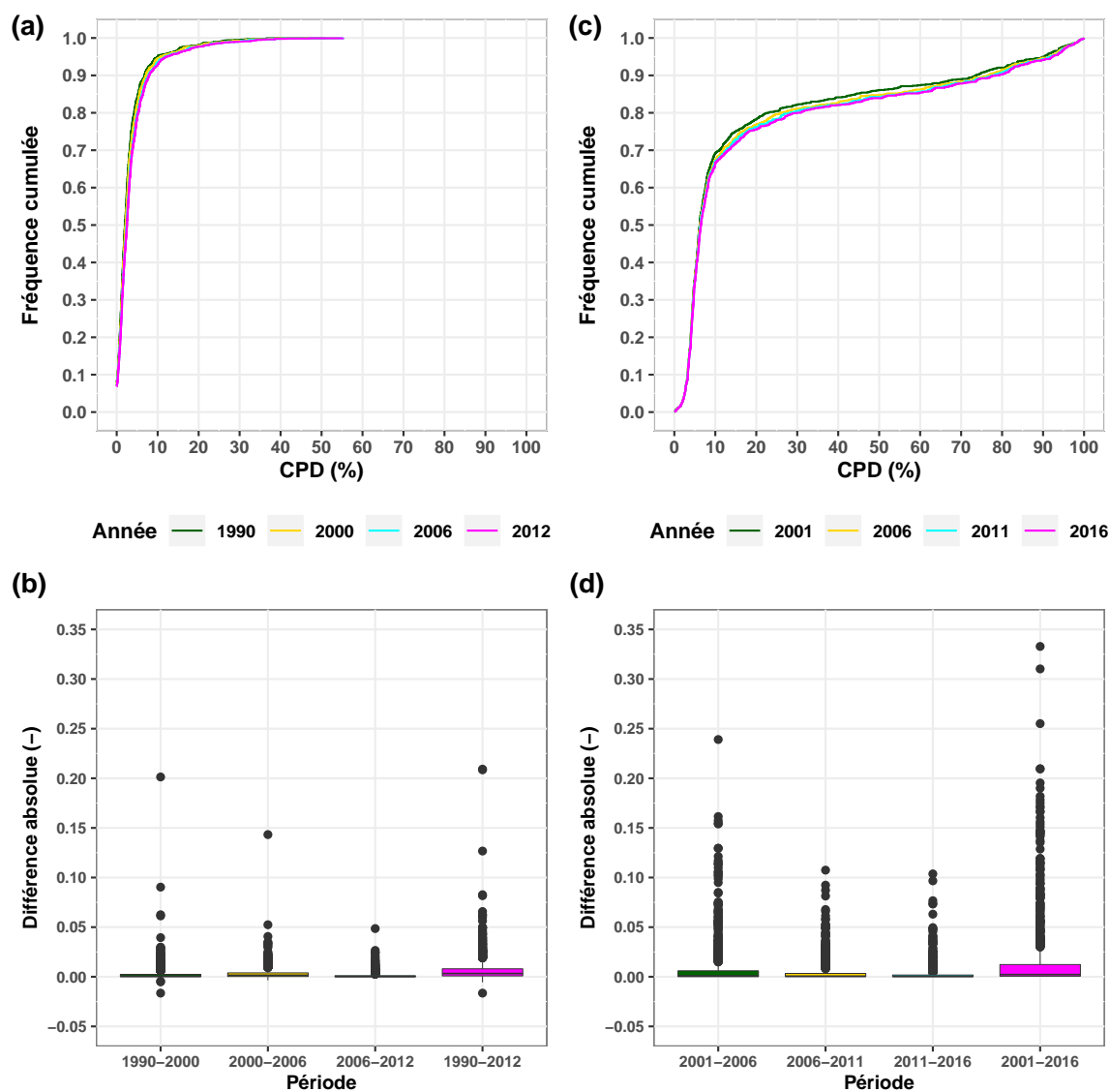


Figure 2.11. En haut, évolution des courbes cumulées de la fraction du couvert urbain *CPD* pour (a) les bassins versants situés en France, à partir de CORINE Land Cover 1990, 2000, 2006 et 2012, et (c) les bassins versants situés aux États-Unis, déterminés à partir de NLCD 2001, 2006, 2011 et 2016. En bas, distributions de l'évolution temporelle du couvert urbain des bassins versants situés (b) en France et (d) aux États-Unis.

climatique (à partir des Figures 2.5 et 2.6). Nous allons néanmoins confirmer ce résultat à travers des métriques déterminées pour chaque bassin versant. Une fois les métriques présentées, nous montrons leur variabilité sur le territoire métropolitain, puis sur le sud-est des États-Unis. Ensuite, grâce à des courbes cumulées, nous comparons l'échantillon français à l'échantillon nord-américain en termes hydroclimatiques. Enfin, nous regardons dans quelle mesure les bassins les plus fortement urbanisés dans l'échantillon conservent la même diversité climatique que l'échantillon total.

2.5.1 Caractéristiques hydroclimatiques choisies

Pour caractériser le climat de chaque bassin versant, nous avons adopté 5 mesures simples se basant sur la précipitation et l'ETP, préparées comme décrit en Section 2.2. Ces mesures sont :

1. La précipitation moyenne interannuelle P_m (mm/an), prise comme la moyenne des précipitations enregistrées au pas de temps horaire, que l'on a multipliée par 24 pour avoir une moyenne journalière, puis par 365.25 pour avoir une moyenne interannuelle en mm/an.
2. L'ETP moyenne interannuelle ETP_m (mm/an), prise comme la moyenne des évapotranspirations potentielles estimées au pas de temps horaire, multipliée par 24 pour avoir une moyenne journalière, puis par 365.25 pour avoir une moyenne interannuelle en mm/an.
3. L'indice d'aridité, pris comme le rapport ETP_m/P_m (-) de l'ETP moyenne sur la précipitation moyenne.
4. La flashiness⁹ de précipitation FP (-), estimée par la relation (Baker *et al.*, 2004) :

$$FP = \frac{\sum_h |P_h - P_{h-1}|}{\sum_h P_h}$$

où P_h est la lame précipitée à l'heure h . FP est d'autant plus importante que le signal de précipitation contient des variations importantes (c'est-à-dire un signal rugueux), alors qu'elle est faible pour un signal plutôt lisse.

5. Le coefficient d'intensité de précipitation P_{99}/P_m (-), utilisé par Ficchi (2017, p. 79) au pas de temps journalier, où P_{99} (mm/h) est le quantile 99% des valeurs de précipitations horaires pour un bassin donné, et que l'on peut interpréter comme la lame d'eau de précipitation dépassée en moyenne une fois sur 100. P_m est la précipitation moyenne interannuelle (mm/h). Ce coefficient donne une indication sur les événements pluviométriques extrêmes rencontrés sur le bassin versant.

⁹Anglicisme conservé car très répandu dans la littérature en anglais. Il peut être remplacé par « soudaineté ». Ce terme a été hérité des travaux de thèse de Poncelet (2016).

En plus des caractéristiques climatiques, nous avons adopté 4 autres caractéristiques de la réponse hydrologique :

1. L'écoulement moyen interannuel ou module Q_m (mm/an), pris comme la moyenne des débits estimés au pas de temps horaire, que l'on a multipliée par 24 pour avoir une moyenne journalière, puis par 365.25 pour avoir une moyenne interannuelle, ramenée à la surface topographique du bassin versant.
2. L'indice d'écoulement de base IEB (-), calculé comme le volume de l'écoulement de base divisé par le volume total écoulé. Pour obtenir le débit de base, nous avons utilisé un filtre numérique récursif (Collischonn et Fan, 2013; Eckhardt, 2005), comprenant deux paramètres. Le premier est fonction de la constante de récession désignée communément par k , figurant dans l'équation linéaire $\frac{dQ}{dt} = Qe^{-kt}$ (Blume, Zehe et Bronstert, 2007). Son estimation a été effectuée à l'aide d'une régression. Le deuxième paramètre noté BFI_{max} est estimé itérativement à l'aide d'une enveloppe de débit de base, délimitée par l'utilisation de filtres numériques, qui permet notamment de déterminer les points de début et de fin des crues, et par conséquent les points de début et de fin de l'écoulement provenant de la surface (Mei et Anagnostou, 2015).
3. L'indice de débit maximal horaire Q_{99}/Q_m (-), où Q_{99} (mm/h) est le quantile 99% de la distribution des débits moyens horaires observés à l'exutoire du bassin versant, et Q_m (mm/h) l'écoulement moyen interannuel. Cet indice donne une information sur les valeurs extrêmes que peut atteindre la réponse du bassin versant et sa violence par rapport aux conditions habituelles localement.
4. Le temps caractéristique de la réponse hydrologique, noté DT (en heures). Il est considéré comme le décalage à appliquer sur les séries de débit horaire pour maximiser la corrélation avec les séries de précipitation horaire. Ce temps donne une idée sur la rapidité de la réponse du bassin versant.

La variabilité spatiale de chacune de ces caractéristiques hydroclimatiques (sauf DT) est montrée pour les bassins situés en France, puis pour les bassins nord-américains, dans la prochaine partie.

2.5.2 Variabilité spatiale des caractéristiques hydroclimatiques en France et aux États-Unis

Les caractéristiques hydroclimatiques sont montrées sur la Figure 2.12 pour les bassins situés en France. À l'exception des très hautes altitudes et de la Corse, ces bassins couvrent la majeure partie du territoire métropolitain. Les précipitations moyennes varient de 510 à

1790 mm/an, avec des moyennes plus importantes sur les reliefs (le massif armoricain, les côtes normandes, le massif central, les Vosges et le Jura à l'est, le pied des Pyrénées). L'ETP varie de 590 à 910 mm/an, avec des valeurs plus importantes dans le centre, le bassin aquitain et le sud-est. L'indice d'aridité combine les deux forçages : il est généralement inférieur à 1 pour les bassins français caractérisés par des climats tempérés et océaniques (type Cfb dans la classification Köppen-Geiger, voir Beck *et al.*, 2018). En revanche, l'indice est relativement plus important pour les bassins en région méditerranéenne, en Occitanie et dans le centre. La variabilité spatiale de l'écoulement moyen interannuel suit assez fidèlement la variabilité spatiale des précipitations, avec 95% des valeurs entre 80 et 1050 mm/an. La variabilité spatiale de l'indice d'écoulement de base est moins facile à discerner, vu sa dépendance présumée aux caractéristiques lithologiques. Il est relativement important pour la majorité des bassins français, atteignant en médiane 0.77, avec des valeurs proches de 1 pour des bassins localisés en majeure partie dans le bassin de la Somme, au nord et dans la région des Grands Causses, où les formations calcaires sont prédominantes. Des valeurs très faibles de cet indice sont enregistrées en Vendée, dans le Jura et au sud. La flashiness de précipitation indique plus de variabilité dans les séries de précipitation observées dans la moitié nord du territoire et dans le nord-est du massif central, opposée par de faibles valeurs au sud du massif central, dans le massif armoricain et à proximité des Alpes. Les extrêmes de précipitation, traduits par P_{99}/P_m , varient de 10 jusqu'à 30 fois la précipitation moyenne interannuelle. Les plus violents épisodes sont rencontrés en Vendée, au centre de la France où le climat est continental, et plus vigoureusement dans le sud-est et la région méditerranéenne, connue pour ses épisodes de crues éclair (Gaume *et al.*, 2009). Les extrêmes hydrologiques retraduisent ces extrêmes climatiques sur la région méditerranéenne et en Vendée, mais sont atténués par rapport à l'écoulement moyen dans le nord et en particulier dans le bassin de la Somme.

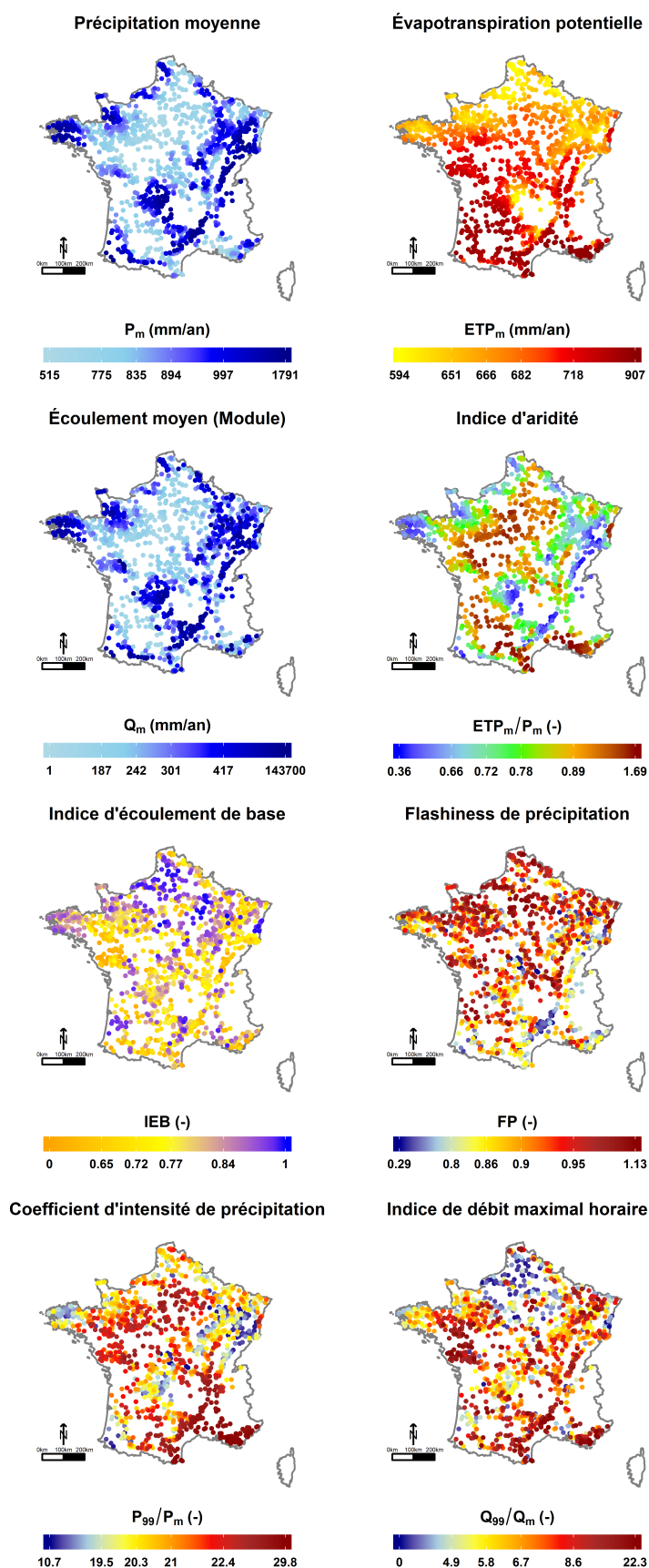


Figure 2.12. Caractéristiques hydroclimatiques des 1355 bassins versants français. Les points représentent les localisations des centroïdes des bassins versants. Pour chaque variable, on classe d'abord les bassins versants. Ensuite, la couleur est attribuée en fonction du rang du bassin versant (de 1 jusqu'à 1355). Cela permet de mieux dégager les contrastes régionaux pour chacune des variables.

Les 870 bassins nord-américains ne couvrent que les États-Unis du sud-est, pour lesquels les caractéristiques climatiques sont montrées sur la Figure 2.13. Les précipitations ont une plage de variation de 190 à 1790 mm/an, avec des précipitations relativement faibles sur les bassins de la plaine côtière atlantique et du centre (Texas, Oklahoma, Kansas). Les précipitations les plus importantes sont enregistrées à l'ouest du massif appalachien, dans les plaines centrales et sur les états côtiers du golfe du Mexique. L'ETP est caractérisée par un gradient nord-sud avec des valeurs de plus en plus importantes en allant vers le sud, surtout en Floride et au Texas. Les lames moyennes d'ETP vont de 750 à 1400 mm/an. L'indice d'aridité suit la variabilité des précipitations et de l'ETP, avec des valeurs plus faibles dans le nord-ouest du massif appalachien et plus fortes dans la plaine atlantique et le centre, où l'on atteint des indices supérieurs à 5, ce qui indique la présence de conditions climatiques très arides. Comme pour les bassins français, l'écoulement moyen suit bien la variabilité des précipitations, sauf une forte atténuation sur les bassins du Texas à cause d'indices d'aridité très forts. Sur ces bassins, on atteint des niveaux d'écoulement moyen quasiment nuls, ce qui est une indication de la présence de cours d'eau éphémères. L'indice d'écoulement de base couvre toute la plage possible de variation, avec une médiane de 0.56. Il indique que les apports en base sont importants autour du massif appalachien, en Floride, et au sud des états de la Géorgie et de l'Alabama, contrairement aux zones très arides, comme le sud du Texas. La flashiness de précipitation indique une répartition géographique uniforme, avec des valeurs très importantes sur le massif appalachien, en Floride, au Centre, au Texas et sur la confluence des quatre états d'Arkansas, Missouri, Kansas et Oklahoma (monts Ozarks). On remarque que c'est sensiblement les mêmes points où l'on a les plus forts extrêmes de précipitations (si l'on enlève le Texas), avec des coefficients P_{99}/P_m allant jusqu'à 30. En revanche, les extrêmes hydrologiques par rapport à l'écoulement moyen de chaque bassin sont plutôt rencontrés sur les bassins arides du centre, les bassins de la Floride et les bassins à l'est du massif appalachien.

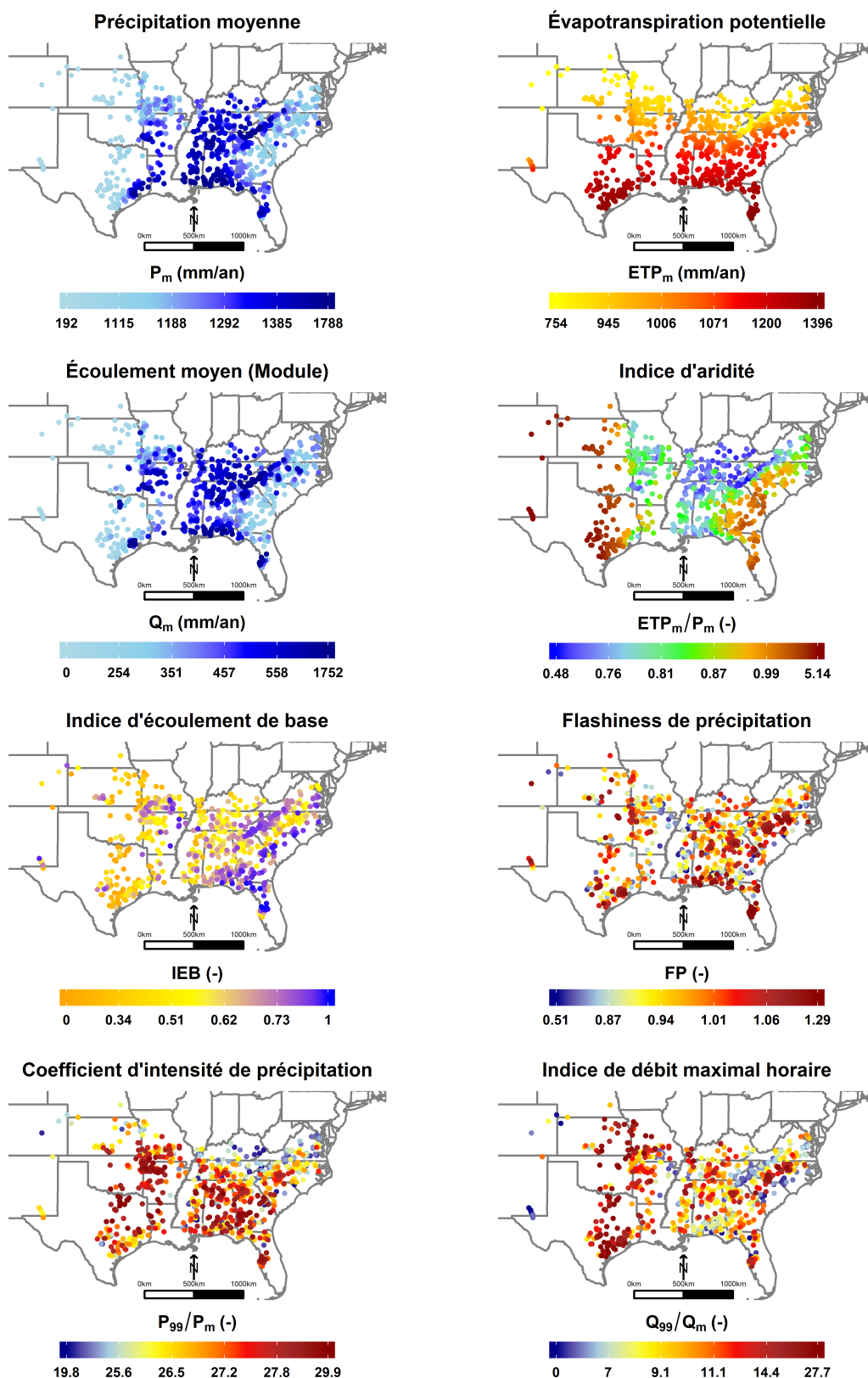


Figure 2.13. Caractéristiques hydroclimatiques des 870 bassins versants nord-américains. Les points représentent les localisations des centroïdes des bassins versants. Pour chaque variable, on classe d'abord les bassins versants. Ensuite, la couleur est attribuée en fonction du rang du bassin versant (de 1 jusqu'à 870). Cela permet de mieux dégager les contrastes régionaux pour chacune des variables.

À partir de cette description géographique des caractéristiques climatiques et hydrologiques de l'échantillon sélectionné, nous remarquons une diversité hydroclimatique importante des bassins versants choisis. On couvre une plage de régime climatique allant de l'aride jusqu'au tempéré, en passant par le méditerranéen. Cette diversité climatique est accompagnée d'une diversité hydrologique, notamment en termes d'écoulement moyen qui suit sensiblement la variabilité des précipitations et de l'ETP, également en termes de soutien par la nappe à travers l'écoulement de base. Les extrêmes hydrologiques sont en général associés aux régions arides et méditerranéennes.

Maintenant que la variabilité spatiale des conditions hydroclimatiques est décrite pour chaque pays, il est intéressant de voir, comme pour l'occupation du sol, les différences hydroclimatiques entre l'échantillon nord-américain et français.

2.5.3 Contrastes hydroclimatiques entre les bassins situés en France et les bassins situés aux États-Unis

À partir de la Figure 2.14, on remarque l'existence d'un contraste climatique important entre les bassins situés en France et aux États-Unis. Les précipitations sont plus fortes sur les bassins nord-américains que sur les bassins français (1240 mm/an contre 890 mm/an, valeurs médianes), ce qui est le cas aussi pour l'ETP, pour laquelle 90% des bassins situés aux États-Unis dépassent le maximum d'ETP enregistré par les bassins situés en France. Ceci est confirmé par les différences en médiane : 1040 mm/an sur les bassins nord-américains contre 680 mm/an sur les bassins français. En revanche, les indices d'aridité sont plutôt similaires, excepté quelques bassins nord-américains qui représentent des valeurs extrêmes dépassant 2. Cette similarité inclut également les écoulements moyens, avec une médiane d'environ 400 mm/an aux États-Unis contre 300 mm/an en France, et les temps caractéristiques de réponse (13 heures en médiane sur les bassins nord-américains contre 15 heures sur les bassins français). Notons que la majeure partie des bassins versants ont un temps de réponse infra-journalier (1517/2225 bassins versants), avec uniquement 4 bassins versants qui ont un temps de réponse caractéristique infra-horaire. Les apports par écoulement de base sont relativement plus importants pour l'échantillon des bassins français par rapport à l'échantillon nord-américain. Les distributions de la flashiness de précipitation sont assez proches entre les deux pays, avec des valeurs légèrement plus importantes aux États-Unis. En revanche, le coefficient d'intensité de précipitation P_{99}/P_m est beaucoup plus important sur les bassins nord-américains, mais la majorité des valeurs de P_{99}/P_m rencontrées sur ces bassins sont également atteintes sur les bassins français. Ces différences sont atténuées mais persistantes pour l'indice du débit maximal horaire, avec des valeurs relativement plus fortes aux États-Unis.

De cette comparaison, on retient deux constats importants : (1) des contrastes climatiques

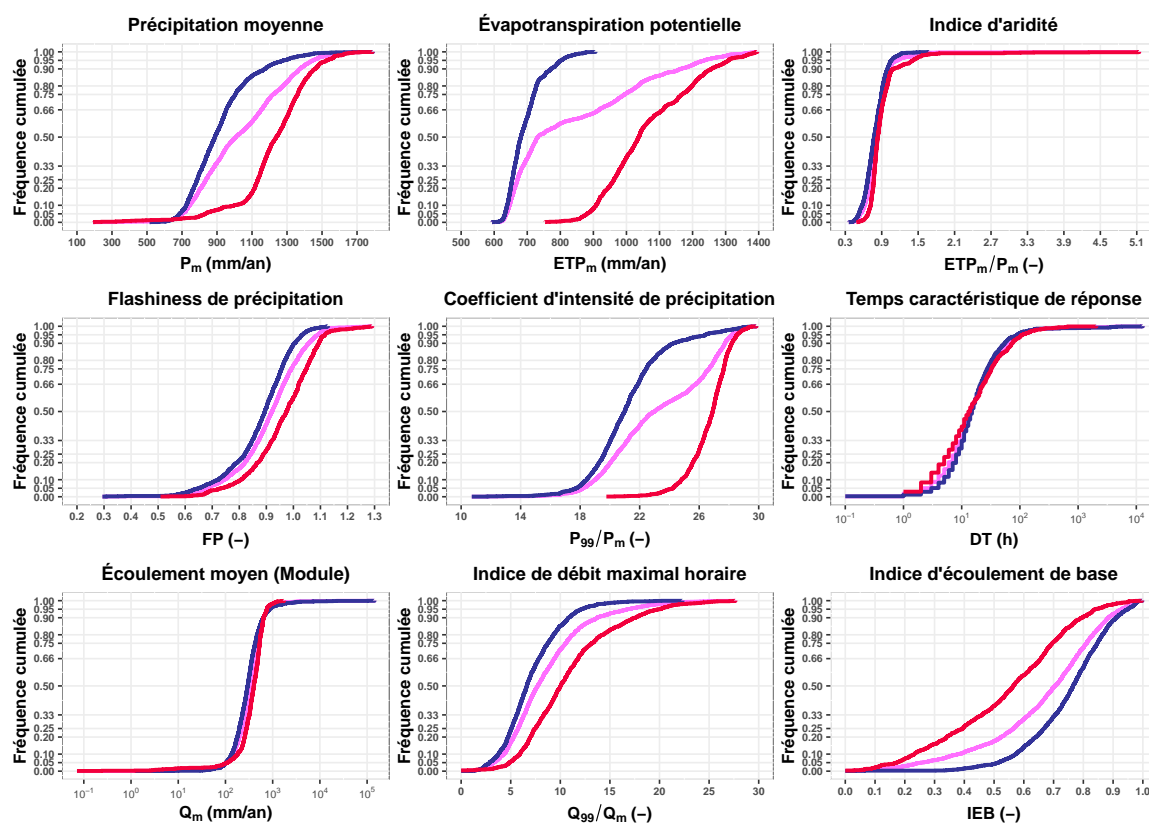


Figure 2.14. Courbes cumulées des caractéristiques hydroclimatiques pour l'échantillon total (en magenta), les bassins situés en France (en bleu) et les bassins situés aux États-Unis (en rouge).

existent entre les bassins nord-américains et français, ce qui contribue à la richesse de l'échantillon total. L'exemple de l'ETP (ou de l'indice d'aridité) montre que l'on ne peut pas enlever l'un ou l'autre des sous-échantillons sans appauvrir considérablement l'échantillon total, par élimination de cas très arides par exemple. Ceci est bénéfique au développement du modèle, lequel doit reproduire l'écoulement observé de manière satisfaisante tout au long de ce gradient climatique. **(2)** Ces contrastes doivent être gardés en mémoire lorsqu'on veut analyser l'impact des différences de paysage urbain, d'un pays à l'autre, sur l'hydrologie. Sinon, on pourrait attribuer de manière erronée une différence de comportement hydrologique (si différence il y a) à la structuration du paysage urbain alors qu'elle découle d'un contraste climatique.

2.5.4 Caractéristiques hydroclimatiques des bassins fortement urbanisés

Pour sélectionner les bassins versants fortement urbanisés, nous nous sommes basés sur deux seuils de taux moyen d'imperméabilité totale TIA : 5% et 10%. Ces deux seuils sont fréquemment reportés dans la littérature comme valeurs au-delà desquelles le changement dans le comportement hydrologique à cause de l'urbanisation devient significatif (Booth et Jackson, 1997; Brabec, Schulte et Richards, 2002; Salvadore, Bronders et Batelaan, 2015).

Les différences des caractéristiques hydroclimatiques des deux échantillons par rapport à l'échantillon total sont montrées sur la Figure 2.15.

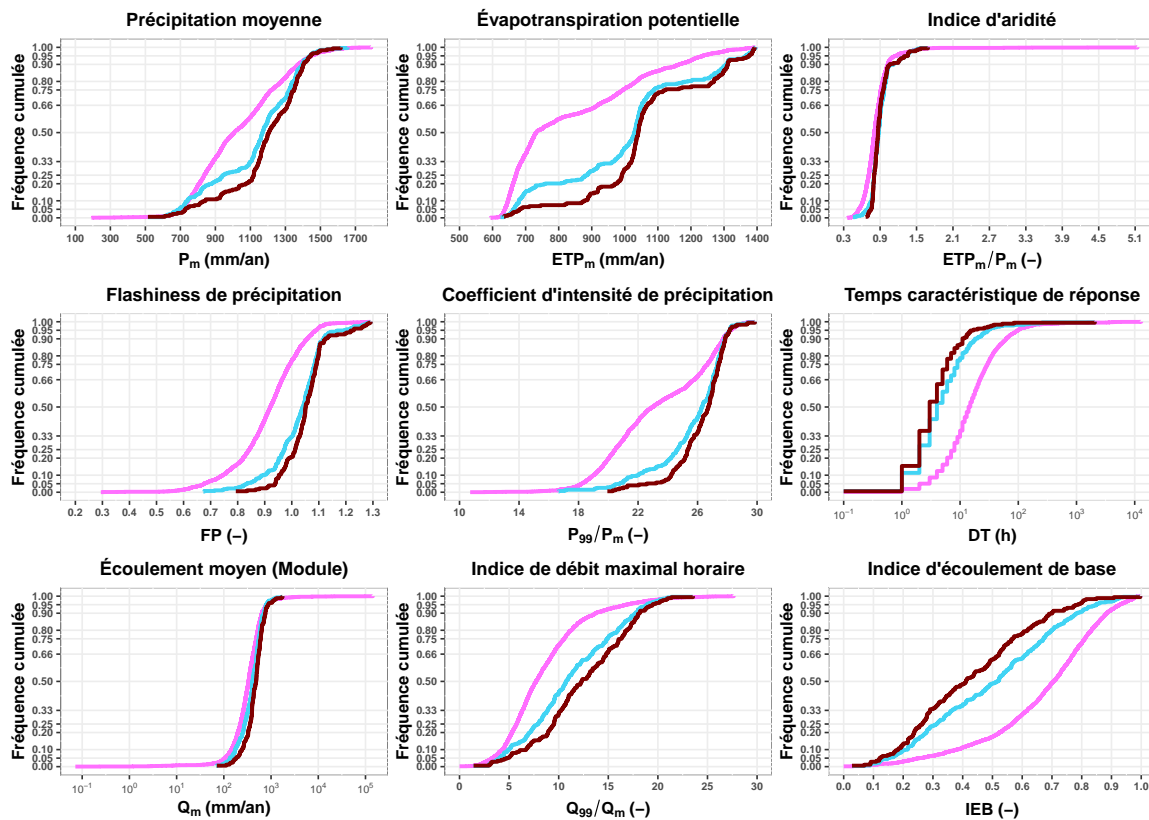


Figure 2.15. Courbes cumulées des caractéristiques hydroclimatiques pour l'échantillon total (en magenta), les bassins fortement urbanisés correspondant à $TIA \geq 5\%$ (en cyan) et à $TIA \geq 10\%$ (en marron).

Nous remarquons que les deux sous-échantillons fortement urbanisés couvrent la majeure partie des valeurs rencontrées dans l'échantillon total pour toutes les caractéristiques, ce qui confirme que la diversité climatique de l'échantillon est conservée même si on restreint l'étude sur les cas fortement urbanisés. En outre, nous constatons que les distributions des forçages climatiques sont un peu décalées par rapport à l'échantillon total, du fait que la majeure partie des bassins fortement urbanisés sont nord-américains (voir Figure 2.14). Sur les caractéristiques de la réponse hydrologique, l'écoulement moyen interannuel est plus important sur les bassins fortement urbanisés, mais les différences ne sont pas flagrantes : 330 mm/an en médiane pour l'échantillon total contre 400-480 mm/an pour l'échantillon des bassins pour lesquels $TIA \geq 5\%$ et $TIA \geq 10\%$. Ces différences sont plus importantes sur l'indice du débit maximal horaire, l'indice de l'écoulement de base et le temps caractéristique de la réponse, indiquant des extrêmes hydrologiques plus forts sur les bassins fortement urbanisés avec des temps caractéristiques plus courts et des écoulements de base plus faibles. La comparaison avec la Figure 2.14 indique que ces courbes sont plus décalées de l'échantillon total que celles des bassins nord-américains, ce qui suggère un effet lié à l'urbanisation (analysé plus en détail dans le Chapitre 4). Les différences en termes de temps caractéristique peuvent

être liées en outre à la faible taille des bassins fortement urbanisés (47-60 km² en médiane pour les bassins fortement urbanisés contre 222 km² pour l'échantillon total).

Nous concluons que l'échantillon des bassins fortement urbanisés est caractérisé par un gradient riche de situations hydroclimatiques. La variabilité spatiale de la réponse hydrologique (écoulement moyen interannuel) suit assez bien la variabilité spatiale des précipitations. Or, l'importance des apports en écoulement de base n'est pas cohérente avec cette distribution dans l'espace, car elle est non seulement fonction du régime climatique mais aussi de la morphologie et de la lithologie du bassin. Comme nous venons de le faire pour les caractéristiques climatiques, nous décrivons dans les deux prochaines sections les caractéristiques morphologiques et géopédologiques des bassins sélectionnés.

2.6 Caractéristiques morphologiques de l'échantillon

Nous désignons par morphologique toute caractéristique relevant de la forme géométrique (en deux ou trois dimensions) du bassin versant ou de son réseau de drainage. Cela comprend la surface, le périmètre, la pente, la densité de drainage, etc. Il s'agit d'un grand nombre de caractéristiques qui régissent la répartition des précipitations entre ruissellement et infiltration, et la dynamique de la réponse hydrologique.

Dans cette partie, nous allons décrire la variabilité spatiale des 4 métriques choisies pour caractériser la morphologie des bassins sélectionnés. Nous commençons par détailler ces métriques, puis nous décrivons leur variabilité en France et aux États-Unis. Ensuite, nous regardons les différences morphologiques entre les bassins français et nord-américains. Finalement, nous analysons si la diversité morphologique de l'échantillon total est conservée quand on focalise sur les bassins fortement urbanisés.

2.6.1 Caractéristiques morphologiques choisies

Nous avons calculé, à l'échelle de chaque bassin versant, les caractéristiques morphologiques suivantes :

1. La pente moyenne, notée *Pente* (-), extraite à partir de la base de données HDMA en utilisant un modèle numérique de terrain basé sur les produits HydroSHEDS, GMTED2010 et SRTM (Verdin, 2017). Elle est disponible sur tout le globe à une résolution de 3 secondes d'arc (environ 90 m) pour des latitudes inférieures à 60°. La pente moyenne pour chaque bassin a été prise comme la moyenne arithmétique des valeurs de pente

à partir des pixels du bassin versant. Une pente importante traduit des changements d'altitude importants au sein du bassin versant.

2. L'altitude moyenne, notée Alt (en m), extraite à partir de la base de données HDMA (Verdin, 2017). Comme pour la pente, il s'agit de la moyenne arithmétique des valeurs d'altitude au-dessus du niveau de la mer enregistrées par les pixels du bassin versant. Une altitude moyenne importante traduit la présence de reliefs dans le bassin.
3. L'indice topographique, noté CTI (pour *Compound Topographic Index*, sans unité¹⁰). Formulé au sein de TOPMODEL (Beven et Kirkby, 1979), il est supposé proportionnel aux variations spatiales du niveau de la nappe, ce qui permet de caractériser sa profondeur en connaissant la topographie. Pour cela, cet indice compare l'aire de recharge à l'amont de chaque pixel (c'est-à-dire l'aire topographique qui alimente le pixel) avec la pente locale de ce dernier suivant la définition :

$$CTI_i = \ln\left(\frac{A_i}{s_i}\right)$$

où A_i est l'aire spécifique en m (aire de recharge divisée par la longueur de parcours de l'amont jusqu'au pixel) et s_i (–) la pente locale du pixel i . Nous avons gardé cette formulation classique, malgré l'existence de formulations plus élaborées, permettant notamment de s'affranchir de la dépendance à la résolution du modèle numérique de terrain (Ducharne, 2009). À l'échelle du bassin versant, nous avons pris la médiane des valeurs CTI_i des pixels du bassin, extraites à partir de la base de données HDMA. Des valeurs de CTI importantes sont rencontrées dans les plaines en aval caractérisées par des pentes faibles et des surfaces de recharge en amont importantes.

4. La densité de drainage, notée DD (en km/km²), définie comme le rapport entre la longueur totale des cours d'eau pérennes du bassin versant (en km) et sa surface topographique (en km²). Des valeurs de DD importantes caractérisent un réseau d'eau pérenne dense sur le bassin, c'est-à-dire une lithologie moins perméable. L'estimation de DD peut se faire également en utilisant des modèles numériques de terrain se basant sur le concept de l'aire de drainage critique au-dessus de laquelle un pixel peut être considéré comme un cours d'eau (Schneider *et al.*, 2017). Cette surface critique peut même être utilisée pour remplacer la métrique classique de la densité de drainage (Le Moine, 2008, Annexe F). En milieu urbain, le réseau d'écoulement artificiel peut être intégré dans le réseau hydrographique par exemple en diminuant numériquement les altitudes des pixels sur lesquels se trouve le réseau artificiel (opération désignée par *stream burning*, voir Gironás *et al.*, 2010; Rodriguez, Bocher et Chancibault, 2013). Pour notre cas, nous avons utilisé les bases de données cartographiant les cours d'eau de chaque pays. En France, les cours d'eau sont extraits à partir de la BD Carthage® de l'IGN de 2014 (échelle de 1:50000), en utilisant la couche *Tronçon Hydrographique*. Seuls les

¹⁰Plus rigoureusement, l'unité est le mètre.

cours d'eau labellisés « Permanent » ont été pris en compte dans le calcul de la densité de drainage. Pour les bassins nord-américains, nous avons utilisé la base de données NHD de l'USGS (échelle de 1:24000), en exploitant la couche *NHDFlowline*, dont les réseaux ont été téléchargés grâce au package en R *FedData* (Bocinsky, Beaudette et Chamberlain, 2019) sur la période allant du 22 février 2019 au 06 mars 2019. Seuls les tronçons labellisés « *stream/river perennial* » (code : 46006¹¹) ont été considérés dans le calcul de la densité de drainage. L'estimation de cette mesure ne permet pas, malheureusement, de distinguer les cours d'eau naturels des cours d'eau artificiels.

En outre, la surface du bassin versant comme caractéristique morphologique de base a déjà été caractérisée dans la Section 2.3 et nous ne la mentionnons pas dans cette section. Elle a été estimée à partir des contours des bassins versants fournies par GAGES-II (Falcone, 2011) aux États-Unis, et à partir de la base de données des 3701 bassins situés en France (Bourgin *et al.*, 2010).

Se limitant donc à 4 caractéristiques, nous montrons dans la prochaine partie leur variabilité spatiale en France et aux États-Unis.

2.6.2 Variabilité spatiale des caractéristiques morphologiques en France et aux États-Unis

Les caractéristiques morphologiques des bassins versants situés en France sont montrées sur la Figure 2.16. À partir de l'altitude moyenne, on peut facilement discerner les bassins drainant une partie des hauts reliefs : le massif central, les Alpes dans le sud-est, les Pyrénées dans le sud-ouest, les Vosges et le Jura dans l'est. Des altitudes plus faibles mais traduisant malgré tout un relief sont observées dans le nord-ouest et dans le massif armoricain. Les altitudes moyennes varient entre 20 m et 1120 m au-dessus du niveau de la mer. La pente suit assez bien la variabilité de l'altitude : les bassins caractérisés par des altitudes moyennes importantes ont des pentes plus fortes. On remarque également que les régions caractérisées par de hauts reliefs et des pentes importantes possèdent des densités de drainage importantes, ce qui traduit le lien entre la pente et la présence de cours d'eau. Des conclusions similaires sont apportées par la variabilité spatiale du *CTI*, lequel enregistre des valeurs plus importantes dans les moyennes et basses altitudes vers la moitié nord/nord-ouest du pays et sur une partie du sud-ouest.

Comme pour la France, on distingue la présence de relief à partir de la variabilité spatiale de l'altitude moyenne aux États-Unis (Figure 2.17), à savoir les monts Ozarks au centre (dans les États d'Arkansas, Oklahoma, Kansas et Missouri) et les reliefs appalachiens dans l'est, où l'on

¹¹Voir la nomenclature complète sur : https://nhd.usgs.gov/userGuide/Robohelpfiles/NHD_User_Guide/Feature_Catalog/Hydrography_Dataset/Complete_FCode_List.htm.

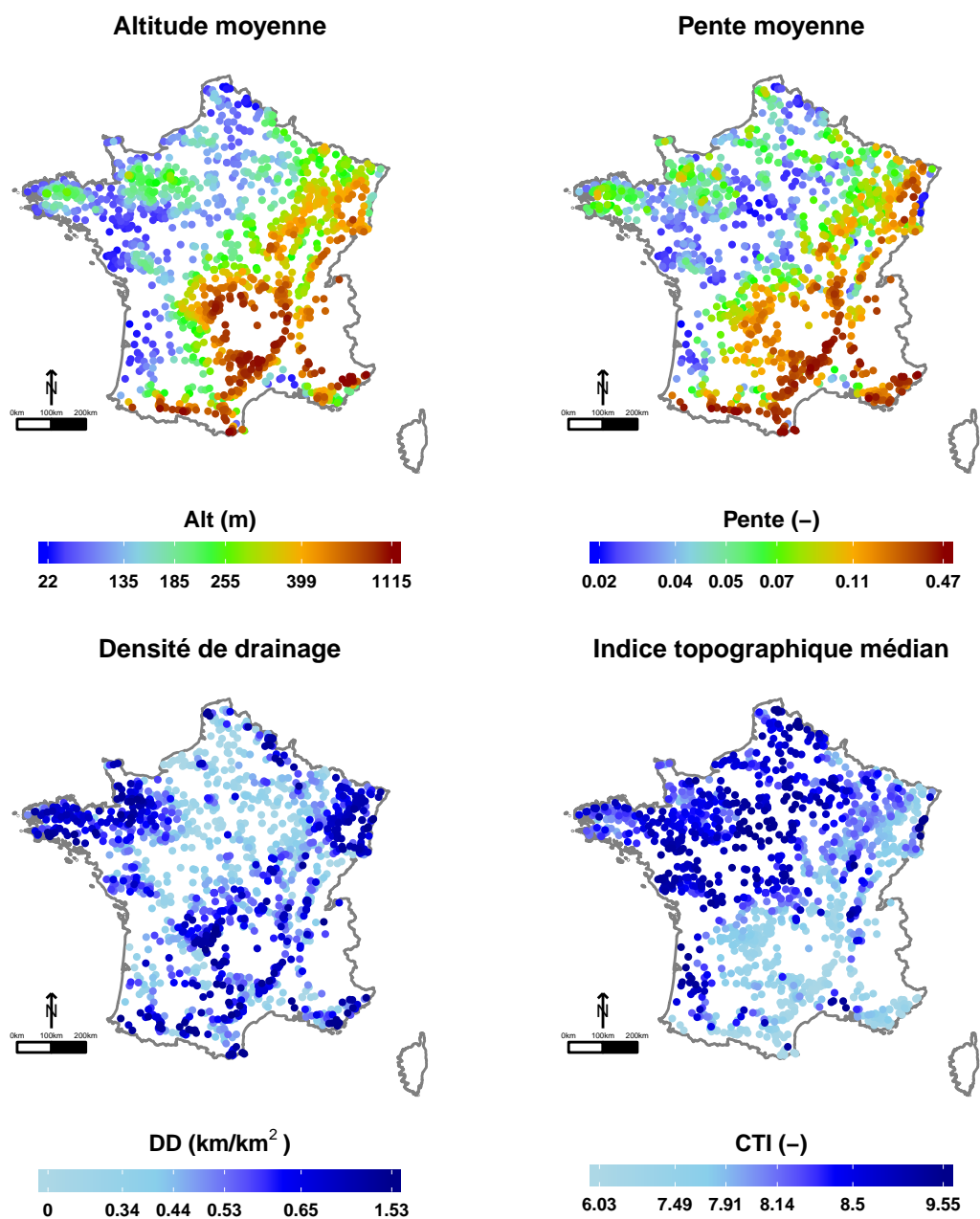


Figure 2.16. Caractéristiques morphologiques des 1355 bassins versants situés en France. Les points représentent les localisations des centroïdes des bassins versants. Pour chaque variable, on classe d'abord les bassins versants. Ensuite, la couleur est attribuée en fonction du rang du bassin versant (de 1 jusqu'à 1355). Cela permet de mieux dégager les contrastes régionaux pour chacune des variables.

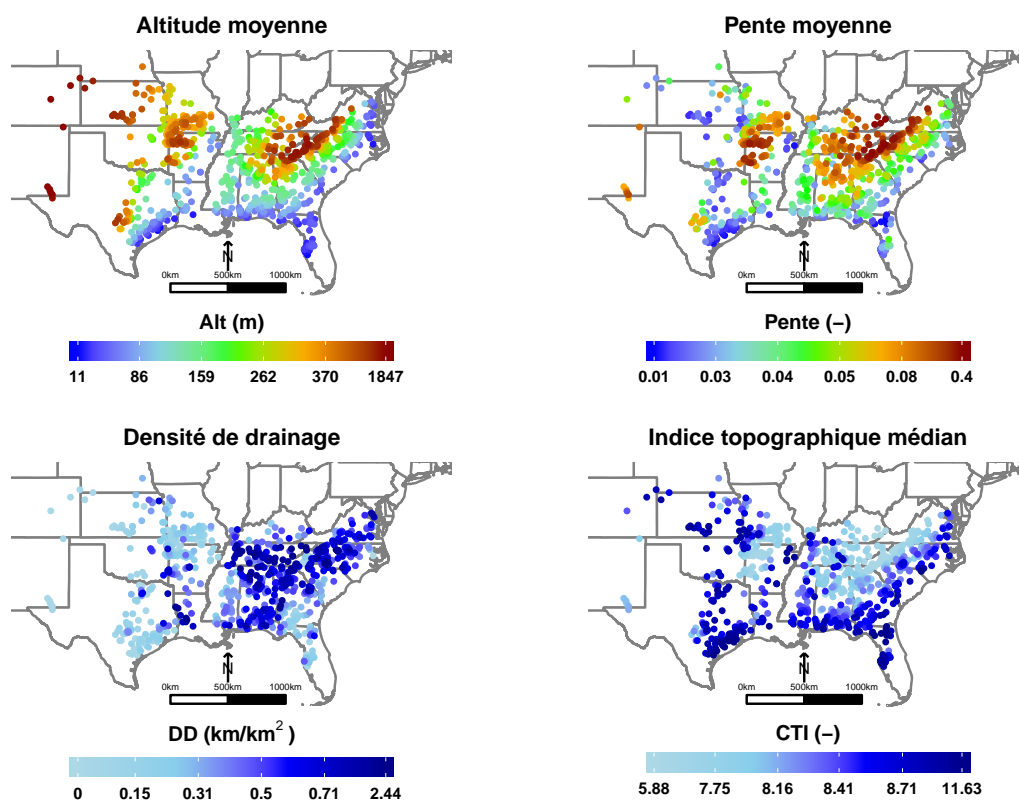


Figure 2.17. Caractéristiques morphologiques des 870 bassins versants situés aux États-Unis. Les points représentent les centroïdes des bassins versants. Pour chaque variable, on classe d'abord les bassins versants. Ensuite, la couleur est attribuée en fonction du rang du bassin versant (de 1 jusqu'à 870). Cela permet de mieux dégager les contrastes régionaux pour chacune des variables.

dépasse 1800 m d'altitude. La variabilité spatiale de la pente moyenne suit assez sensiblement celle de l'altitude moyenne, sauf pour les bassins au centre dans le Kansas, le Texas, le Colorado et le Nebraska. L'indice topographique varie de manière inversement proportionnelle à la pente, avec des valeurs plus importantes sur les plaines. La densité de drainage se trouve plus importante sur la rive gauche du Mississippi et sur les reliefs appalachiens, et plus faible dans les zones arides au Texas, au centre et vers la côte atlantique en Floride et au sud-est de la Géorgie.

Nous constatons donc la diversité de l'échantillon en termes d'altitude moyenne, en couvrant des bassins en plaine (environ 20 m au-dessus de la mer) comme en montagne (jusqu'à 1800 m d'altitude). Pour les deux pays, la pente suit fidèlement l'altitude, avec des pentes plus fortes pour des bassins à hauts reliefs, sauf pour quelques bassins nord-américains du centre à l'est des Rocheuses. Par ailleurs, l'indice topographique *CTI* se trouve plutôt important sur des pentes douces. La densité de drainage suit bien les reliefs dans le contexte tempéré en France, mais elle est plus faible sur les bassins arides au Texas et dans la partie centrale des États-Unis du centre à cause de l'assèchement des cours d'eau. Nous allons voir dans quelle mesure les contrastes morphologiques sont différents entre les deux pays.

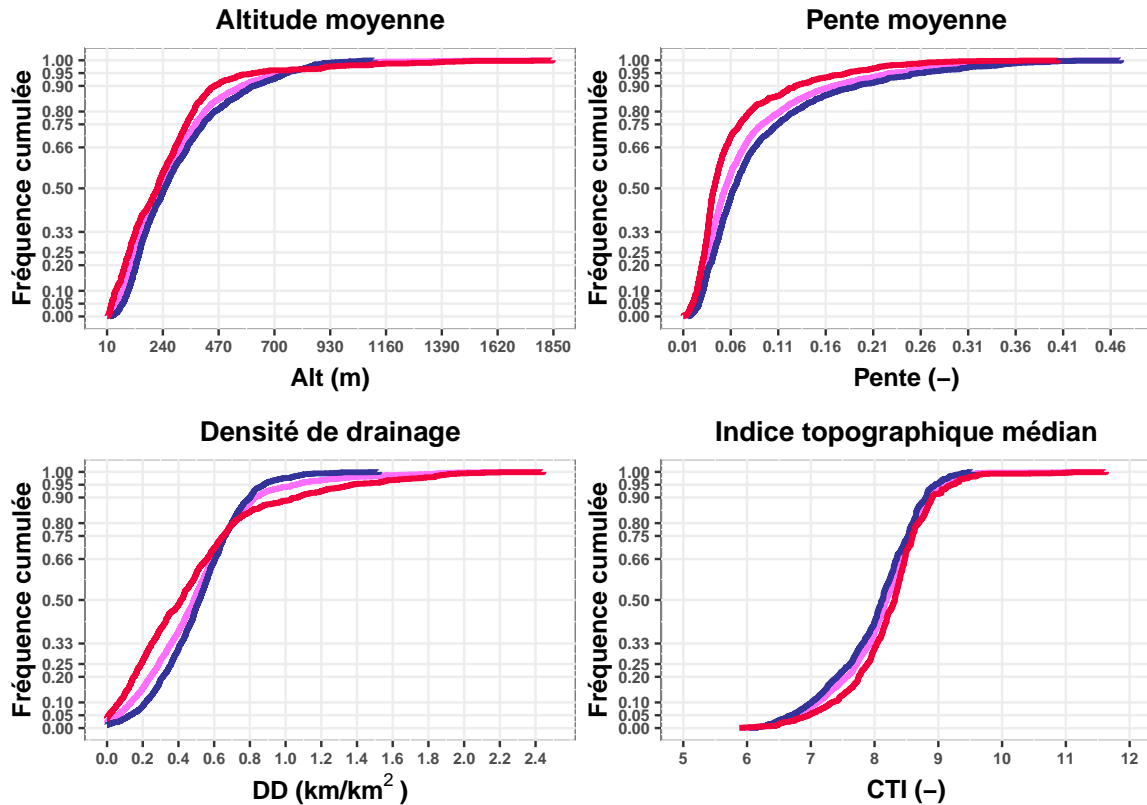


Figure 2.18. Courbes cumulées des caractéristiques morphologiques de l'échantillon total (en magenta), du sous-échantillon des bassins situés en France (en bleu) et du sous-échantillon des bassins situés aux États-Unis (en rouge).

2.6.3 Différences morphologiques entre les bassins situés en France et les bassins situés aux États-Unis

Quelques bassins versants nord-américains contribuent à la richesse de l'échantillon en couvrant des altitudes, des densités de drainage et des indices topographiques importants, comme on peut le voir sur la Figure 2.18. En dehors de ces quelques bassins, peu de différences existent entre les sous-échantillons français et nord-américain, comme c'est confirmé par les courbes assez serrées de la Figure 2.18 et les valeurs médianes similaires en termes de pente (0.04 pour l'échantillon nord-américain contre 0.06 en France), d'altitude moyenne (218 m pour l'échantillon nord-américain contre 247 m en France), de densité de drainage (0.41 km/km² pour l'échantillon nord-américain contre 0.52 km/km² en France), et d'indice topographique médian (8.32 pour l'échantillon nord-américain contre 8.12 en France). Dans la prochaine partie, nous analysons si les bassins versants fortement urbanisés couvrent les mêmes plages de variation que l'échantillon total en termes de caractéristiques morphologiques.

2.6.4 Caractéristiques morphologiques des bassins fortement urbanisés

Les différences des caractéristiques morphologiques de deux échantillons de bassins versants fortement urbanisés par rapport à l'échantillon total sont montrées sur la Figure 2.19. Rappelons que ces deux échantillons ont été sélectionnés sur la base de deux seuils de TIA de 5% et 10%. Nous constatons qu'en général, les bassins fortement urbanisés couvrent plus que 90% des valeurs des caractéristiques morphologiques de l'échantillon total. Les bassins fortement urbanisés se situent plutôt dans des régions de faible altitude et de faible pente. Ils sont caractérisés par un indice topographique et une densité de drainage légèrement plus importants que l'échantillon total.

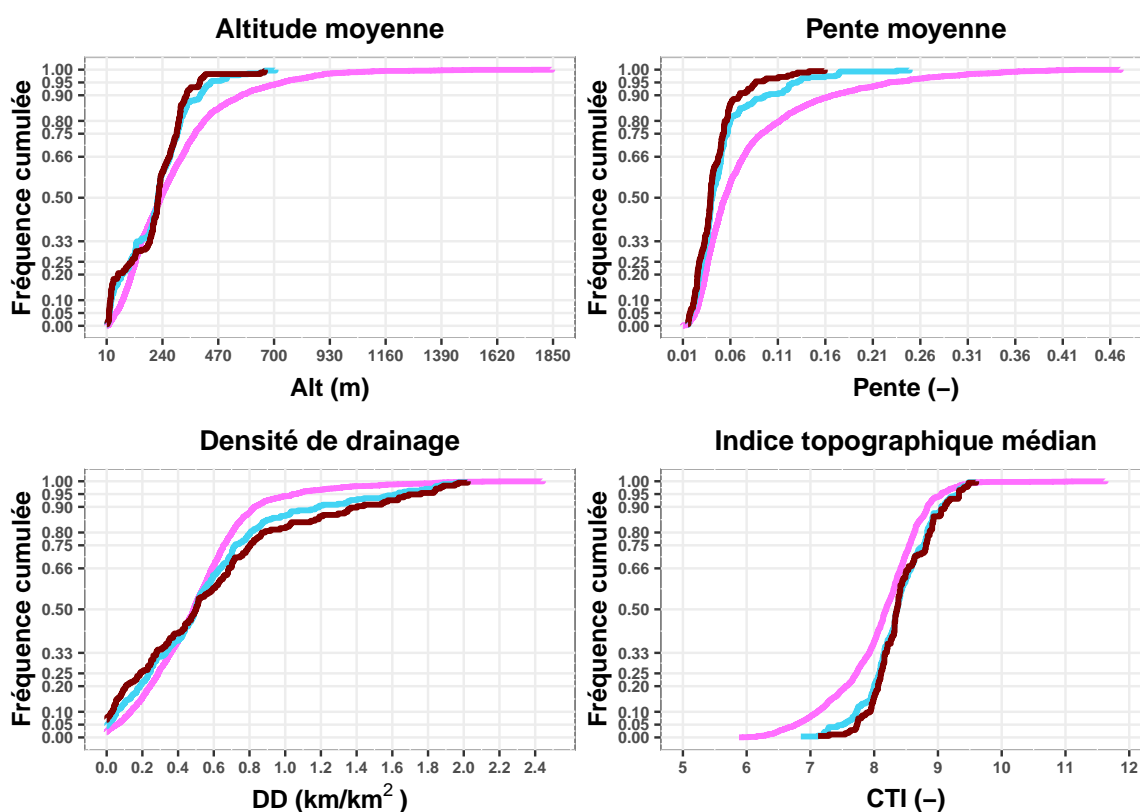


Figure 2.19. Courbes cumulées des caractéristiques morphologiques pour l'échantillon total (en magenta), les bassins fortement urbanisés correspondant à $TIA \geq 5\%$ (en cyan) et à $TIA \geq 10\%$ (en marron).

Ainsi, on remarque une richesse au niveau de l'échantillon total en termes morphologiques avec peu de différences entre les bassins français et nord-américains et entre l'échantillon total et les sous-échantillons de bassins fortement urbanisés. Une comparaison au niveau des caractéristiques géopédologiques complète ci-dessous la caractérisation de l'échantillon.

2.7 Caractéristiques géopédologiques de l'échantillon

Les caractéristiques géopédologiques concernent tout ce qui relève de la texture du sol et des couches géologiques du sous-sol. Elles contrôlent la répartition entre infiltration et ruissellement, la dynamique des écoulements, le stockage souterrain et les échanges nappe-rivière. Leur description complète donc le tableau dressé pour les autres classes de caractéristiques (occupation du sol, climat, régime hydrologique et morphologie).

La caractérisation précise de la géologie et de la pédologie d'un bassin versant ne sera pas abordée ici. Nous nous sommes contentés donc de quelques paramètres hydrodynamiques importants (porosité, perméabilité) et de texture du sol (contenu en gravier, sable, argile et limon), qui sont faciles à obtenir et issus de bases de données homogénéisées sur les deux continents. Nous conservons le même schéma de présentation ; d'abord, une définition des 6 métriques utilisées, puis une description de la variabilité spatiale en France et aux États-Unis, ensuite une discussion des différences entre les bassins nord-américains et français, et enfin un focus sur les bassins fortement urbanisés.

2.7.1 Caractéristiques géopédologiques choisies

Pour chaque bassin versant, nous avons déterminé deux caractéristiques hydrodynamiques, la porosité totale et la perméabilité intrinsèque, ainsi que quatre caractéristiques de la texture du sol et sous-sol, mesurant le contenu en gravier, en sable, en argile et en limon.

Les paramètres hydrodynamiques proviennent de la base de données GLHYMPS (Gleeson *et al.*, 2014), issue d'une compilation de données lithologiques à haute résolution à l'échelle du globe. Cette base de données caractérise la lithologie sur tout le globe à l'aide de deux paramètres : la porosité totale (*Por*, sans unité) et la perméabilité intrinsèque (*Perm*, en m^2) grâce à un ensemble de polygones, dont la taille médiane est de 100 km^2 . Pour chaque bassin, l'intersection avec ces polygones est effectuée. Puis, en pondérant par l'aire intersectée avec chaque polygone, nous avons déterminé la moyenne de la porosité totale et de la perméabilité intrinsèque.

Pour la texture du sol, nous avons utilisé la base de données raster HWSD, version 1.2, produite conjointement par les organismes FAO, IIASA, ISRIC, ISS-CAS et JRC (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012). Cette base de données contient des caractéristiques physico-chimiques des couches du sol (*top soil*, 0-30 cm) et du sous-sol (*sub soil*, 30-100 cm), couvrant tout le globe avec une résolution de 1/120°. Pour chaque unité homogène, définie à l'aide d'un identifiant global, nous avons déterminé le contenu en gravier (diamètre supérieur à 2 mm), sable (diamètre entre 0.0625 mm et 2 mm), argile (diamètre inférieur à 0.002 mm) et limon (diamètre entre 0.002 mm

et 0.0625 mm) sur la hauteur du sol, avec une pondération de 30% et 70% des contenus du *top soil* et du *sub soil*, respectivement. Pour chaque bassin versant, nous avons pondéré les contenus de chacune des unités par la proportion du bassin versant qu'elles occupent. Ceci a permis d'obtenir des contenus moyens à l'échelle du bassin versant en gravier (en %, noté *Gravier*), sable (en %, noté *Sable*), argile (en %, noté *Argile*) et limon (en %, noté *Limon*).

Nous allons donc regarder la variabilité spatiale de ces six caractéristiques sur le territoire français, puis sur la partie des États-Unis que couvre l'échantillon nord-américain.

2.7.2 Variabilité spatiale des caractéristiques géopédologiques en France et aux États-Unis

En France, on peut discerner des régions avec d'importantes valeurs de porosité (Figure 2.20) : le massif armoricain, une partie du bassin parisien, les Vosges, et dans le bassin aquitain. Les bassins dans le nord-ouest du massif central ainsi qu'en Bourgogne montrent des valeurs plutôt faibles. La porosité moyenne couvre une large gamme allant de 0.01 jusqu'à 0.27. En revanche, la perméabilité intrinsèque est très faible sur le massif armoricain et dans le bassin aquitain, contrairement au bassin parisien où l'on rencontre des valeurs importantes. La perméabilité intrinsèque varie de 10^{-11} m² jusqu'à 10^{-8} m² (valeurs équivalentes à des conductivités hydrauliques à saturation de 10^{-4} m/s à 10^{-1} m/s). Des contenus importants en gravier et en sable sont retrouvés dans presque les mêmes régions : le massif armoricain, la Vendée, le bassin aquitain et autour du massif central. Les zones de forte teneur en argile et en limon sont opposées aux zones où le contenu en sable prédomine : les bassins situés en Bourgogne et au massif armoricain ont les plus fortes teneurs en argile et en limon, avec une préférence du limon à l'ouest et une préférence de l'argile à l'est. On peut conclure que les bassins versants en France sont caractérisés par une prédominance du contenu en sable et en limon par rapport au contenu en argile.

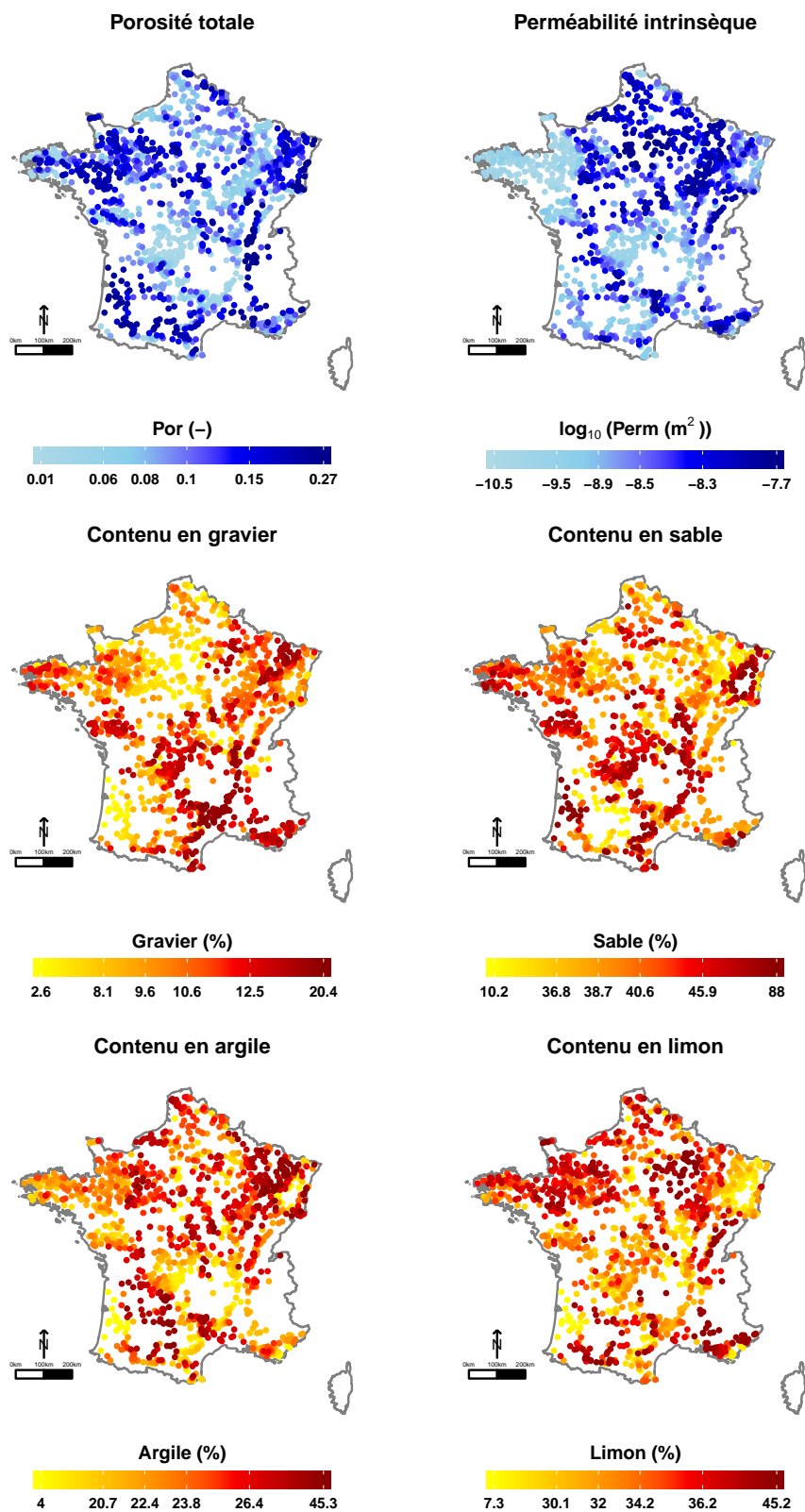


Figure 2.20. Caractéristiques géopédologiques des 1355 bassins versants situés en France. Les points représentent les centroïdes des bassins versants. Pour chaque variable, on classe d'abord les bassins versants. Ensuite, la couleur est attribuée en fonction du rang du bassin versant (de 1 jusqu'à 1355). Cela permet de mieux dégager les contrastes régionaux pour chacune des variables.

Aux États-Unis (Figure 2.21), de faibles valeurs de porosité sont rencontrées sur le massif appalachien, dans les états du centre (le Missouri) et plus à l'ouest (Texas). Sur la partie des bassins situés plus à l'ouest, la perméabilité intrinsèque est plus importante au nord mais très faible au Texas dans le sud et sur le massif appalachien. La porosité et la perméabilité sont relativement importantes sur les plaines. Leurs valeurs varient de 0.01 jusqu'à 0.28 en porosité, et de 10^{-11} m² jusqu'à 10^{-8} m² en perméabilité intrinsèque (valeurs équivalentes à des conductivités hydrauliques à saturation de 10^{-4} m/s à 10^{-1} m/s). Des contenus importants en sable et en gravier sont observés sur les plaines côtières atlantiques et sur la Floride, avec des valeurs moins importantes à proximité du Mississippi et beaucoup plus faibles en allant vers l'ouest. En revanche, l'argile prédomine plutôt à l'ouest dans le Texas et à proximité du Mississippi, mais avec des valeurs plus faibles sur les bassins du Nouveau-Mexique. Quant au limon, il est plutôt prédominant au centre nord dans le Kansas et à l'ouest au centre du Texas. On constate une prédominance du contenu en sable, suivie par le contenu en argile puis le contenu en limon.

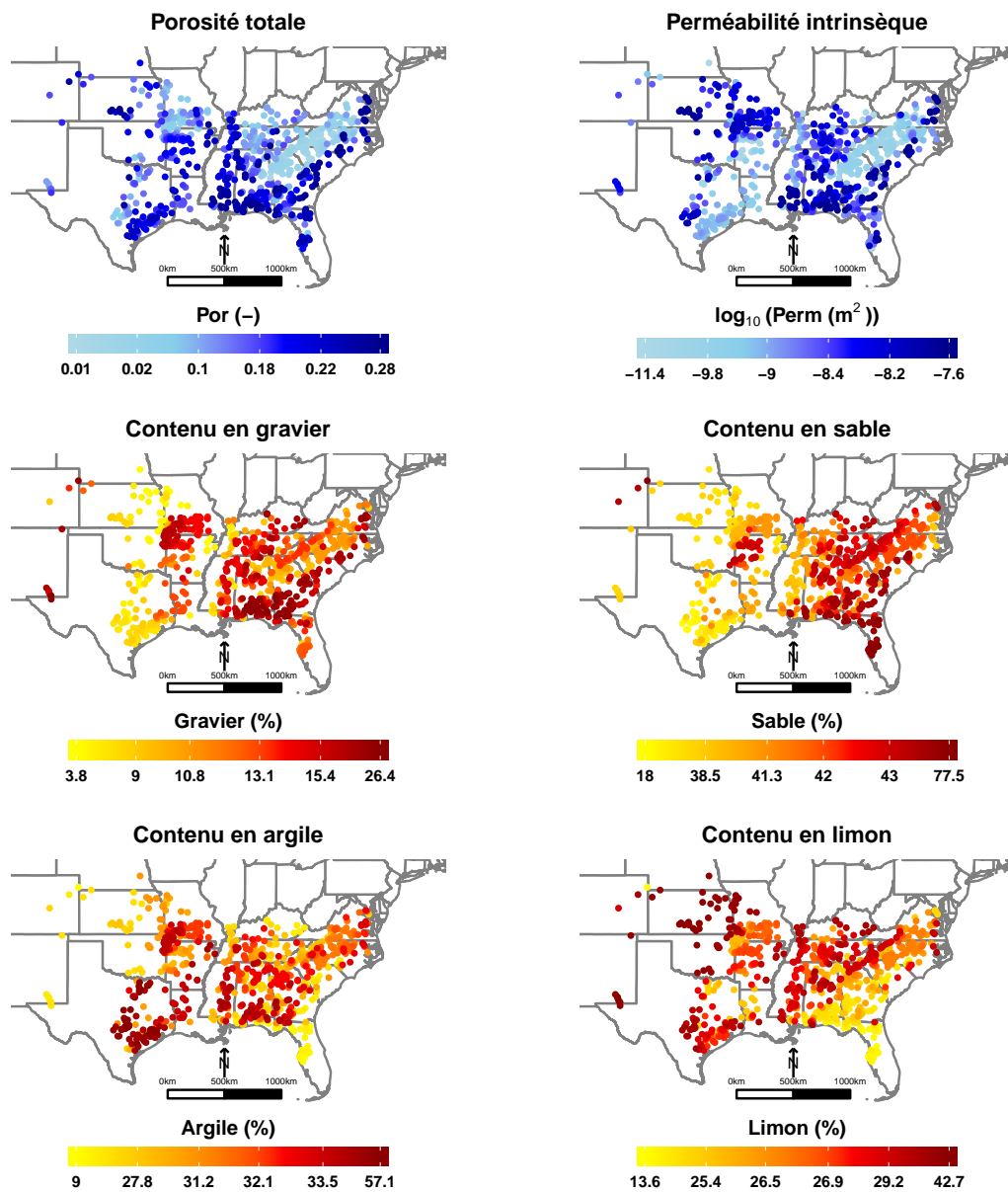


Figure 2.21. Caractéristiques géopédologiques des 870 bassins versants situés aux États-Unis. Pour chaque variable, on classe d'abord les bassins versants. Ensuite, la couleur est attribuée en fonction du rang du bassin versant (de 1 jusqu'à 870). Cela permet de mieux dégager les contrastes régionaux pour chacune des variables.

2.7.3 Contrastes géopédologiques entre les bassins situés en France et les bassins situés aux États-Unis

Les bassins versants français montrent des distributions de caractéristiques géopédologiques similaires aux bassins nord-américains (Figure 2.22). Néanmoins, on note des valeurs de porosité plus importantes, globalement, sur les bassins nord-américains (valeur médiane : 0.15 contre 0.1 en France). Cette différence peut être mise en relation avec les différences de teneur en argile, pour laquelle les bassins nord-américains montrent des taux de 10% en plus en médiane par rapport à l'échantillon français. En médiane, les bassins sélectionnés sont plutôt sableux (40%), avec plus de contenu en argile aux États-Unis et plus de contenu en limon aux bassins français.

2.7.4 Caractéristiques géopédologiques des bassins fortement urbanisés

Comme nous l'avons constaté pour les caractéristiques hydroclimatiques et morphologiques, on ne perd pas la richesse de l'échantillon en termes géopédologiques quand on focalise sur les bassins fortement urbanisés (Figure 2.23). La porosité totale et la perméabilité intrinsèque sont légèrement plus faibles sur les bassins fortement urbanisés. Les contenus en gravier et en sable sont plutôt similaires, alors que l'on retrouve un peu la particularité des bassins versants nord-américains en analysant les contenus en argile et en limon. Ces légères différences n'empêchent pas de conclure sur la diversité des situations géopédologiques qu'offre les bassins fortement urbanisés.

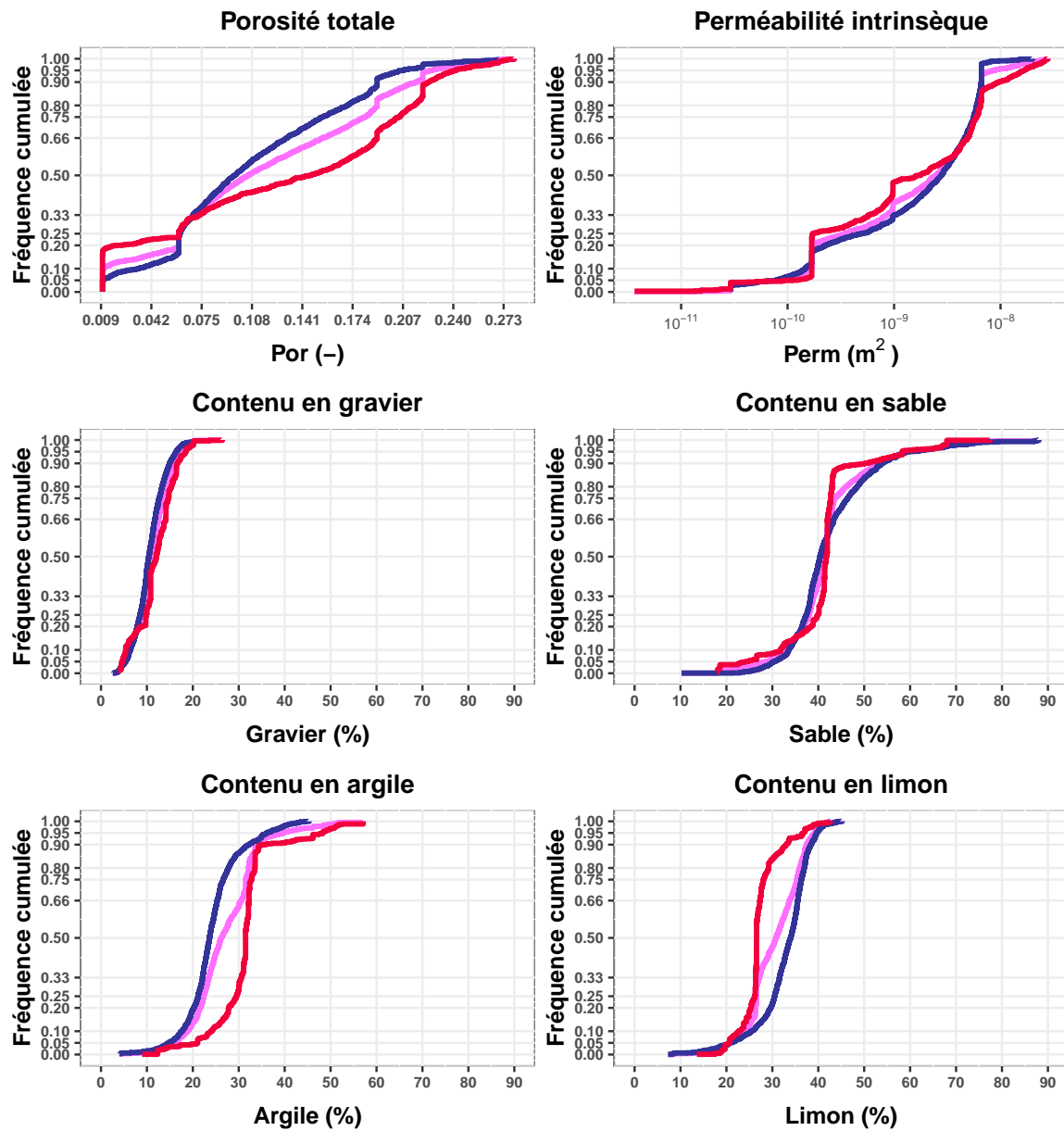


Figure 2.22. Courbes cumulées des caractéristiques géopédologiques de l'échantillon total (en magenta), des bassins situés en France (en bleu) et des bassins situés aux États-Unis (en rouge).

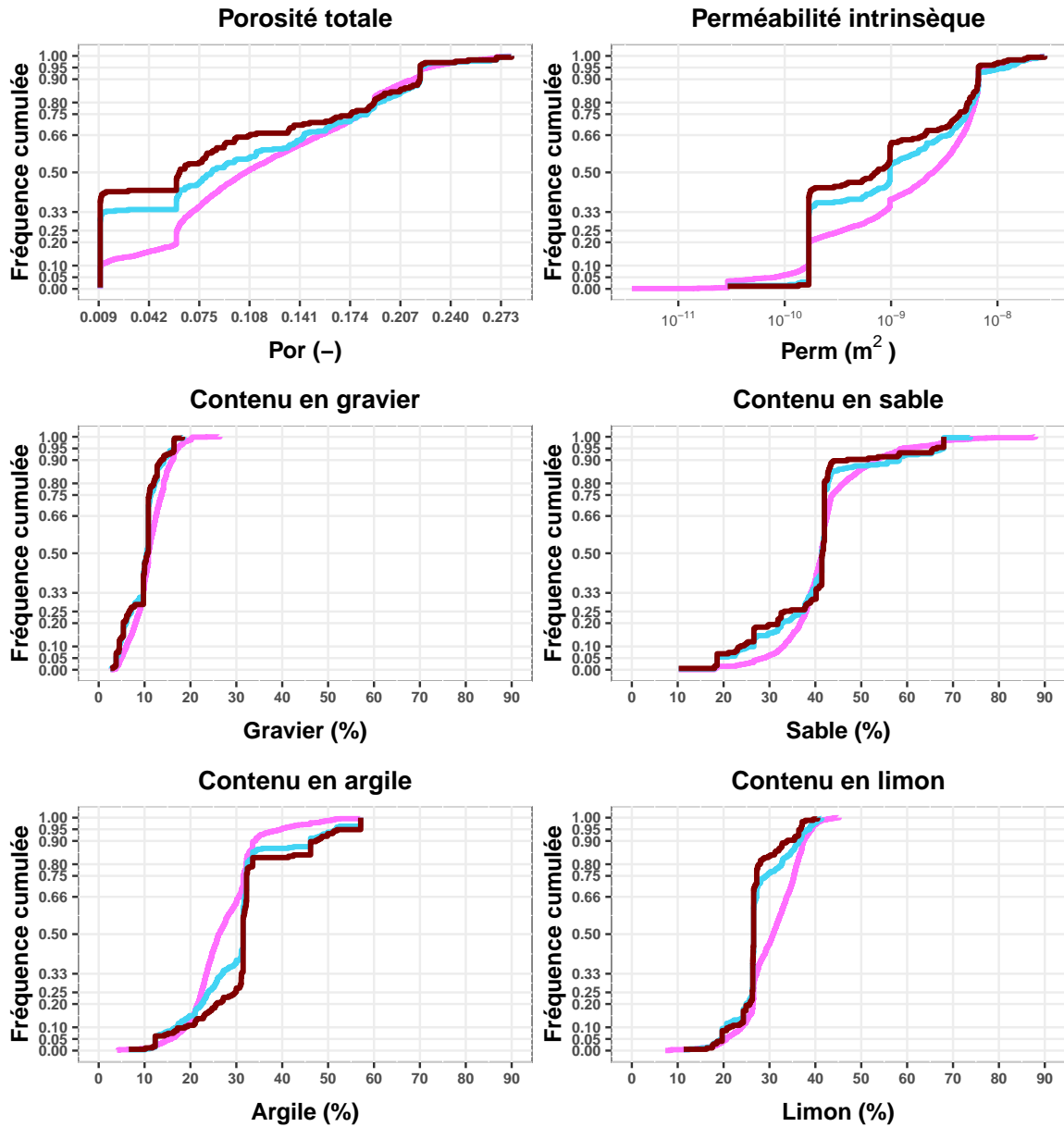


Figure 2.23. Courbes cumulées des caractéristiques géopédologiques pour l'échantillon total (en magenta), les bassins fortement urbanisés avec $TIA \geq 5\%$ (en cyan) et avec $TIA \geq 10\%$ (en marron).

2.8 Synthèse

2.8.1 Un large échantillon avec des caractéristiques climatiques et d'urbanisation très variées

Notre recherche d'une structure de modèle pour prédire l'impact de l'urbanisation sur l'hydrologie se fait avec un large échantillon de bassins versants pour lesquels le modèle doit reproduire de manière satisfaisante les écoulements observés. Dans le présent chapitre, nous avons montré la procédure de préparation des séries hydroclimatiques nécessaires au calage et à la validation du modèle. Puis, nous avons choisi trois critères pour sélectionner 2225 bassins versants à partir d'un échantillon initial de 12768 bassins versants situés en France et aux États-Unis. Ensuite, nous avons discuté des caractéristiques d'occupation du sol, en particulier, le couvert urbain, qui est le centre d'intérêt de nos travaux. Cette caractérisation d'occupation du sol a été complétée par une description des caractéristiques hydroclimatiques, morphologiques et géopédologiques, en France, aux États-Unis, en illustrant les différences entre les deux pays, et plus particulièrement les contrastes entre l'échantillon total et les sous-échantillons de bassins fortement urbanisés (273 bassins versants avec un taux moyen d'imperméabilité totale $TIA \geq 5\%$, et 175 bassins versants avec $TIA \geq 10\%$). Nous en tirons les conclusions suivantes :

1. L'échantillon représente une grande diversité de situations d'urbanisation, matérialisée par des fractions de couvert urbain allant de 0% à 100%. Si la distribution n'est pas pour autant uniforme, à cause de la prédominance des bassins ruraux, le nombre de bassins fortement urbanisés est important et suffisamment large pour couvrir la plupart des situations d'urbanisation à l'échelle du bassin versant. En outre, l'inclusion des bassins ruraux est utile car elle permet de traiter le problème de manière relative dans un cadre comparatif fortement/faiblement urbanisé. Cela permet également de conserver une continuité dans le gradient d'urbanisation, ce qui n'était pas possible dans le cas d'un échantillon de faible taille.
2. L'échantillon est caractérisé par une diversité importante de situations climatiques, morphologiques et géopédologiques. Les bassins versants sélectionnés proviennent des régions sous des climats arides, tempérés et méditerranéens. Ils couvrent des altitudes de 10 m à 1800 m, leurs tailles vont de 1 km² jusqu'à 10⁵ km², et les caractéristiques de leurs sols et aquifères couvrent des plages de valeurs suffisamment larges. Ceci se traduit par des densités de drainage contrastées et des réponses hydrologiques très variées. Enfin, une sélection des bassins versants fortement urbanisés à partir des seuils de TIA de 5% et 10% n'engendre pas des différences importantes avec l'échantillon total.

Nous avons donc collecté une grande diversité de bassins versants pour tester la robustesse du modèle et son aptitude à reproduire des comportements hydrologiques associés à des situations climatiques et d'urbanisation très diversifiées. Nous synthétisons dans le Tableau 2.4 les distributions des différentes caractéristiques présentées dans ce chapitre.

Tableau 2.4. Synthèse des caractéristiques de l'échantillon des 2225 bassins versants sélectionnés. Les taux de chevauchement représentent la fraction de l'échantillon total dont les valeurs de la caractéristique en question tombent dans l'intervalle défini par la sélection des bassins fortement urbanisés. Plus les taux sont importants, moins on perd en diversité quand on focalise sur l'échantillon des bassins fortement urbanisés.

Classe de caractéristiques	Désignation	Notation et unité	Min	Q25%	Q50%	Q75%	Max	Taux de chevauchement (%)	
								$T/A \geq 5\%$	$T/A \geq 10\%$
-	Longueur de la série hydroclimatique	Longueur (ans)	8	15	16	16	16	100	100
	Surface	S (km ²)	0.49	80	222	680	110000	91	76.2
	Pente moyenne	$Pente$ (-)	0.01	0.04	0.05	0.09	0.47	95.3	88.4
	Altitude moyenne	Alt (m)	11	129	233	374	1847	94.4	93.3
	Densité de drainage	DD (km/km ²)	0	0.29	0.49	0.65	2.44	99.8	99.8
Morphologie	Indice topographique	CTI (-)	5.88	7.72	8.19	8.55	11.63	93.5	89.3
	Précipitation moyenne	P_m (mm/an)	192	836	1007	1219	1791	99.1	98.7
	ETP moyenne	ETP_m (mm/an)	594	670	731	994	1396	99.2	96
	Écoulement moyen	Q_m (mm)	0.07	224	334	486	143700	96.7	96.6
	Indice d'aridité	ETP_m/P_m (-)	0.36	0.71	0.81	0.91	5.14	98.7	83.6
Climat et régime hydrologique	Flashiness de précipitation	FP (-)	0.29	0.84	0.92	0.99	1.29	95.4	84.4
	Indice d'écoulement de base	IEB (-)	0	0.56	0.71	0.81	1	99.8	99.8
	Coefficient d'intensité de précipitation	P_{99}/P_m (-)	10.72	20.46	23.01	26.66	29.86	98.7	81.2
	Indice de débit maximal horaire	Q_{99}/Q_m (-)	0.00	5.75	7.71	10.53	27.66	98.9	98.2

Suite à la page suivante...

Tableau 2.4 suite : Synthèse des caractéristiques de l'échantillon des 2225 bassins versants sélectionnés.

Classe de caractéristiques	Désignation	Notation et unité	Min	Q25%	Q50%	Q75%	Max	Taux de chevauchement (%) <i>TIA</i> ≥ 5%	<i>TIA</i> ≥ 10%
	Temps de réponse caractéristique	<i>DT</i> (h)	0	8	14	29	12946	99.5	99.5
Occupation du sol	Fraction du couvert urbain	<i>CPD</i> (-)	0.00	0.071	0.038	0.069	1.00	24.6	9.8
	Taux moyen d'imperméabilité totale	<i>TIA</i> (-)	0	0.005	0.009	0.019	0.588	12.3	7.9
	Fraction du bassin avec une imperméabilité > 80%	<i>IMP80</i> (-)	0	0.001	0.002	0.005	0.387	64.9	32.8
	Irrégularité de l'imperméabilité	<i>IrIMP</i> (-)	0	0.006	0.01	0.018	0.465	88.8	88.4
	Fraction du couvert forestier	<i>fFOR</i> (-)	0	0.122	0.289	0.498	1.00	89.9	77.8
	Fraction d'eau	<i>fW</i> (-)	0	0	0.001	0.003	0.217	99.8	99.8
Géopédologie	Porosité totale	<i>Por</i> (-)	0.01	0.06	0.106	0.182	0.280	99.9	99.9
	Perméabilité intrinsèque	<i>Perm</i> ($\log_{10} \text{ m}^2$)	-11.4	-9.5	-8.6	-8.2	-7.6	98.8	98.8
	Contenu en gravier	<i>Gravier</i> (%)	2.6	8.9	10.8	13.7	26.4	97.8	96.8
	Contenu en sable	<i>Sable</i> (%)	10.2	38	41.3	43.7	88	99.1	98.3
	Contenu en limon	<i>Limon</i> (%)	7.3	26.6	30.8	35.4	45.2	98.5	97.5
	Contenu en argile	<i>Argile</i> (%)	4	22.4	26.2	31.5	57.1	99.5	99.5

Fin du Tableau 2.4.

2.8.2 Discussion sur les choix de modélisation : représentation globale au pas de temps horaire

Nous avons choisi de représenter la relation pluie-débit au pas de temps horaire. Nous remarquons que peu de bassins versants ont eu un temps caractéristique moyen, déterminé à partir de la corrélation pluie-débit sur une période d'au moins 8 ans, inférieur à 1 heure (4/2225). Si l'on suit les exigences d'Obled, Zin et Hingray (2009) pour une modélisation parcimonieuse, le pas de temps de discrétisation doit être inférieur à un tiers du temps caractéristique de la réponse. Sur les 2225 bassins versants, le pas de temps horaire est satisfaisant pour 2113 bassins. Les 112 bassins versants restants pour lesquels le pas de temps horaire est en dehors de cette limite (mais dont la réaction est supra-horaire) contiennent 75 bassins versants ayant un taux moyen d'imperméabilité totale qui dépasse 5% (sur un total de 273 bassins versants ayant *TIA* supérieure à 5%). En outre, sur la base d'une caractérisation des champs de précipitations par des méthodes de krigeage, le choix de discrétisation spatiale recommandé par Obled, Zin et Hingray (2009) considère qu'au pas de temps horaire, la pluie peut être considérée homogène sur une entité spatiale allant jusqu'à 300 km². Le nombre de bassins versants ayant une taille supérieure à 300, 600 et 1200 km² est de 937, 613, et 340, respectivement. L'approche globale est satisfaisante pour la majorité mais il existe un nombre important de bassins versants qui pourraient être discrétisés si l'on voulait tenir compte de l'hétérogénéité spatiale du forçage.

Ce commentaire des choix de pas de temps et d'espace par rapport à l'échantillon de bassins versants collecté montre que le pas de temps horaire est acceptable dans la majorité des cas. Si pour une bonne partie de l'échantillon le choix de la discrétisation spatiale n'est pas suffisant, il est néanmoins imposé par nos objectifs de travailler à l'échelle du bassin versant selon une approche top-down.

2.8.3 Échantillonnage des bassins versants pour les prochains chapitres

L'échantillon des bassins versants présenté ne sera pas complètement mobilisé pour tous les chapitres que l'on présentera par la suite. Nous avons procédé par quelques sélections exigées par les objectifs de chaque chapitre.

Au Chapitre 3, nous montrons une étude de la variabilité des coefficients de ruissellement en fonction du taux moyen d'imperméabilité totale *TIA* et d'autres caractéristiques temporelles (comme l'humidité du sol) et spatiales (comme la morphologie et le climat). L'étude est effectuée dans un cadre événementiel à partir des événements dont les caractéristiques sont montrées dans la Figure 2.7. Nous avons écarté l'échantillon français car il présentait moins de situations d'urbanisation que l'échantillon nord-américain. Ainsi, nous n'avons utilisé que les

70227 événements nord-américains, collectés à partir de 852 bassins versants (18/870 n'ont pas eu d'événements sélectionnés).

Au Chapitre 4, nous comparons les performances de deux modèles hydrologiques conceptuels appliqués à des bassins versants urbanisés avec leurs performances sur les bassins versants non urbanisés. Pour cela, nous avons choisi les bassins versants les plus fortement urbanisés dans l'échantillon, suivant un seuil minimal de $TIA = 10\%$. Comme échantillon témoin, nous avons choisi leurs voisins ruraux les plus proches qui ont un TIA inférieur à 5%. Ce choix a donné 175 bassins versants urbanisés et 175 non urbanisés. Sur les 350 bassins mobilisés, 312 sont situés aux États-Unis et 38 en France.

Au Chapitre 5, nous avons utilisé tout l'échantillon pour régionaliser la structure d'un modèle conceptuel. Ceci a permis d'exploiter toute l'information contenue dans l'échantillon collecté pour définir des relations de régression en utilisant un algorithme d'apprentissage automatique : les forêts aléatoires (*random forest*).

Au Chapitre 6, nous avons testé plusieurs modifications de structure d'un modèle conceptuel pour tenir compte des spécificités de l'urbanisation. Le test a donc visé les bassins dont les niveaux d'imperméabilité sont les plus importants. Nous avons choisi un seuil de 5% de TIA , lequel est reporté dans la littérature comme le seuil minimal au-delà duquel le changement dans le comportement hydrologique, à cause de l'urbanisation, devient significatif. Il permet en outre d'effectuer les tests sur une large gamme de TIA . Sur les 2225 bassins versants, 273 satisfont ce seuil, dont 205 sont situés aux États-Unis et 68 en France.

La synthèse des sous-échantillons utilisés dans chaque chapitre est donnée dans le Tableau 2.5.

Tableau 2.5. Synthèse des échantillons de bassins versants utilisés dans chacun des prochains chapitres.

Chapitre	Objectif	Nombre de bassins	Nord-américains	Français
Chapitre 3	Étudier le rôle de contrôles spatiaux et temporels dans la variabilité du coefficient de ruissellement	852	852	0
Chapitre 4	Comparaison des performances de deux modèles conceptuels sur des bassins versants urbanisés et ruraux	350, dont 175 avec $TIA \geq 0.1$ et 175 avec $TIA \leq 0.05$	312	38
Chapitre 5	Régionalisation des paramètres d'un modèle conceptuel	2225	870	1355
Chapitre 6	Test de plusieurs modifications d'un modèle conceptuel pour tenir compte des spécificités des bassins urbanisés	273, pour lesquels $TIA \geq 0.05$	205	68

CHAPITRE 3

PROCESSUS DE GÉNÉRATION DU RUISSELLEMENT EN MILIEU URBANISÉ : RÔLES DE L'IMPERMÉABILITÉ ET DE L'HUMIDITÉ DU SOL

Sommaire

3.1	Introduction	136
3.2	Dataset and methods	137
3.3	Results	153
3.4	Discussion	158
3.5	Conclusion	161

Une des questions centrales de cette thèse est de déterminer si les processus hydrologiques des bassins urbanisés sont très différents de ceux des bassins non urbanisés. Si ces différences sont importantes, nous devons alors en tenir compte dans la structure des modèles hydrologiques à développer pour les bassins versants urbanisés.

Avant d'entamer les travaux de test et de développement de modèle, nous proposons dans ce chapitre d'analyser la variabilité d'une métrique caractérisant la réponse des bassins versants : le coefficient de ruissellement. Cette métrique donne une indication sur les processus intervenant dans la génération du ruissellement. Nous avons choisi cette métrique car les facteurs impactant sa variabilité sont *a priori* sensiblement différents entre (1) les surfaces urbaines imperméables où le ruissellement est traditionnellement conçu comme dépendant uniquement du forçage pluviométrique, et (2) les surfaces naturelles où l'état d'humidité du sol intervient comme facteur important dans la génération du ruissellement. Sur un bassin urbanisé, donc composé d'une mosaïque des deux types de surface, la conjugaison des deux contrôles de l'imperméabilité et des conditions d'humidité antérieures doit être explorée pour comprendre si l'humidité joue encore un rôle déterminant même si le bassin versant est fortement urbanisé. Dans ce chapitre, nous regardons aussi l'importance des caractéristiques de l'événement pluviométrique et des caractéristiques spatiales du bassin versant (c'est-à-dire le climat, la morphologie et la géopédologie) dans l'explication de la variabilité temporelle et spatiale des coefficients de ruissellement.

Le présent chapitre correspond à un article publié dans *Water Resources Research*¹², dans lequel nous abordons les questions suivantes :

1. Quelle est l'importance relative du taux moyen d'imperméabilité totale (en tant que facteur spatial) et des conditions antérieures d'humidité du sol (en tant que facteur temporel) dans l'explication de la variabilité spatiotemporelle du coefficient de ruissellement?
2. Dans quelle mesure les conditions d'humidité antérieures jouent un rôle dans la variabilité des coefficients de ruissellement sur un large gradient d'imperméabilité totale?
3. À l'échelle du bassin versant, existe-t-il un seuil d'imperméabilité à partir duquel l'humidité du sol ne joue plus un rôle dans la détermination du coefficient de ruissellement?

Pour répondre à ces questions, nous nous servons d'un échantillon de 70227 événements extraits à partir des séries pluie-débit de 852 bassins versants nord-américains. Cet échantillon couvre une gamme du taux moyen d'imperméabilité totale allant de 0% à 59% de la surface du bassin versant avec diverses conditions d'humidité du sol précédant les événements de pluie.

¹²Sous le titre : "Beyond imperviousness: the role of antecedent wetness in runoff generation in urbanized catchments", doi : [10.1029/2020WR028060](https://doi.org/10.1029/2020WR028060).

Abstract

Explaining the response of catchments with a mix of pervious and impervious surfaces requires the inclusion of controls other than catchment imperviousness. In this study, we made use of a large sample of 70,227 events in 852 United States catchments to analyze the interplay of total impervious area (*TIA*) and antecedent soil moisture (*ASM*) in determining the event runoff ratio (*RR*). First, we investigated the importance of *TIA* as a spatial control compared to other catchment attributes (describing climate and catchment landscape) in explaining the differences between catchments in terms of median event-scale *RR*. Second, we examined the importance of *ASM* as a temporal control of the variability of *RR* in comparison with spatial settings and other event characteristics. Third, events were grouped into 30 classes of *TIA* and 30 classes of *ASM*, and linear regression models were used to analyze the evolution of *RR* depending on each one of the two independent variables. Results showed the following: **(1)** Analysis of the importance of space-varying catchment descriptors highlighted the relevance of *TIA* as a control of catchment response, whereas *ASM* showed a stronger ability compared to other event characteristics. **(2)** *ASM* played a significant role in urbanized catchments as in rural ones. The absence of abrupt changes in the linear model parameters indicates that the impact of urbanization on the relationship between *RR* and *ASM* was relatively gradual. **(3)** Differences in *RR* between rural and intensively urbanized catchments were observable not only for dry pre-event conditions, but also for wet short-term conditions.

Plain language summary

Urbanization causes a substantial increase in surface sealing which leads to dramatic impacts on water cycle. Reliable projections of the impact of urbanization requires a good understanding of the multiple factors that shape the response of a catchment along a wide range of urbanization levels. In catchments with a mix of urban and natural areas, the interplay of total impervious area (*TIA*) and antecedent soil moisture (*ASM*) conditions is still not well elucidated. To fill this gap, we attempted to quantify the relative importance of *TIA* and *ASM* in determining the event-scale catchment response using a sample of 852 United States catchments. We found that both controls play an important role in runoff generation. Particularly, *ASM* should be considered even in catchments with a high level of urban cover. We also found that the impact of total impervious area on runoff ratio was relatively gradual, and became significant when *TIA* exceeded 5%. This study provides a solid empirical understanding of the behavior of urbanized catchments, which is a step toward more reliable prediction of the impact of urbanization on hydrology for urban planning policies.

3.1 Introduction

Runoff is an aggregated result of the interaction of numerous hydrological processes within a catchment. Regarded as the catchment response to climatic forcing, it is a complex function of interrelated factors such as climate, morphology, geopedology, and land use. Hierarchizing these factors is necessary in order to understand and predict the behavior of catchments under change. In particular, urbanization and the associated increase in surface imperviousness constitute perhaps one of the most dramatic human-induced changes with regard to landscapes and hydrological processes (Vörösmarty *et al.*, 2010). Yet, quantifying the impact of urbanization on catchment response is still not straightforward due to balancing factors originating from the non-urbanized part of the catchment as well as the dependency on catchment properties.

The event runoff ratio (RR) is a widely used hydrological metric to characterize the catchment response (Blume, Zehe, and Bronstert, 2007). It expresses the fraction of rainfall that is converted into runoff. Numerous factors drive the variability of RR , some of which are variable in time (such as event rainfall depth and intensity, and antecedent soil moisture conditions; see, e.g., Graeff *et al.*, 2012; Hewlett, Fortson, and Cunningham, 1977; Tarasova *et al.*, 2018a), while others reflect the catchment hydrological properties, and thus vary in space from one catchment to another (e.g., climate setting, land use, slope, and soil characteristics; see Addor *et al.*, 2018; Merz and Blöschl, 2009; Tarasova *et al.*, 2018b). In the case of urbanized catchments, RR generally increases with the total impervious area (TIA) of the catchment (Miller and Hess, 2017; Zhou *et al.*, 2017), as a consequence of reduced infiltration due to surface sealing. Although urban-induced imperviousness (measured by TIA) is widely considered as a major control by many predictive regression relationships in urbanized areas (e.g., Survey, 1994), the effect of urbanization on catchment behavior is dependent on the local physical characteristics (Price, 2011). Hierarchizing all involved factors (i.e., soil, climate, topography, and land use) is thus needed to elucidate the relative importance of each one in influencing the hydrological response, particularly that of TIA .

Antecedent soil moisture (ASM) modulates the temporal variability of event-scale RR , and thus it is a widely considered control to explain and predict the hydrological behavior of non-urbanized catchments (Berghuijs *et al.*, 2016). Previous studies attempted to represent the interplay of TIA and ASM in the variability of the RR of urbanized catchments (Ando, Musiaké, and Takahasi, 1984; Brun and Band, 2000; Zhang and Shuster, 2014), while the relevance of ASM as a control of RR in highly urbanized catchments is still debated. Smith *et al.* (2013) found that the 5-day antecedent rainfall, used as a proxy for ASM , did not play a significant role in the generation of RR for seven urbanized catchments in the Baltimore region (United States). Zhou *et al.* (2017) analyzed the importance of ASM for a sample of 16 catchments located in Charlotte, North Carolina, with TIA ranging between 0.08 and 0.50. They found that

the influence of *ASM* on *RR* decreased with increasing *TIA*. Miller and Hess (2017) analyzed the role of *ASM* for eight catchments located in the United Kingdom, and found that *RR* was less dependent on *ASM* in urbanized catchments than in rural ones.

These studies suggest that *ASM* plays a less important role in the runoff generation as the catchment imperviousness increases. However, they focused on a small number of catchments, with a range of imperviousness that was not rich enough to reach a general conclusion, nor to quantify the explanatory power of *ASM* in comparison with *TIA*.

To address these gaps, we propose to analyze the relationship between the catchment response, materialized by *RR*, and *ASM* across a wide range of *TIA*. The novelty of this study is the use of a large sample of 70,227 rainfall-runoff events from 852 urbanized and rural catchments located in the United States (US), which offers a large diversity of urbanization levels resulting in a wide range of *TIA*. Another key novelty consists in considering both the space-varying factors (defined as catchment attributes, which vary in space from one catchment to another) and the time-varying factors (defined as factors that vary from one event to another, i.e., event rainfall characteristics and *ASM*) by analyzing their relative importance in explaining the spatiotemporal variability of *RR* along a continuum of catchment imperviousness.

This paper is structured as follows. In Section 3.2, we present the catchment sample, and we describe how *RR*, *TIA*, and *ASM* were estimated. We explain how we assessed the importance of *TIA* (as a space-varying control) and *ASM* (as a time-varying control) in predicting *RR*. Then, we detail how the variability of *RR* was analyzed with regard to *ASM* and *TIA*, as well as how the sensitivity of *RR* to each variable was assessed. The results are presented and discussed in Sections 3.3 and 3.4, and Section 3.5 concludes with the main findings and some perspectives.

3.2 Dataset and methods

3.2.1 Catchment sample and spatial characteristics

We used a sample of 852 catchments located in the United States (Saadi, Oudin, and Ribstein, 2019), for which the snow effect and the influence of artificial reservoirs were limited. Their sizes ranged between 1.1 km² and 21,300 km², with a median value of 321 km² (interquartile range: 90 km²-926 km²). The majority of these catchments are located in Southeastern United States, and are characterized by a temperate humid subtropical climate (Beck *et al.*, 2018). Their aridity indices varied between 0.48 and 3.4, and a non-negligible number of catchments are located in areas with relatively high aridity indices (Figure 3.1), indicating the richness of

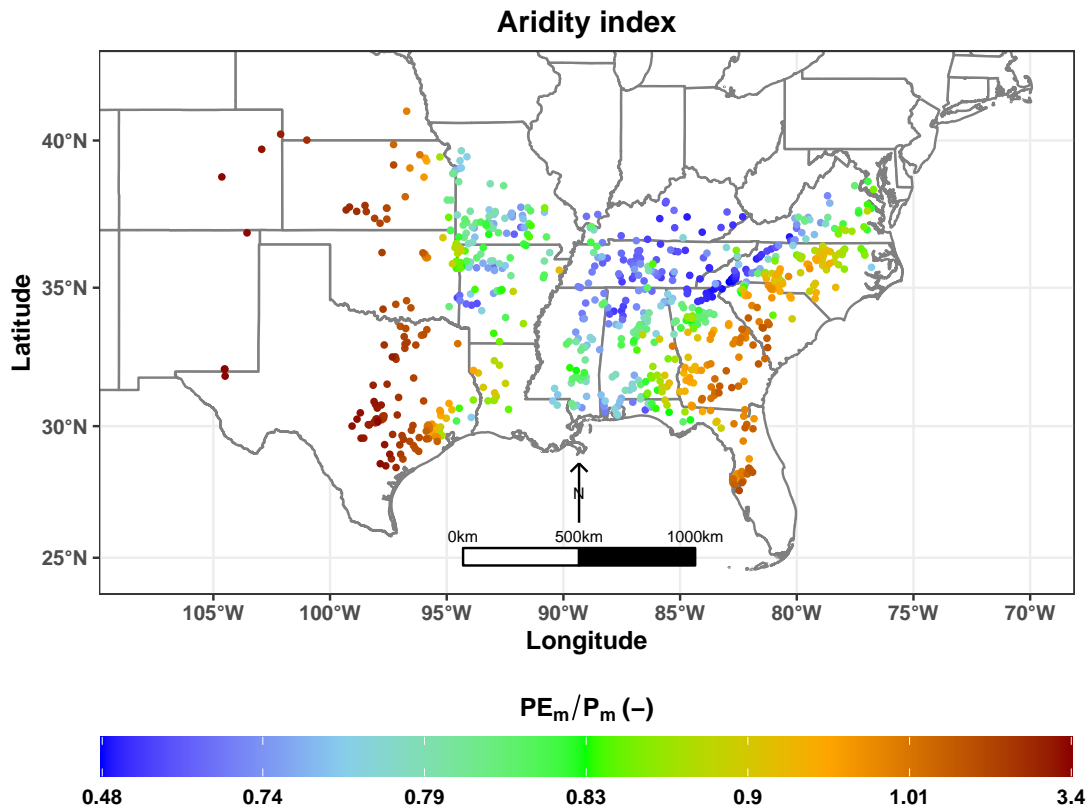


Figure 3.1. Location and aridity index of the 852 US catchments. Dots represent catchment centroids. Colors indicate aridity indices, computed as the ratio of average annual potential evapotranspiration PE_m (mm/year) to average annual precipitation P_m (mm/year).

the catchment sample with regard to climatic settings.

The hourly mean areal precipitation depths were extracted from the National Centers for Environmental Prediction’s Stage-IV dataset using the `geoknife` R package (Read *et al.*, 2015; Lin and Mitchell, 2005). This dataset was constructed by mosaicking regional multi-sensor (i.e., radar and gages) precipitation analyses into a national product, gridded over the United States at a 4-km resolution. Hourly mean streamflow data were prepared by processing the instantaneous streamflow measurements, which we extracted from the GAGES-II database maintained by the United States Geological Survey (Falcone, 2011) using the `dataRetrieval` R package (Cicco *et al.*, 2018). Daily temperature was extracted from the Daymet product (Thornton *et al.*, 2016), available over the United States at a 1-km resolution. Recorded periods span between 2002 and 2017, with lengths ranging from 8 to 16 years.

Landscape and climate settings of each catchment were characterized using an initial set of 23 descriptors (Table 3.1). The land cover characteristics were extracted from the Multi-Resolution Land Characteristics Consortium’s NLCD database, available for the years 2001, 2006, 2011, and 2016 (e.g., Homer *et al.*, 2015) at a 30-m resolution. These included the mean TIA (computed as the arithmetic mean of percentages of sealed surfaces estimated at each catchment pixel),

catchment percent developed (*CPD*, computed as the proportion of catchment area occupied by urbanization classes), the fraction of catchment area with an imperviousness value higher than 0.8 (*IMP80*), the fraction of forest (*fFOR*), and the fraction of open water (*fW*). The distribution of imperviousness values of the catchment was characterized using the irregularity of imperviousness (*IrIMP*). A linear interpolation was used to estimate these characteristics for missing years.

Climate conditions were described using six metrics, including long-term mean precipitation (P_m), flashiness of precipitation (*FP*), hourly precipitation intensity coefficient (P_{99}/P_m), long-term mean potential evapotranspiration (PE_m), flashiness of potential evapotranspiration (*FPE*), and aridity index (*AI*). In addition to the catchment area (*Area*), the drainage density (*DD*), the catchment mean slope (*Slope*), the mean elevation ($Elev_m$), and the median value of compound topographic index (*CTI*) of catchment pixels were chosen to characterize the morphological settings. The mean content of sand (*Sand*), gravel (*Gravel*), silt (*Silt*), and clay (*Clay*), the mean soil porosity (*Porosity*) and the mean intrinsic permeability (*Perm*) were used as geopedological characteristics. Table 3.1 summarizes the catchment characteristics, their estimation methods, and their corresponding sources.

Table 3.1.1. Description of catchment characteristics, their estimation methods, and their sources.

Type	Variable name	Notation and unit	Computation	Data source	Selected/rejected? ^a
Land use and imperviousness	Mean total impervious area	TIA (-)	Arithmetic mean of imperviousness values (i.e., percentage of developed surface) from catchment pixels	National Land Cover Database (NLCD) 2001, 2006, 2011, and 2016 (30-m resolution)	Selected
	Catchment percent developed	CPD (-)	Number of pixels occupied by developed classes divided by the total number of pixels		Rejected, highly correlated with TIA
	Fraction of catchment area with imperviousness value higher than 0.8	$IMP80$ (-)	Number of pixels with an imperviousness value higher than 0.8 divided by the total number of pixels		Rejected, highly correlated with TIA
	Fraction of forest	$fFOR$ (-)	Number of pixels occupied by forest classes divided by the total number of pixels		Selected
	Fraction of open water	fW (-)	Number of pixels occupied by open water class divided by the total number of pixels		Selected
	Irregularity of imperviousness	$IrIMP$ (-)	$IrIMP = \frac{\sqrt{\sum_j IMP_j^2}}{\sum_j IMP_j}$, where IMP_j is the imperviousness value of the pixel j of the catchment		Selected

Continued on next page...

Table 3.1 continued: Description of catchment characteristics, their estimation methods, and their sources.

Type	Variable name	Notation and unit	Computation	Data source	Selected/rejected? ^a
Climate	Long-term mean precipitation	P_m (mm/h)	Total precipitation depth of the recorded period, divided by the number of hours	NCEP Stage-IV dataset, 4-km resolution, extracted using the <code>geoknife</code> R package (Read <i>et al.</i> , 2015)	Selected
	Flashiness of precipitation	FP (-)	$FP = \frac{\sum_h P_h - P_{h-1} }{\sum_h P_h}$, with P_h the precipitation depth (mm) at hour h (Baker <i>et al.</i> , 2004; Holko <i>et al.</i> , 2011)		
	Hourly precipitation intensity coefficient	P_{99}/P_m (-)	The ratio of the 99 th quantile of hourly precipitation P_{99} (mm/h) to long-term mean precipitation P_m (mm/h)	Estimated using a temperature-based formula (Oudin <i>et al.</i> , 2005). Daily temperature was extracted from Daymet dataset, 1-km resolution (Thornton <i>et al.</i> , 2016)	Selected
	Long-term mean potential evapotranspiration	PE_m (mm/h)	Total potential evapotranspiration depth of the recorded period, divided by the number of hours		

Continued on next page...

Table 3.1 continued: Description of catchment characteristics, their estimation methods, and their sources.

Type	Variable name	Notation and unit	Computation	Data source	Selected/rejected? ^a
	Flashiness of potential evapotranspiration	FPE (-)	$FPE = \frac{\sum_h PE_h - PE_{h-1} }{\sum_h PE_h}$, with PE_h the hourly depth of potential evapotranspiration (mm) at hour h (Baker <i>et al.</i> , 2004; Holko <i>et al.</i> , 2011)		Rejected, highly correlated with PE_m
	Aridity index	AI (-)	$AI = PE_m / P_m$	PE_m and P_m data sources	Selected
Morphology	Catchment area	$Area$ (km ²)	-	(Falcone, 2011)	Rejected, highly anticorrelated with FP
	Drainage density	DD (km/km ²)	$DD = \frac{\sum_i L_i}{Area}$, with L_i length of stream i (km) and $Area$ the catchment area (km ²)	The hydrographic networks were extracted from the National Hydrography Dataset (NHD) using the FedData R package (Bocinsky, Beaudette, and Chamberlain, 2019)	Selected
	Mean slope	$Slope$ (-)	Arithmetic mean of slope values from catchment pixels		Rejected, highly correlated with $fFOR$

Continued on next page...

Table 3.1 continued: Description of catchment characteristics, their estimation methods, and their sources.

Type	Variable name	Notation and unit	Computation	Data source	Selected/rejected? ^a
	Mean elevation	$Elev_m$ (m)	Arithmetic mean of elevation values from catchment pixels	Hydrologic Derivatives for Modeling Applications (HDMA) database (Verdin, 2017)	Rejected, highly anticorrelated with PE_m
	Median compound topographic index	CTI (-)	$CTI = \text{median} \left(\log \left(\frac{A_{s,i}}{\tan(\beta_i)} \right) \right)$, with $A_{s,i}$ the specific contributing area (m) of the catchment pixel i and $\tan(\beta_i)$ its topographic slope		
Geopedology	Mean sand content of soil and subsoil geologic units	$Sand$ (%)	Weighted arithmetic mean of sand content values of catchment pixels (30% of topsoil and 70% of subsoil)	Harmonized World Soil Database (HWSD, Version 1.2; FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012)	Selected
	Mean gravel content of soil and subsoil geologic units	$Gravel$ (%)	Weighted arithmetic mean of gravel content values of catchment pixels (30% of topsoil and 70% of subsoil)		
	Mean silt content of soil and subsoil geologic units	$Silt$ (%)	Weighted arithmetic mean of silt content values of catchment pixels (30% of topsoil and 70% of subsoil)		

Continued on next page...

Table 3.1 continued: Description of catchment characteristics, their estimation methods, and their sources.

Type	Variable name	Notation and unit	Computation	Data source	Selected/rejected? ^a
	Mean clay content of soil and subsoil geologic units	<i>Clay</i> (%)	Weighted arithmetic mean of clay content values of catchment pixels (30% of topsoil and 70% of subsoil)		Selected
	Mean porosity of the catchment soil and subsoil geologic units	<i>Porosity</i> (-)	Volume of voids divided by the total volume of soil and subsoil, estimated over sets of homogeneous polygons. For each catchment, mean porosity was estimated by area-weighted mean of values extracted from polygons in intersection with the catchment polygon	GLobal HYdrogeology MaPS (GLHYMPS; Gleeson <i>et al.</i> , 2014)	Selected
	Mean logarithmic values of intrinsic permeability	<i>Perm</i> (m ²)	As for porosity, mean intrinsic permeability was estimated by area-weighted mean of logarithmic values extracted from polygons in intersection with the catchment polygon		Selected

^a By choosing a Spearman's rank correlation coefficient threshold $|r_{s,p}| > 0.7$, variables that were highly correlated with the selected ones were rejected in order not to influence variable importance scores, as explained in Section 3.2.3.

End of Table 3.1.

3.2.2 Time-varying event and catchment characteristics

Many event-separation algorithms are reported in the literature (Mei and Anagnostou, 2015; Merz and Blöschl, 2009; Norbiato *et al.*, 2009). The event-separation method used in this paper is detailed in Appendix A, only the three main steps are mentioned hereafter. First, we proceeded by separating the baseflow from the direct flow, using a combination of the constant- k method (Blume, Zehe, and Bronstert, 2007) and a recursive digital filtering algorithm (Collischonn and Fan, 2013; Eckhardt, 2005; Mei and Anagnostou, 2015). Second, the direct flow time series, which is the difference between the total streamflow and the baseflow, was iteratively screened to look for peak flows in the vicinity of which the start and the end of each event were estimated using direct flow thresholds and a bunch of time characteristics estimated from the cross-correlation of direct flow and rainfall. Third, the event was retained after satisfying additional conditions related to peak flow importance, duration, and total rainfall depth. The event RR was computed as the ratio of direct flow to the total rainfall depth. This event-separation algorithm yielded 70,227 events, and the number of events per catchment ranged between 1 and 615, with a median of 50 events per catchment (interquartile range: 23-104).

For each event, a set of time-varying characteristics summarized in Table 3.2 were estimated. The first group of characteristics is composed of event rainfall depth P_{tot} (mm), rainfall intensity P_{int} (mm/h), and maximum hourly rainfall depth of the event P_{max} (mm/h). The second group is a set of variables derived to estimate catchment ASM . To this end, four types of proxies were tested:

- The discharge at the start of the event can be considered to reflect the state of catchment moisture (Cappus, 1960; Kirchner, 2009; Tarasova *et al.*, 2018a). Two measures were considered: the total discharge and the baseflow at the start of the event, both normalized by the mean annual streamflow of the catchment.
- The commonly applied Natural Resources Conservation Service's Curve Number (NRCS-CN) method traditionally considers the antecedent precipitation (AP) as a proxy of ASM , namely 5-day AP (Isik *et al.*, 2013; Soulis *et al.*, 2009). For each event, we have estimated a set of AP measures corresponding to different lengths of time windows before the start of the event, including 2, 5, 10, 15, and 20 days, in addition to a specific period Lag (in hours) estimated for each catchment as the delay Δt (in hours) that maximized the correlation between the precipitation time series $P(t)$ and direct flow $Q_d(t + \Delta t)$ (see Appendix A).
- The antecedent precipitations index (API) has been widely used as a proxy of ASM (Berthet *et al.*, 2009; Heggen, 2001; Kohler and Linsley, 1951). It is computed as a weighted sum of the antecedent rainfall depths according to their antecedence in

time. As with *AP*, we considered six *APIs* including 2, 5, 10, 15, and 20 days as well as *Lag* hours of antecedent precipitation. The weighting parameter was estimated corresponding to a 10% weight for the earliest considered depth in each *API* (e.g., the 120th and the 360th hour before the start of each event for 5-day *API* and 15-day *API*, respectively).

- Measures based on soil moisture accounting reservoirs (*SMAR*; Berthet *et al.*, 2009; Merz, Blöschl, and Parajka, 2006a). We used the *SMAR* of the hourly hydrological model GR4H (Ficchi, Perrin, and Andréassian, 2019), due to its ability to represent *ASM* as shown in previous studies (Anctil *et al.*, 2004; Javelle *et al.*, 2010). It enables us to make full use of the precipitation (*P*) and potential evapotranspiration (*PE*) hourly time series to get information on the current hydric state of the catchment. The equations of the *SMAR* are detailed in Appendix B. The pre-event catchment *ASM* is taken as the ratio $\frac{S}{S_{max}}$ at the start of the event, where *S* is the updated *SMAR* state (mm) and *S_{max}* (mm) the maximum capacity. This ratio varies between 0 (extremely dry condition, i.e., the *SMAR* is empty) and 1 (extremely wet condition, i.e., the *SMAR* is full). *S_{max}* regulates the *SMAR* dynamics and memory; setting *S_{max}* low allows us to target the short-term wetness conditions (i.e., a few days before the event), whereas large values of *S_{max}* make it possible to target wetness conditions on longer terms. Four values of *S_{max}* were tested in this study: 10, 50, 100, and 200 mm.

Table 3.2. Summary of the time-varying event characteristics and antecedent soil moisture (ASM) proxies.

Type	Variable name	Notation and unit	Computation	Tested configurations	Selected/rejected? ^a
Event characteristics	Total rainfall depth	P_{tot} (mm)	$P_{tot} = \sum_{h \in event} P_h$, where P_h is the precipitation depth at hour h	–	Selected
	Rainfall intensity	P_{int} (mm/h)	$P_{int} = \frac{P_{tot}}{h_{end} - h_{start}}$, where P_{tot} is the event's total rainfall depth, h_{start} and h_{end} are the hours of event start and end	–	Selected
	Maximum hourly rainfall depth	P_{max} (mm/h)	$P_{max} = \max_{h \in event} P_h$, where P_h is the precipitation depth at hour h	–	Rejected, highly correlated with P_{tot}
Antecedent soil moisture conditions (ASM)	Pre-event total flow and baseflow	Q_E (–) and Q_{bE} (–)	The total streamflow and the baseflow at the start of the event, divided by the mean annual streamflow	–	Q_E selected Q_{bE} rejected, highly correlated with Q_E

Continued on next page...

Table 3.2 continued: Summary of the time-varying event characteristics and antecedent soil moisture (*ASM*) proxies.

Type	Variable name	Notation and unit	Computation	Tested configurations	Selected/rejected? ^a
	Antecedent precipitation	<i>AP-D</i> days (mm)	$AP = \sum_{h=1}^N P_{h_{start}-h}$, with h_{start} the start of the event, P_h is the precipitation depth at hour h , and $N = 24 \cdot D$ is the period length in hours	$N \in \{48 \text{ h}, 120 \text{ h}, 240 \text{ h}, 360 \text{ h}, 480 \text{ h}, Lag\}$, corresponding to <i>AP-02</i> , <i>AP-05</i> , <i>AP-10</i> , <i>AP-15</i> , <i>AP-20</i> , and <i>AP-Lag</i>	<i>AP-Lag</i> selected <i>AP-02</i> rejected, highly correlated with <i>API-02</i> <i>AP-05</i> and <i>AP-10</i> rejected, highly correlated with <i>API-10</i> <i>AP-15</i> and <i>AP-20</i> rejected, highly correlated with <i>API-15</i> and <i>API-20</i> (which in turn are highly correlated with <i>API-10</i>)
	Antecedent precipitation index	<i>API-D</i> days (mm)	$API = \sum_{h=1}^N \alpha^h P_{h_{start}-h}$, with h_{start} the start of the event, P_h is the precipitation depth at hour h , and α chosen as $\alpha^N = 0.1$, where $N = 24 \cdot D$ is the period length in hours	$N \in \{48 \text{ h}, 120 \text{ h}, 240 \text{ h}, 360 \text{ h}, 480 \text{ h}, Lag\}$, corresponding to <i>API-02</i> , <i>API-05</i> , <i>API-10</i> , <i>API-15</i> , <i>API-20</i> , and <i>API-Lag</i>	<i>API-02</i> and <i>API-10</i> selected <i>API-05</i> , <i>API-15</i> , and <i>API-20</i> rejected, highly correlated with <i>API-10</i> <i>API-Lag</i> rejected, highly correlated with <i>AP-Lag</i>

Continued on next page...

Table 3.2 continued: Summary of the time-varying event characteristics and antecedent soil moisture (*ASM*) proxies.

Type	Variable name	Notation and unit	Computation	Tested configurations	Selected/rejected? ^a
	Soil moisture accounting reservoir	$SMAR-S_{max}$ (-)	Presented in Appendix B	$S_{max} \in \{10 \text{ mm}, 50 \text{ mm}, 100 \text{ mm}, 200 \text{ mm}\}$ corresponding to <i>SMAR-10</i> , <i>SMAR-50</i> , <i>SMAR-100</i> , and <i>SMAR-200</i>	<i>SMAR-50</i> selected <i>SMAR-10</i> , <i>SMAR-100</i> , and <i>SMAR-200</i> rejected, highly correlated with <i>SMAR-50</i>

^a By choosing a Spearman's rank correlation coefficient threshold $|r_{sp}| > 0.7$, variables that were highly correlated with the selected ones were rejected in order not to influence variable importance scores, as explained in Section 3.2.3.

End of Table 3.2.

3.2.3 Relevance of catchment imperviousness and ASM as controls of the spatial and temporal variability of RR

Before investigating the interplay of catchment imperviousness and *ASM* in determining event *RR*, we questioned their relevance in comparison with the remaining catchment and event characteristics. First, we ranked the space-varying factors (Table 3.1) based on their importance as explanatory variables of the evolution of median *RR* from one catchment to another. This will help to unveil the importance of catchment imperviousness in controlling average catchment hydrological behavior. Second, the time-varying factors (Table 3.2) were ranked based on their importance in predicting event *RR*, which helped to reveal the variables that mostly contributed to event-to-event variability of catchment response among spatial settings, event rainfall characteristics, and *ASM* proxies.

In our case, variable ranking was a difficult task as all space-varying and time-varying factors simultaneously and nonlinearly interact to decide the event-scale *RR*. To deal with the nonlinear and multidimensional aspects of the problem, we have employed the random forest algorithm (Breiman, 2001a), which has been applied by recent studies that exploited large samples of catchments characterized by a relatively large number of descriptors (Addor *et al.*, 2018; Zhang *et al.*, 2018). In addition, random forest reduces the risk of overfitting and robustly handles nonlinear relationships between predictors while allowing for variable importance assessment. These advantages are achieved thanks to a column and row subsampling strategy. A random forest is constructed by growing a number of regression trees. At each tree growing, a number of independent variables is randomly sampled at each split (generally one third of variables), while a subset of data is left out of the training sample (36.8% of cases) in order to assess the accuracy of the constructed random forest, called out-of-bag (OOB) data. OOB data is used to assess the importance of each variable by randomly shuffling the values of the OOB cases specific to each tree. This random shuffling would lead to a decrease in the accuracy of the constructed random forest, and the greater the decrease in accuracy, the more important the variable in question. The decreases in accuracy are averaged over all the trees that included the considered variable, and the obtained mean is normalized by the standard deviation of decreases to compute the variable's *Z*-score of importance. As an accuracy measure, the mean squared error (*MSE*) is utilized, and the most important variables are the ones with the highest increases in *MSE*.

Using this method, simply ranking variables does not allow for deciding which ones are significantly important. To accomplish this, we used the Boruta algorithm (Kursa and Rudnicki, 2010) implemented in the `Boruta` R package (Kursa and Rudnicki, 2020). This algorithm creates copies of all variables, called “shadow variables”, by shuffling the values of each one of the original variables. Then, both original and shadow variables are combined to run the random forest algorithm implemented in the `randomForest` R package (Liaw and Wiener,

2002) and obtain variable importance scores. The maximum Z-score of the shadow variables $MZSV$ is used to assess whether a variable is statistically relevant or not. A variable obtaining a Z-score higher than $MZSV$ scores a hit. At each iteration, the ratio of the cumulative number of hits to the total number of trials is used to assess the statistical significance of the variable importance by means of a binomial test. The outcome of this statistical test decides whether the variable is significantly relevant, significantly irrelevant, or without decision.

Boruta algorithm was used to rank both the space-varying and the time-varying variables, and decide whether a variable was significantly relevant or irrelevant in predicting RR (p -value < 0.01). As variable ranking is influenced by the multicollinearity of variables, we identified the ones that showed high Spearman's rank correlation coefficient (i.e., $|r_{sp}| > 0.7$). Among the space-varying descriptors, we selected the ones that were best correlated (in the sense of Pearson's correlation coefficient) with the 852 median event-scale RR per catchment. This has led to a final selection of 16 out of the initial 23 space-varying descriptors, as detailed in Table 3.1. Similarly, we selected only the time-varying variables that were not highly correlated (Table 3.2) by preferring the ones that had the best Pearson's correlation with the 70,227 event RR s. This has led to a rejection of 14 initially considered variables, as detailed in Table 3.2.

Once the 16 space-varying variables were ranked, the ones that were considered to be relevant by the Boruta algorithm were used to predict the median event-scale RR for each catchment. This predicted median RR , noted *Spatial RR*, was combined with the 7 selected time-varying variables (Table 3.2) to rank them based on their importance in predicting the 70,227 event RR s. This helped to compare the importance of *ASM* variables with that of event rainfall characteristics and catchment descriptors reflected by *Spatial RR*.

3.2.4 Investigation of the relationship between RR and ASM for different levels of catchment imperviousness

In this section, we present the method used to analyze the relationship between RR and two *a priori* driving characteristics: a time-varying characteristic representing catchment wetness (*ASM*) and a space-varying characteristic representing the catchment imperviousness (*TIA*). The 70,227 rainfall-runoff events are grouped into $N_{TIA} \cdot N_{ASM}$ classes, where N_{TIA} is the number of classes of *TIA* and N_{ASM} the number of classes of the *ASM* variable (Table 3.2). For each independent variable (*TIA* or *ASM*), the classes were delimited by variable quantiles ($N_{TIA} + 1$ quantiles for *TIA*, and $N_{ASM} + 1$ quantiles for *ASM*) extracted from the event dataset. This pre-clustering aims at: (i) facilitating the visualization of the RR evolution across *TIA* and *ASM* classes, and (ii) deriving robust regression models by filtering RR variability among a given class of *TIA* and *ASM*.

First, the matrix

$$MRR = \begin{pmatrix} MRR_{1,1} & \dots & MRR_{1,N_{ASM}} \\ \dots & MRR_{i,j} & \dots \\ MRR_{N_{TIA},1} & \dots & MRR_{N_{TIA},N_{ASM}} \end{pmatrix}_{\substack{1 \leq i \leq N_{TIA} \\ 1 \leq j \leq N_{ASM}}} \quad (3.2.1)$$

is constructed, where $MRR_{i,j}$ represents the mean of runoff ratios RR for the class $\{i, j\}$ of events for which TIA values are located between the i^{th} and $(i + 1)^{th}$ quantile of TIA (the $(N_{TIA} + 1)^{th}$ quantile being the max of TIA of all events) and ASM values are located between the j^{th} and $(j + 1)^{th}$ quantile of ASM (the $(N_{ASM} + 1)^{th}$ quantile being the max of ASM of all events), with $1 \leq i \leq N_{TIA}$ and $1 \leq j \leq N_{ASM}$. This matrix is constructed to qualitatively understand how RR varies along the TIA axis (i.e., urbanization-induced contrasts) and ASM axis (i.e., wetness state-induced contrasts). In particular, $MRR_{1,1}$ and $MRR_{N_{TIA},1}$ represent the mean RR under dry conditions for the least and highest urbanized cases, respectively, whereas $MRR_{1,N_{ASM}}$ and $MRR_{N_{TIA},N_{ASM}}$ represent the mean RR under wet conditions for the least and highest urbanized cases, respectively.

Second, in order to quantify the RR gradient along TIA and ASM axes, a weighted least squares (WLS) regression model was fitted to each row (i.e., RR as a linear function of ASM for a given class of urbanization) and each column (i.e., RR as a linear function of TIA for a given class of antecedent wetness conditions). The goodness of fit is estimated using the coefficient of determination R^2 . In particular, the explanatory power of ASM across all the classes of TIA $R^2(RR, ASM)$ is computed as:

$$\begin{cases} R_i^2(RR, ASM) = 1 - \frac{\sum_{j=1}^{N_{ASM}} w_{i,j} \cdot (MRR_{i,j} - \widehat{MRR}_{i,j}(ASM))^2}{\sum_{j=1}^{N_{ASM}} w_{i,j} \cdot (MRR_{i,j} - \overline{MRR}_i)^2} & ; 1 \leq i \leq N_{TIA} \\ R^2(RR, ASM) = \text{median}_{1 \leq i \leq N_{TIA}} R_i^2(RR, ASM) \end{cases} \quad (3.2.2)$$

where $MRR_{i,j}$ is the element $\{i, j\}$ of matrix MRR (Equation 3.2.1), $\{\widehat{MRR}_{i,j}(ASM)\}_{1 \leq j \leq N_{ASM}}$ are estimations from the regression line $RR = f(ASM)$ for the class i of TIA , and \overline{MRR}_i is the weighted mean of $\{MRR_{i,j}\}_{1 \leq j \leq N_{ASM}}$ for the class i . The weights $w_{i,j}$ were chosen as the inverse of the variance of RR of each class $\{i, j\}$ to account for the intraclass variability of RR . Similarly, the explanatory power of TIA across the ASM classes is estimated as:

$$\begin{cases} R_j^2(RR, TIA) = 1 - \frac{\sum_{i=1}^{N_{TIA}} w_{i,j} \cdot (MRR_{i,j} - \widehat{MRR}_{i,j}(TIA))^2}{\sum_{i=1}^{N_{TIA}} w_{i,j} \cdot (MRR_{i,j} - \overline{MRR}_j)^2} & ; 1 \leq j \leq N_{ASM} \\ R^2(RR, TIA) = \text{median}_{1 \leq j \leq N_{ASM}} R_j^2(RR, TIA) \end{cases} \quad (3.2.3)$$

where $\{\widehat{MRR}_{i,j}(TIA)\}_{1 \leq i \leq N_{TIA}}$ are estimations from the regression line $RR = f(TIA)$ for the class j of ASM , and \overline{MRR}_j is the weighted mean of $\{MRR_{i,j}\}_{1 \leq i \leq N_{TIA}}$ for the class j .

Then, we analyzed the evolution of the parameters (i.e., slope and intercept) of the regression line $RR = f(ASM)$ along the TIA classes, as well as the evolution of the parameters of the regression line $RR = f(TIA)$ along the ASM classes.

3.3 Results

3.3.1 Assessing the relevance of catchment imperviousness and ASM as explanatory variables of event-scale RR

The catchment land-use characteristics, namely TIA and fraction of forest ($fFOR$), were the most important variables in explaining the spatial variability of median event RR , as indicated by Figure 3.2a. Aridity index (AI) obtained almost similar importance scores to those of $fFOR$, followed by mean content of gravel ($Gravel$), coefficient of hourly precipitation intensity (P_{99}/P_m), and mean intrinsic permeability ($Perm$), proving the strong controls of climatic and geopedological conditions over the catchment response. The following group of controls had equivalent importance scores, with mean content of silt ($Silt$) and flashiness of precipitation (FP) slightly ahead of mean content of sand ($Sand$), clay ($Clay$), and long-term averages of precipitation and potential evapotranspiration (P_m and PE_m). The last group of controls was constituted by the drainage density (DD) and the remaining land-use characteristics (fW and $IrIMP$). Nonetheless, all the variables were confirmed significantly important in explaining the spatial variability of RR , illustrating the multi-dependency of RR on catchment attributes.

Spatial RR, the predicted median RR for each one of the 852 catchments by random forest using the space-varying characteristics, proved to be the most relevant feature to explain the spatiotemporal variability of event RR , as shown in Figure 3.2b. This underlines the importance of the physical properties of the catchment in modulating its event-scale RR . The second most important control after *Spatial RR* was *SMAR-50*, underscoring the high importance of ASM as a primary control of RR ahead of rainfall characteristics, and the usefulness of the *SMAR* technique as a proxy for ASM in predicting RR . *Spatial RR* and *SMAR-50* were followed by the pre-event discharge (Q_E) and event characteristics (P_{int} and P_{tot}). A final group of time-varying variables was formed by the remaining ASM proxies (*API-02*, *API-10*, and *AP-Lag*), suggesting their relatively low relevance in explaining RR compared to other event characteristics. Finally, all employed variables were confirmed to be significantly important.

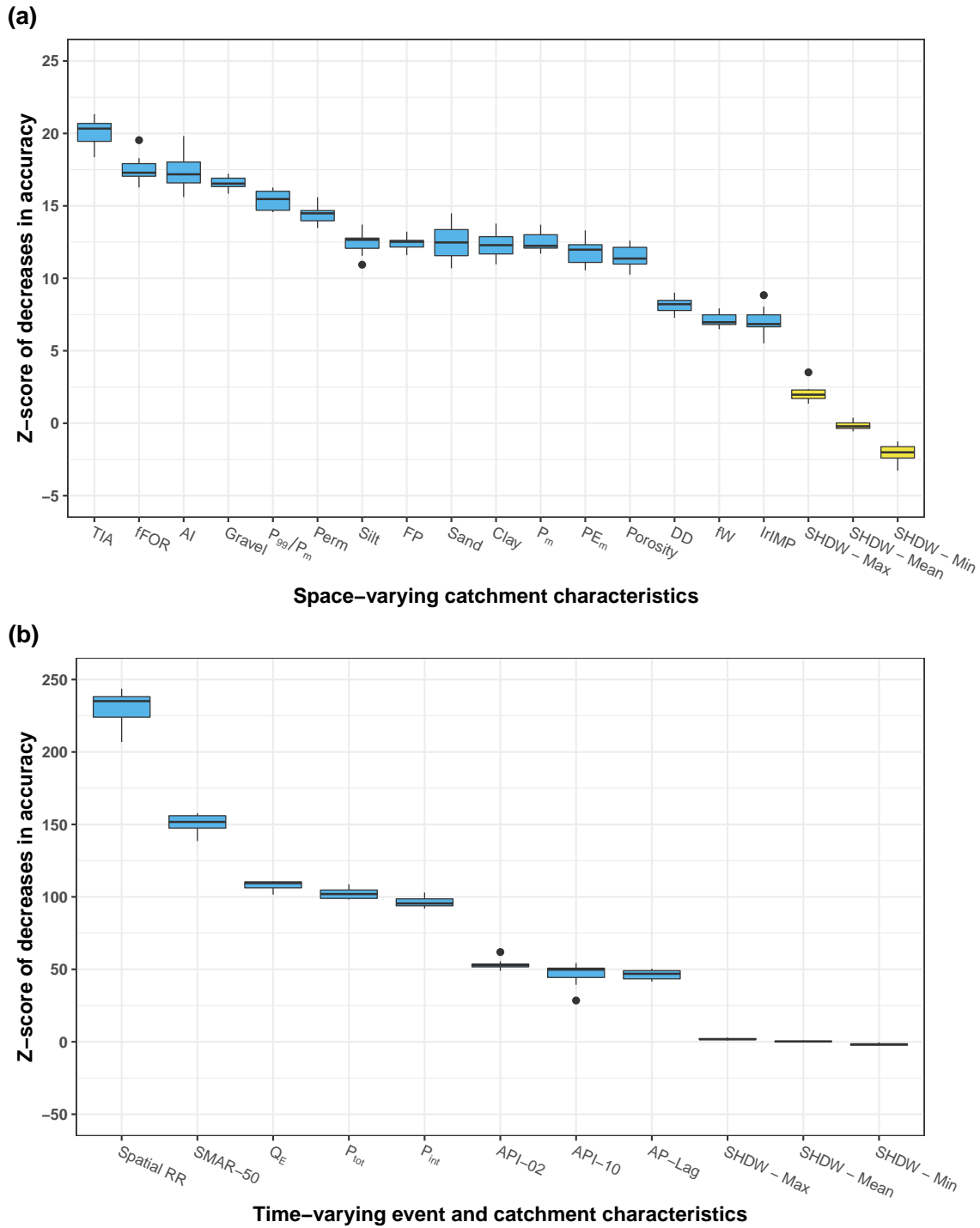


Figure 3.2. (a) Rankings of space-varying catchment characteristics (Table 3.1), based on their importance in explaining the variability of the 852 median catchment event-scale runoff ratios, and (b) rankings of *Spatial RR* and time-varying event and catchment characteristics (Table 3.2), based on their importance in explaining the variability of the 70,227 event runoff ratios. Importance is estimated using the Z-score of decreases in accuracy, i.e., of increases in mean squared errors (*MSE*) of the predictions of random forests. *Spatial RR* is the prediction of median catchment runoff ratio estimated by random forest using the space-varying catchment descriptors. *SHDW-Max*, *SHDW-Mean* and *SHDW-Min* are the distributions of maximum, mean and minimum values of Z-score of shadow variables, obtained using the Boruta R package.

3.3.2 Analyzing RR variability using TIA and SMAR as a surrogate of ASM

RR is similarly controlled by both the imperviousness level, i.e., *TIA*, and the short-term *ASM*, estimated by *SMAR-50*, as shown in Figure 3.3. In particular, the *RR* in highly urbanized catchments (i.e., high *TIA*) was sensitive to changes in *SMAR-50*. Mean *RR* per each class of *TIA* and *SMAR-50* varied between 0.01 and 0.43, the lowest values being attained for the rural cases during the dry period (bottom-left area in Figure 3.3), whereas the highest values were attained for the most urbanized cases during the wet period (top-right area in Figure 3.3). Moving from dry to wet conditions, mean *RR* increased for all catchments regardless of their level of *TIA*. Similarly, higher values of *TIA* always resulted in higher values of mean *RR* for all *ASM* conditions. Similar patterns were obtained using pre-event discharge measures, *AP* and *API* metrics instead of *SMAR-50* (not shown here).

The obtained WLS linear regression models were satisfactory for both *SMAR-50* and *TIA*, with median $R^2(RR, SMAR - 50) = 0.90$ over the 30 classes of *TIA* (range: 0.78-0.95) and $R^2(RR, TIA) = 0.76$ over the 30 classes of *SMAR-50* (range: 0.30-0.92). The goodness of fit is illustrated by examples in Figure 3.4 for the least and most urbanized cases (bottom and top rows of Figure 3.3) and the lowest and highest *ASM* conditions (left and right columns of Figure 3.3). The first two examples (Figures 3.4a and 3.4b) illustrate the effect of *SMAR-50* on mean event *RR*, where an increase in *SMAR-50* was accompanied by increasing mean *RR* similarly for both the least urbanized cases and the most urbanized cases. The last two examples (Figures 3.4c and 3.4d) indicate the effect of *TIA* on mean event *RR*, where *RR* got increased with increased *TIA*. The effect of *TIA* is visually more pronounced amid dry conditions (i.e., very low *SMAR-50*) than in wet conditions (i.e., very high *SMAR-50*), but in both cases the differences in *TIA* were accompanied by differences in catchment response.

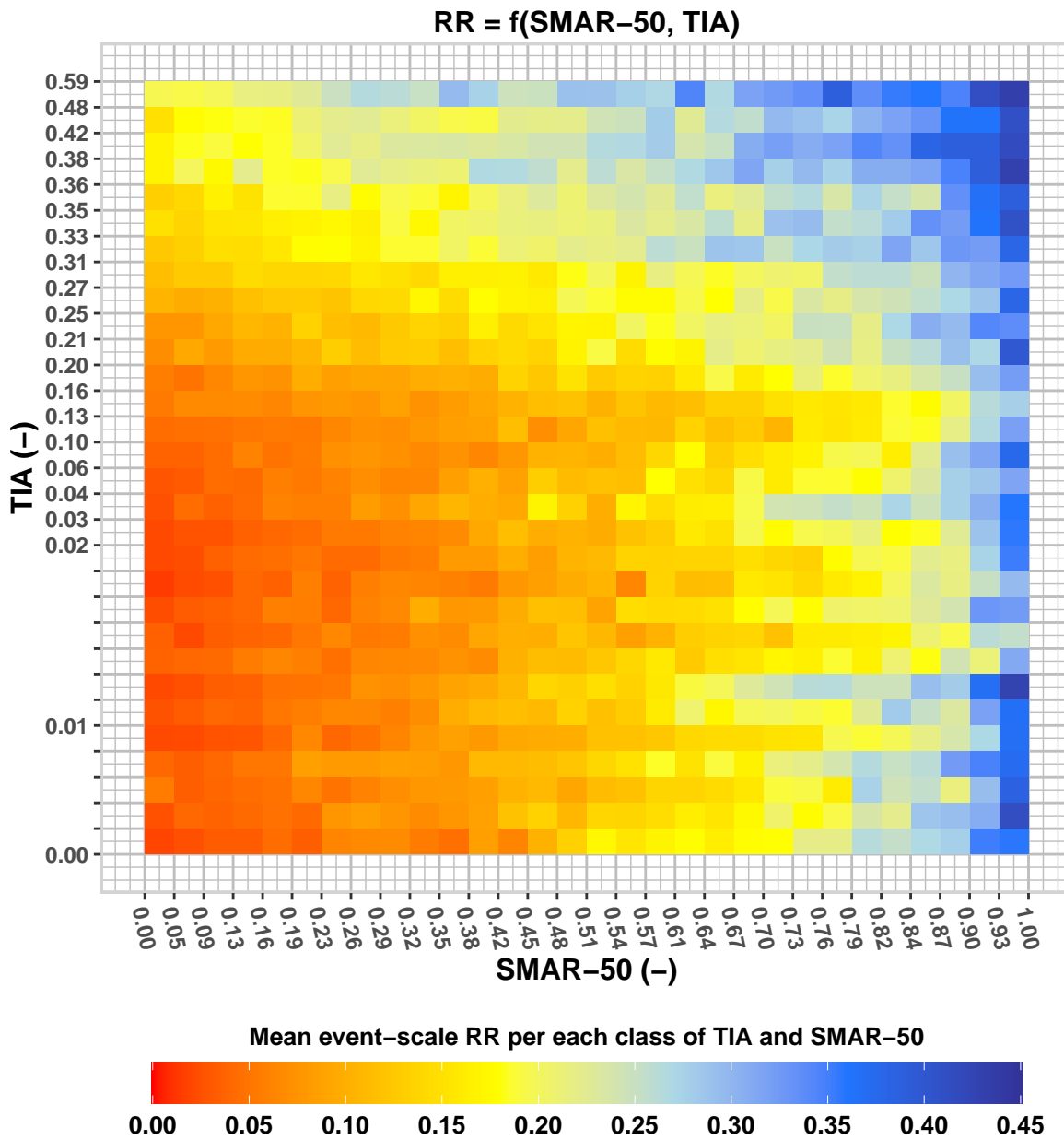


Figure 3.3. Runoff ratio (RR) dependence on antecedent soil moisture (ASM) conditions, assessed using the level of the soil moisture accounting reservoir ($SMAR-50$, Appendix B), and the catchment total impervious area (TIA). This matrix was constructed by setting $N_{TIA} = N_{SMAR-50} = 30$, yielding 900 classes of $TIA-SMAR-50$. The bottom row shows the evolution of mean RR for rural cases (i.e., TIA close to 0), and the top row shows mean RR for intensively urbanized ones (i.e., TIA between 0.48 and 0.59). The left column shows mean RR for dry conditions (i.e., $SMAR-50$ close to 0), and the right column shows mean RR for wet conditions (i.e., $SMAR-50$ close to 1).

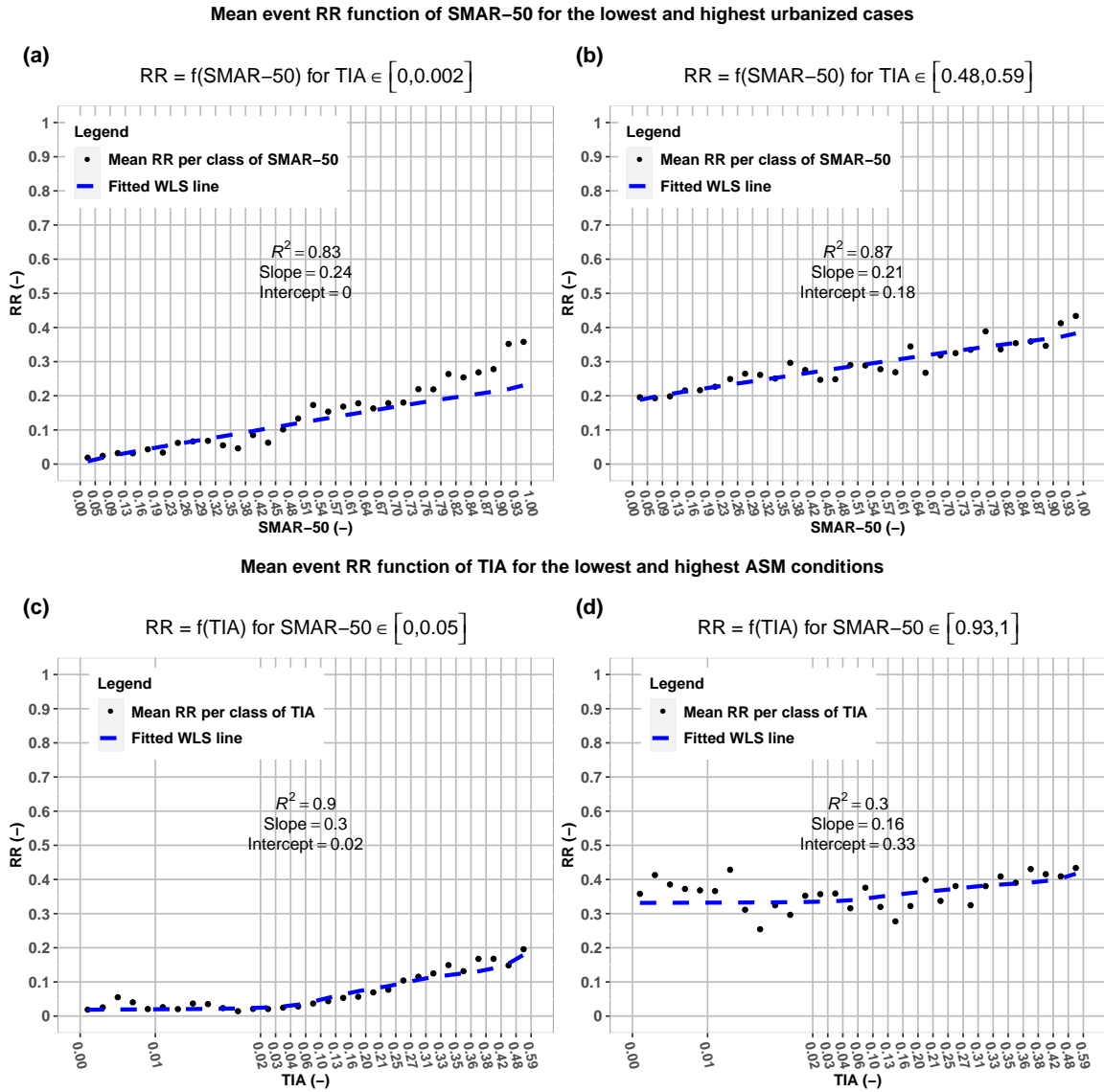


Figure 3.4. Scatter plots showing the variability of mean runoff ratio (RR) as a function of total impervious area (TIA) and antecedent soil moisture (ASM), assessed using the level of the soil moisture accounting reservoir ($SMAR-50$, Appendix B). The first row shows $RR = f(SMAR - 50)$ for (a) the lowest TIA level and (b) the highest TIA level. The second row shows $RR = f(TIA)$ for (c) the driest ASM conditions and (d) the wettest ASM conditions. Black dots show mean RR per class of TIA and $SMAR-50$, and blue dashed lines show the fitted weighted least squares (WLS) regression models.

An overview of the evolution of the WLS model parameters suggests that the control of $SMAR-50$ on RR was not influenced by TIA , as manifested by the behavior of the slope of the WLS regression line $RR = f(SMAR - 50)$ in Figure 3.5a. In general, the slope did not deviate significantly from 0.21 for all classes of TIA . This means that even for the most urbanized catchments, short-term ASM still played a significant role in determining RR as for the rural ones. The differences between rural and urbanized catchments were most reflected in the intercept parameter, which could be interpreted as the value of RR in very dry conditions (i.e., $SMAR-50$ close to 0). The intercept was around zero for catchments with TIA less than 5%, and increased somewhat linearly to reach a value of 0.18 for high levels of TIA . This suggests that in dry conditions, RR is solely controlled by the impervious part of the catchment. Varying intercept with almost constant slope is equivalent to shifting upwardly the line $RR = f(SMAR - 50)$ as TIA gets higher, as illustrated in Figure 3.5b.

3.4 Discussion

3.4.1 Importance of ASM as a control of RR in urbanized catchments

This study analyzed the role of ASM , in addition to TIA , as controls of RR . The limited number of studies that have addressed the role of ASM in urbanized catchments suggested that ASM had a less important control on RR compared with TIA (Miller and Hess, 2017; Smith *et al.*, 2013; Zhou *et al.*, 2017). Our results showed that ASM is in fact an important control to consider for predicting the event RR even in catchments with high levels of TIA . As reported in some studies (Sillanpää and Koivusalo, 2015; Zhou *et al.*, 2017), the differences between rural and urbanized catchments were essentially during the dry period. We showed that this was also the case for very wet short-term ASM (Figures 3.3, 3.4, and 3.5). Although some of the findings are intuitive, they are purely data-based, and their importance lies most in providing means of revising methods in the practice of hydrology that do not account for the effect of ASM in urbanized areas. In addition, the findings confirm previously obtained results concerning the interplay of ASM and imperviousness, namely the logistic function model proposed by Brun and Band (2000), or the scenario-based conclusions of Zhang and Shuster (2014) and Yang *et al.* (2011). Methodologically, this study is a stepping stone to overcome the issue of the “rural-urban dichotomy” (McGrane, 2016), by analyzing the evolution of catchment behavior along a continuum of TIA , which would not have been possible without the use of a large sample of catchments.

Using $SMAR$ as a proxy for ASM (Appendix B) yielded better predictive power than event characteristics (i.e., P_{tot} and P_{int}) or other ASM proxies that were based on pre-event discharge and antecedent rainfall. The relevance of $SMAR$ as a proxy of ASM has been confirmed in

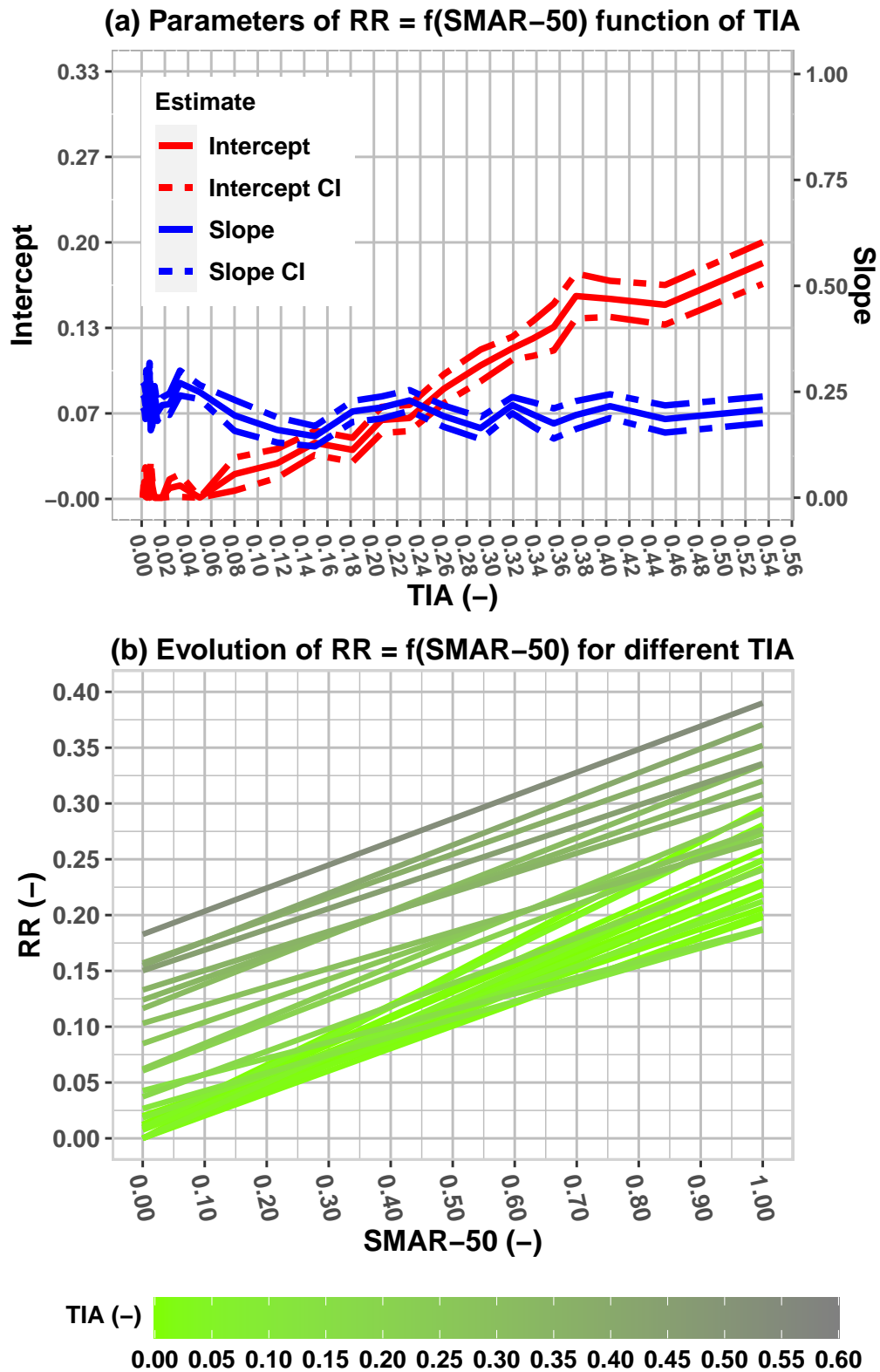


Figure 3.5. (a) Evolution of the parameters of the WLS regression line representing mean runoff ratio (RR) as a function of $SMAR-50$ for different levels of total impervious area (TIA). Estimates are shown with solid lines, and 95% confidence intervals (CI) are shown with two-dashed lines. Parameter estimates that were not statistically different from zero (p -value threshold at 0.05) were set to zero. (b) Evolution of the corresponding lines for different levels of TIA .

other studies (Anctil *et al.*, 2004; Merz and Blöschl, 2009). Moreover, the *SMAR* allows us to follow the seasonal variability of *P* and *PE*, in contrast to *API* that does not distinguish between hot and cold seasons, and the baseflow that is dependent on local configurations other than climate (e.g., soil type, connectivity of surface water and groundwater). The ranking of *ASM* proxies in this study is very helpful for event-based techniques applied to estimate runoff in urbanized areas. In particular, the NRCS-CN method would be enhanced by using *SMAR-50* or *API* measures in estimating *ASM* instead of 5-day *AP*.

The large range of *TIA* values tested in this study showed that the impact of increasing *TIA* on the relationship between *RR* and *ASM* was more gradual than threshold-like. As illustrated by the evolution of the parameters of the regression line representing *RR* as a function of *SMAR-50* (Figure 3.5a), for different values of *TIA*, the intercept started to deviate from zero as *TIA* exceeded 0.05, but no abrupt change was noticed. This advocates the continuousness of the effect of urbanization (see the discussion in Brabec, Schulte, and Richards, 2002), and implies that it would be difficult to detect a significant change in catchment behavior (in response to urbanization) if the absolute change in *TIA* or the difference between analyzed situations (as in a paired-catchment framework) is less than 10%-5%. The threshold of *TIA* above which the intercept became significantly different from zero (i.e., 0.05) has been reported in literature reviews (Alley and Veenhuis, 1983; Brabec, Schulte, and Richards, 2002; Salvatore, Bronders, and Batelaan, 2015).

Some would argue that *TIA* did not exceed 0.6 for the sample used here, leaving a relatively large range of *TIA* uncovered. When examining the proportion of the catchment that is occupied by urbanization classes, i.e., *CPD* (Falcone, 2011; Miller and Hess, 2017), this proportion varied between 0% and 100%, with a median value of 29%, meaning that the range of *TIA* used here covered a large variety of situations of catchment-scale urban sprawl, including catchments completely occupied by urbanization classes. Not reaching a *TIA* higher than 0.6 is perhaps a result of focusing on the catchment scale, as the minimum catchment size in this study is above 1 km², which implies the existence of more natural surfaces (i.e., with low imperviousness values) even for intensively urbanized catchments.

3.4.2 Relevance of other catchment and event characteristics in explaining the variability of *RR*

In addition to *TIA*, *fFOR* had the highest importance in explaining the variability of median event *RR* for the catchment set, highlighting the primary control of land-use characteristics on catchment behavior (Merz and Blöschl, 2009; Tarasova *et al.*, 2018b). Climate characteristics (the aridity index and the coefficient of hourly precipitation intensity P_{99}/P_m) and geopedological settings were the next most relevant spatial controls, underscoring the interplay of climate

and landscape features in defining the catchment long-term hydrological behavior (Gao *et al.*, 2018; Merz and Blöschl, 2009). Comparing the importance of spatial and temporal controls on the variability of event-scale RR illustrated the primacy of regional settings over ASM and event rainfall characteristics (Figure 3.2b), as the former defines the long-term behavior of the catchment while the latter shapes the short-term high-frequency variability (Merz and Blöschl, 2009). Event rainfall characteristics have been regularly chosen as explanatory variables (Blume, Zehe, and Bronstert, 2007; Hewlett, Fortson, and Cunningham, 1977; Tarasova *et al.*, 2018a), which has been confirmed by our study results (Figure 3.2b), but their importance was less than that of regional settings and ASM conditions.

Lastly, it has been shown that not only TIA influences catchment response, but also the spatial arrangement of impervious surfaces (Debbage and Shepherd, 2018; Mejía and Moglen, 2010), which has motivated to search for a more hydrologically relevant imperviousness measure, such as the effective impervious area (EIA) or the landscape fragmentation measures (Alley and Veenhuis, 1983; Ebrahimian, Wilson, and Gulliver, 2016; Oudin *et al.*, 2018). Moreover, not all urbanized cover is impervious to water (Berthier, Andrieu, and Creutin, 2004; Ragab *et al.*, 2003b; Ramier, Berthier, and Andrieu, 2011; Redfern *et al.*, 2016), and different types of urban cover are characterized by a wide range of imperviousness. In this respect, we used a second-order measure of imperviousness, i.e., $IrIMP$, which was a relevant indicator but showed a lesser importance compared to the remaining spatial controls in explaining the spatial variability of event-scale RR (Figure 3.2a).

3.5 Conclusion

In this study, we demonstrated the impact of space-varying and time-varying controls on event RR for a large sample of 852 catchments located in the US, characterized by a wide range of TIA . The main finding was that ASM still plays a significant role in shaping RR even for the most intensively urbanized catchments.

As a spatial control, TIA showed the best ability in explaining the spatial variability of catchment response, ahead of the fraction of forest and the remaining climatic, geopedological, and morphological settings. ASM was estimated by the soil moisture accounting reservoir ($SMAR$) of the GR4H model, as it was best able to explain the event-to-event variability of RR compared with event characteristics (P_{tot} and P_{int}), antecedent precipitation (AP and API), and pre-event discharge. Still, regional settings had the highest importance in explaining the variability of RR compared to ASM and event rainfall characteristics.

Using the $SMAR$ as a proxy for ASM , the events were grouped into 900 classes of TIA and ASM . Then, the dependency of RR on TIA and ASM was then quantified by means of linear

regression. Mean event-scale RR per each class was found to increase with either increasing TIA or increasing ASM . In the highly urbanized catchments, RR increased as pre-event conditions became wet. Differences in RR were found between rural and intensively urbanized catchments in dry as well as in wet ASM conditions prior to the event. The impact of TIA was less important on the slope of $RR = f(ASM)$, whereas the intercept became significantly different from zero at $TIA > 0.05$.

The fact that the TIA was not uniformly distributed led us to adopt a quantile-based clustering procedure, which would have been less harmful given a wider TIA range. Also, more diverse climatic conditions are missing in this study (including hotter and colder climates; see, e.g., Sillanpää and Koivusalo, 2015), which should be included in order to get an overall understanding of the interplay between climate and urbanization. The analysis was restrained here to mean RR values for each class, and an analysis of intraclass variability would help to synthesize the effect of TIA and ASM on the overall distribution of RR . For process-based modeling, the findings of this study are highly relevant to better represent/reproduce the rainfall-runoff relationship especially for catchments with evolving urbanization. In addition, they offer means to check the physical soundness of commonly applied regression relationships in practical hydrology in order to enhance the reliability of their predictions.

Synthèse du chapitre

Ce chapitre a comme but d'analyser les contrôles de la variabilité spatiotemporelle des coefficients de ruissellement le long d'un gradient riche de situations d'imperméabilité, grâce à un large échantillon de 852 bassins versants nord-américains. Comme résultat principal, l'humidité du sol joue un rôle important (en plus de l'imperméabilité) dans la réponse hydrologique même à des niveaux forts d'urbanisation. À partir d'un seuil du taux moyen d'imperméabilité totale *TIA* de 5%, l'impact de l'urbanisation sur la réponse hydrologique est devenu significatif. Comme attendu, les caractéristiques de l'événement pluviométrique ont également montré un impact fort sur les coefficients de ruissellement, mais elles ont eu un pouvoir explicatif moins important que l'humidité du sol et les spécificités régionales du bassin versant. La hiérarchisation des contrôles spatiaux (climat, morphologie, géopédologie) a suggéré qu'il est difficile de trancher sur les contrôles les plus déterminants, soulignant le caractère multi-dépendant de la réponse hydrologique quel que soit le niveau d'urbanisation des bassins versants.

CHAPITRE 4

COMMENT LES MODÈLES CONCEPTUELS PLUIE-DÉBIT GÈRENT-ILS LES SPÉCIFICITÉS DES BASSINS VERSANTS URBANISÉS ?

Sommaire

4.1	Introduction	167
4.2	Catchments and methodology	169
4.3	Results	177
4.4	Discussion	186
4.5	Conclusion	188
4.6	Supporting information	189

L'objectif de cette thèse est de développer un modèle hydrologique capable de reproduire la relation pluie-débit à l'échelle d'un bassin versant dont l'urbanisation augmente. Nous questionnons dans le présent chapitre la pertinence de faire un tel développement en analysant les différences entre les bassins versants urbanisés et non urbanisés du point de vue de modèles non dédiés aux environnements urbanisés.

Les modèles non dédiés aux environnements urbanisés que nous proposons de tester dans le présent chapitre prennent en compte l'humidité du sol (dont on a vu l'importance dans le chapitre précédent) comme variable explicative dans la réponse d'un bassin versant, mais leurs structures ne tiennent compte d'aucune spécificité de l'urbanisation. Ces modèles sont-ils capables de reproduire le comportement hydrologique des bassins fortement urbanisés?

Le présent chapitre correspond à un article publié dans *Hydrological Processes*¹³, dans lequel nous abordons les questions suivantes :

1. Que nous apportent certains indicateurs du comportement hydrologique concernant les différences entre les bassins fortement urbanisés et les bassins faiblement ou non urbanisés?
2. Sans prise en compte explicite des spécificités de l'urbanisation, dans quelle mesure un modèle non dédié aux environnements urbanisés est-il capable de reproduire la relation pluie-débit à l'échelle d'un bassin versant urbanisé?
3. Comment les paramètres d'un modèle non dédié aux environnements urbanisés traduisent-ils l'impact de l'urbanisation sur le comportement hydrologique du bassin versant?

Pour cela, nous avons choisi d'analyser deux structures de modèles conceptuels, globaux et fonctionnant au pas de temps horaire : GR4H et IHACRES. Ces deux structures forment différemment les processus d'infiltration, d'évapotranspiration et du ruissellement, et gèrent de manière différente le routage de l'écoulement lent et de l'écoulement rapide. Nous avons testé ces deux modèles sur 175 bassins versants urbanisés nord-américains et français pour lesquels l'imperméabilité totale moyenne (*TIA*) dépassait 10%. Le test des deux modèles a été effectué en comparant les résultats à ceux obtenus sur un échantillon témoin de 175 bassins versants non urbanisés, pour lesquels *TIA* ne dépassait pas 5% et qui se situent à proximité des 175 bassins urbanisés.

¹³Sous le titre : "Crossing the rural-urban boundary in hydrological modeling: How do conceptual rainfall-runoff models handle the specificities of urbanized catchments?", doi : [10.1002/hyp.13808](https://doi.org/10.1002/hyp.13808).

Abstract

Landscape differences induced by urbanization have prompted hydrologists to define a fuzzy boundary between rural and urban-specific hydrological models. We addressed the validity of establishing this boundary, by testing two rural models on a large sample of 175 French and United States (US) urbanized catchments, and their 175 rural neighbors. The impact of urbanization on the hydrological behavior was checked using four metrics. Using a split-sample test, we have compared the performances, parameter distributions and internal fluxes of GR4H and IHACRES, two conceptual and continuous models running at the hourly time step. Both model structures are based on soil moisture accounting reservoirs (infiltration, runoff, and actual evapotranspiration) and quick flow/slow flow routing components, with no consideration of any specific feature related to urbanization. Results showed: **(1)** Except for the ratio of streamflow flashiness to precipitation flashiness, the range of hydrological signature metrics in rural catchments encompassed the specificities of urbanized ones. Overall, the urbanized catchments showed higher ratios of mean streamflow to mean precipitation (median values: 0.39 vs. 0.27) and streamflow flashiness to precipitation flashiness (0.13 vs. 0.03), besides lower baseflow index (0.42 vs. 0.62) and shorter characteristic response time (3 vs. 14 h). **(2)** The performances of GR4H revealed no significant distinction between rural and urbanized catchments in terms of Kling-Gupta Efficiency (*KGE*), whereas IHACRES better simulated urbanized catchments, especially during summer. **(3)** With respect to differences in urbanization level, the GR4H and IHACRES parameters showed different distributions. The differences in parameters were consistent with the differences in hydrological behavior, which is promising for a model-based assessment of the impact of urbanization. **(4)** The models agreed less in reproducing the internal fluxes over the urbanized catchments than over the rural ones. These results demonstrate the flexibility of conceptual models to handle the specificities of urbanized catchments.

4.1 Introduction

4.1.1 From contrasting landscapes to contrasting hydrological modeling practices

One can easily identify the differences between a “rural” and an “urbanized” landscape: green forests, season-colored crops, and vacant soil-colored lands are clearly distinguishable from areas where paved roads and gray buildings predominate. Population density is the first driver for such landscape discrepancies (Chen, Chuang, and Cheng, 2014; Martin, Kelleher, and Wagoner, 2012): As the population that occupies the defined area increases, land conversion takes

place to accommodate for the increasing population needs. From a hydrological standpoint, this land conversion causes a substantial increase in the total impervious area (*TIA*) of the catchment, which impacts the hydrological processes by specifically emphasizing high flows and altering low flows (e.g., DeWalle *et al.*, 2000; Fletcher, Andrieu, and Hamel, 2013; Salvadore, Bronders, and Batelaan, 2015; Oudin *et al.*, 2018).

These differences in hydrological behaviors between rural and urbanized areas have led hydrologists to develop specific modeling tools for urban catchments. Notwithstanding the existence of these differences, we argue that they are potentially not more important than the ones existing between two distinct rural areas or two distinct urban areas. Moreover, the conceptual tools and physically-based equations used in representing the hydrological processes present many similarities for both types of area (Salvadore, Bronders, and Batelaan, 2015). This was also pointed out by Fletcher, Andrieu, and Hamel (2013), who stated that: “Despite its particularity, urban hydrology is not that different from *natural* hydrology”. The dichotomous approach consisting in selecting rural or urban-specific hydrological models is questionable. From an operational standpoint, this dichotomy leads to complicated choices for catchments with mixed urban/rural landscapes. In this context, a continuous modeling approach would be preferable to the dichotomous approach.

Rural conceptual hydrological models have been widely tested in large-sample studies (Addor *et al.*, 2020; Gupta *et al.*, 2014) and proved to be particularly flexible in the simulation of a large panoply of hydrological behaviors through changes in their parameters. However, these studies have often discarded (or not sufficiently emphasized) the catchments that were characterized by a high proportion of urban cover (Addor *et al.*, 2017; Ficchi, Perrin, and Andréassian, 2019; Esse *et al.*, 2013), maintaining an invisible and fuzzy rural-urban boundary in hydrological practices.

4.1.2 Scope of the study

In this paper, we question the existence of such a boundary, addressing two related issues:

1. Whether urbanized and rural catchments clearly exhibit different hydrological behaviors and whether these differences are larger than the ones among rural catchments. We focus on four aspects of the hydrological behavior: **(i)** how much of precipitation is converted into streamflow; **(ii)** how important is the baseflow component; **(iii)** what is the lag time between precipitation and runoff; and **(iv)** what is the ability of the catchment to damp precipitation flashiness (Baker *et al.*, 2004).
2. Whether conceptual hydrological models are flexible enough to account for these different hydrological behaviors. These differences would impact the model performances,

the parameter distributions, or the internal fluxes of the model.

These issues are addressed here through a comparative assessment across a large sample of urbanized and rural catchments. We collected a sample of 175 urbanized catchments (with a mean *TIA* above 10% over the period 1997-2017), which is far larger than the samples used in urbanized catchments studies (see the review by Salvadore, Bronders, and Batelaan (2015)). Benefiting from a paired-catchment framework (Martin, Kelleher, and Wagener, 2012; Van Loon *et al.*, 2019; Zégre *et al.*, 2010), we compared the average behavior of the urbanized catchments with geographically close rural catchments. Hence, an additional 175 rural catchments were used, for which mean *TIA* was less than 5%.

The study is presented as follows: in Section 4.2, we describe the catchment sample and the hydrological signatures that we used to investigate the differences in terms of hydrological behavior between the urbanized catchments and their rural neighbors. We also present the two rural models that were used to analyze the differences from a rural modeling standpoint. In Section 4.3, we show the assessment of both the hydrological signatures and the models. A discussion follows in Section 4.4, while Section 4.5 concludes with the main findings.

4.2 Catchments and methodology

4.2.1 Sample description

A large sample of catchments located in the United States (US) and France were collected. Working with an international sample of urbanized catchments is important from a hydrological perspective, as the differences in urban planning schemes between the two sides of the North Atlantic Ocean could have a significant impact on the water cycle.

The sample selection was made using four criteria:

1. Limited effect of snow, as the models we applied here do not take into account snow-melting processes.
2. Limited influence of dams. Although these impact river regimens (FitzHugh and Vogel, 2011), and are hence related to a particular type of human impact, we deemed this effect to be a particular case that required adequate streamflow naturalization techniques.
3. Recently available hourly precipitation and streamflow time series and daily temperature. We chose a minimum of 8 years of data availability between 1997 and 2017

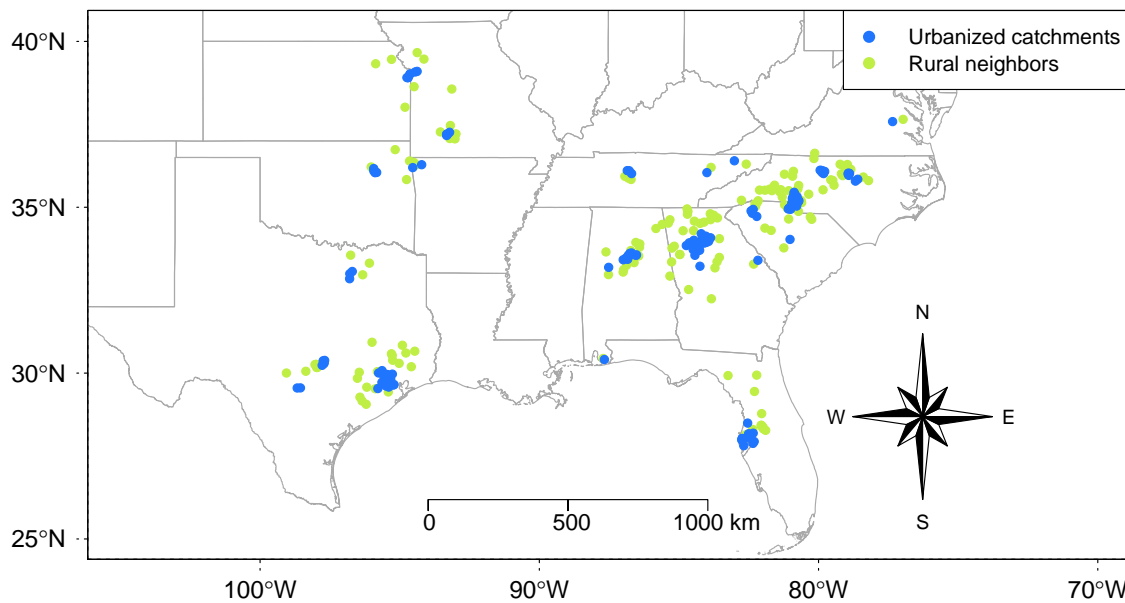


Figure 4.1. Location of the United States urbanized catchment sample (in blue) and their rural neighbors (in green). A dot represents the location of the catchment centroid.

as a requirement, for which complete precipitation and temperature data should be available, and streamflow data should be available for at least 70% of the recorded period. We considered that a minimum of 4 years for calibration was needed for an hourly model (Perrin *et al.*, 2007).

4. The question of imperviousness threshold above which a catchment can be considered to be urbanized was not specifically analyzed in this study. On the basis of previous reviews (Arnold and Gibbons, 1996; Brabec, Schulte, and Richards, 2002; Brun and Band, 2000; Salvatore, Bronders, and Batelaan, 2015), we considered a catchment as urbanized when its mean *TIA* over the whole period was above 10%. Other thresholds such as 15% and 20% were tested and did not significantly change the conclusions (not shown here).

For each urbanized catchment X , the selected rural neighbor catchment Y was the closest rural catchment to X (i.e., *TIA* of Y is less than 5%) in the sense of distance d , which was determined using the distances between the outlets (20% weight) and the centroids (80% weight) of the catchments X and Y (Lebecherel, Andréassian, and Perrin, 2013).

This selection yielded a final sample of 175 urbanized catchments, of which 156 are located in the US and the remaining 19 in France. Figure 4.1 shows the US urbanized catchments and their rural neighbors' location. Figure 4.2 shows the location of the French sample.

The US sample was selected from the GAGES-II database (Falcone, 2011), for which hourly streamflow time series were extracted from the gauges maintained by the US Geological Survey (USGS) using the `dataRetrieval` R package (Cicco *et al.*, 2018). Hourly precipitation depths

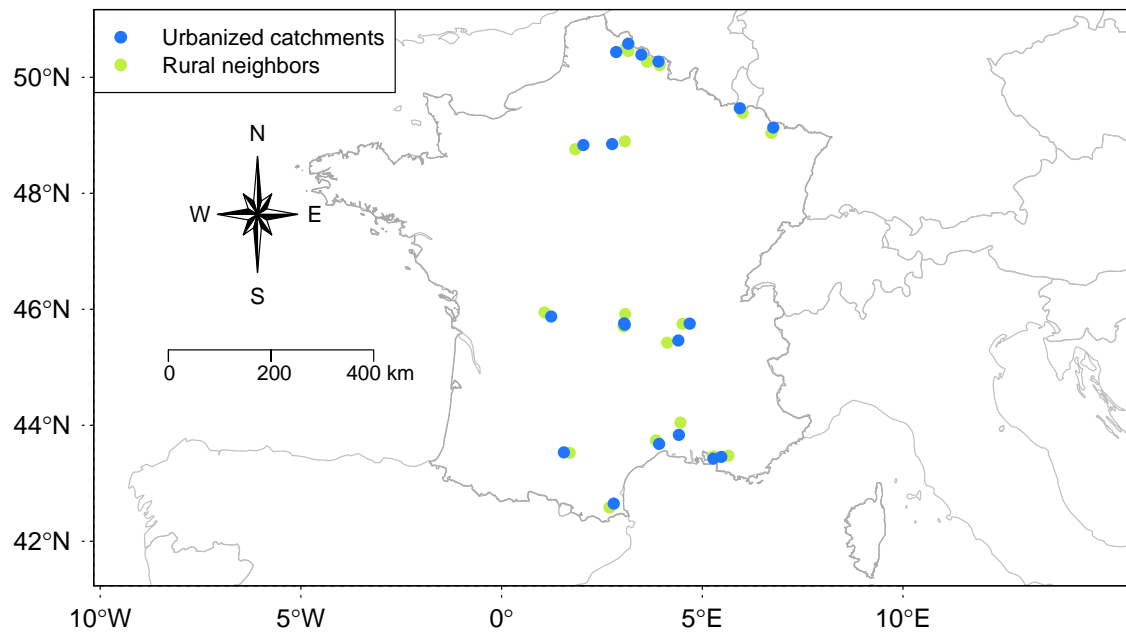


Figure 4.2. Location of the French urbanized catchment sample (in blue) and their rural neighbors (in green). A dot represents the location of the catchment centroid.

were extracted from the NCEP Stage IV product, available at a 4-km resolution using the `geoknife` R package (Lin and Mitchell, 2005; Read *et al.*, 2015). Daily temperature was extracted from the Daymet product available at a 1-km resolution (Thornton *et al.*, 2016). To characterize the degree of imperviousness of the catchment, we used the National Land Cover Database (NLCD) produced by the Multi-Resolution Land Characteristics (MRLC) Consortium for the years 2001, 2006, 2011, and 2016 at a 30-m resolution (e.g., Homer *et al.*, 2015).

The French sample was selected from the Banque HYDRO database, for which hourly streamflow time series were extracted using the `hydro.eaufrance.fr` platform (Leleu *et al.*, 2014). Hourly precipitation depths were prepared using the COMEPHORE product provided by Météo France (Tabary *et al.*, 2013), available at a 1-km resolution across France. Daily temperature was extracted from the SAFRAN product available at an 8-km resolution, provided by Météo France (Vidal *et al.*, 2010). The degree of imperviousness of the catchment was characterized using the Imperviousness Density layers produced by the Copernicus Land Monitoring Service, available for the years 2006, 2009, 2012, and 2015 at a 20-m resolution (Congedo *et al.*, 2016).

For a given catchment, the mean TIA was assessed by an arithmetic mean of the imperviousness values of catchment pixels. These values are percentages of sealed/developed surfaces estimated over every 30-m pixel in the US and 20-m pixel in France. After spatial aggregation, a linear extrapolation helped estimate the mean TIA for the years where land-use characterization was not available.

To estimate the potential evapotranspiration, we used a temperature-based formula (Oudin *et al.*, 2005) to determine daily time series, followed by disaggregation using hourly extraterrestrial radiation to estimate hourly values.

The sizes of the urbanized catchments ranged between 1.1 km² and 727 km² with a median area of 47.7 km², while their rural neighbors' sizes were between 2.4 km² and 4580 km² with a median area of 274 km². The majority of US catchments used here are characterized by a humid subtropical climate (Cfa type in the Köppen-Geiger climate classification system; see Beck *et al.*, 2018), whereas temperate oceanic and mediterranean types of climate characterize the French catchments. The mean precipitation depths for the urbanized/rural pairs of catchments were comparable (mean of 1190 mm/year in urbanized catchments vs. 1170 mm/year falling in rural catchments), and the ratios of the precipitation values of each pair (urbanized to rural) ranged between 0.75 and 1.46 (median ratio: 1.02). The mean potential evapotranspiration differences were similar (1060 mm/year and 1040 mm/year for urbanized and rural catchments, respectively), with ratios of each pair ranging between 0.94 and 1.2 (median ratio: 1.02). The *TIA* values in urbanized catchments ranged between 10% and 59% with a median value of 24.6%, which is an indication that different urbanization stages could be encountered in this sample. The *TIA* range of their rural neighbors was 0%-5% with a median value of 1.2%.

4.2.2 Hydrological signatures used to assess the differences between rural and urbanized catchments

Several hydrological signatures were analyzed to assess the differences in hydrological behavior between rural and urbanized catchments. We focused on four complementary and commonly used hydrological signatures: (1) the catchment yield, i.e., how much of precipitation was converted to streamflow; (2) the importance of the low-frequency component, i.e., the baseflow, with regard to the total flow; (3) the characteristic response time, which informs about how rapid the catchment response was; and (4) the capacity of the catchment in damping precipitation flashiness (Baker *et al.*, 2004). For each one of these four aspects, the hydrological indicator was computed as follows:

1. The Q_{mean}/P_{mean} (–) ratio, where Q_{mean} is the mean annual streamflow depth and P_{mean} the mean annual precipitation depth.
2. The baseflow index *BFI* (–), which is the ratio of total baseflow volume to total streamflow volume. The baseflow signal was extracted from the total streamflow time series using a combination of the constant-*k* method (Blume, Zehe, and Bronstert, 2007) and recursive digital filtering (Mei and Anagnostou, 2015; Eckhardt, 2005; Collischonn and Fan, 2013).

3. The characteristic response time DT (h), which was taken as the applied shift on the streamflow time series that maximizes its correlation with the precipitation time series.
4. The FQ/FP (–) ratio, where FQ (or FP) is the flashiness of observed streamflow (or precipitation). The flashiness FX of a signal X was computed as (Holko *et al.*, 2011; Baker *et al.*, 2004):

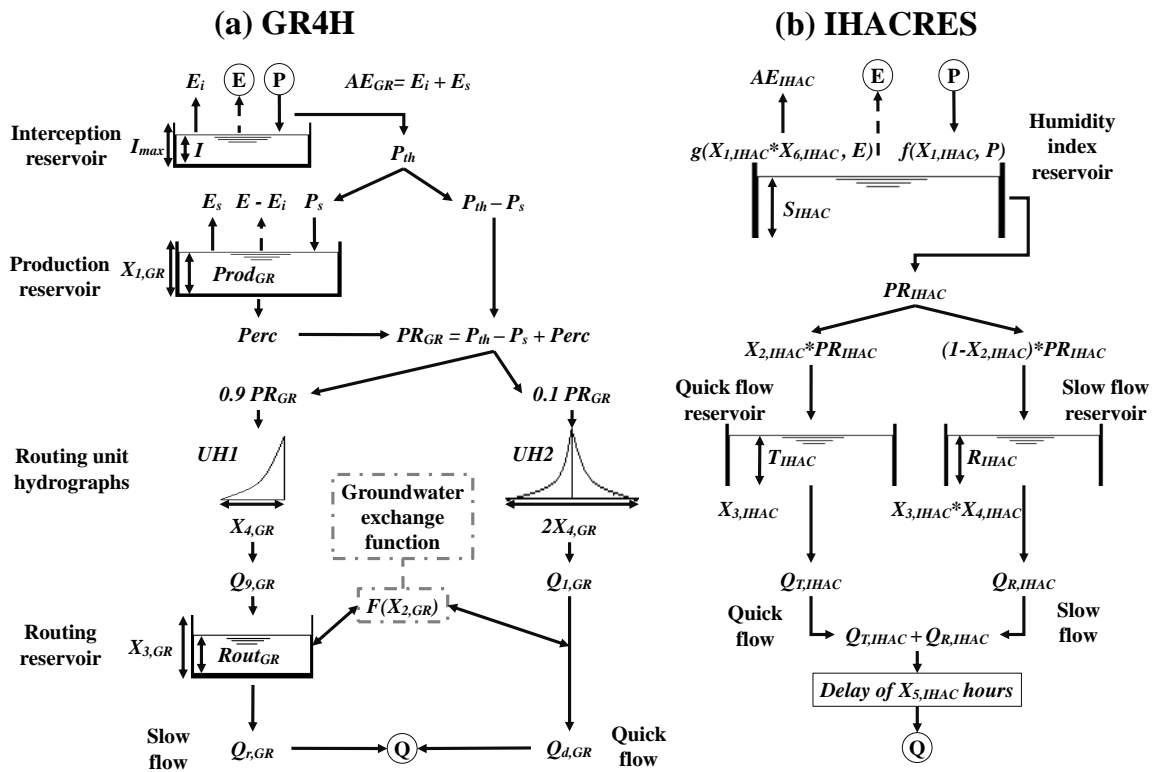
$$FX = \frac{\sum_h |X_h - X_{h-1}|}{\sum_h X_h}$$

where X_h is the value of X at hour h .

4.2.3 Assessment of the differences between urbanized and rural catchments from a rural modeling standpoint

Urbanized catchments are generally characterized by quick runoff generation and propagation due to surface sealing and artificial drainage systems, which imposes a choice of fine modeling time steps (hourly to sub-hourly) to accurately represent the rainfall-runoff relationship in urbanized areas (Dotto *et al.*, 2011; Rodriguez, Andrieu, and Morena, 2008; Salvatore, Bronders, and Batelaan, 2015). We have chosen the hourly time step as a compromise between hydroclimatic data availability and the need for a high temporal resolution to capture as much as possible the dynamics of the hydrological processes in urbanized catchments. Moreover, the characteristic response time DT was at least one hour for the majority of catchments used here (347 out of 350 catchments).

In this study, we used two hourly conceptual rainfall-runoff models, GR4H (Ficchi, Perrin, and Andréassian, 2019) and IHACRES (Jakeman, Littlewood, and Whitehead, 1990). These two models were not developed specifically for urbanized catchments, but they showed their flexibility on a large range of hydrological behaviors. Thus, they were considered to be good candidates to investigate their flexibility over urbanized catchments. The choice of these models was also motivated by their parsimony (GR4H and IHACRES have four and six parameters, respectively, and require only hourly precipitation P and potential evapotranspiration E for forcing; see Figure 4.3). In this way, the model parameter values could be more easily interpreted with respect to catchment behavioral properties. We considered the impact of urbanization on: **(i)** model performance, **(ii)** calibrated model parameter values, and **(iii)** simulated internal fluxes.



Inputs (mm)

P: Precipitation

E: Potential evapotranspiration

GR4H parameters

I_{max} , $X_{1,GR}$, $X_{3,GR}$: Reservoir capacities (mm)

$X_{2,GR}$: Potential exchange parameter (mm/h)

$X_{4,GR}$: Base time of unit hydrographs (h)

GR4H states (mm)

I, $Prod_{GR}$, $Rout_{GR}$: Reservoir states

GR4H internal fluxes and outputs (mm)

E_i , E_s , AE_{GR} : Actual evapotranspiration

P_{th} : Throughfall

P_s : Infiltration

Perc: Percolation

PR_{GR} : Net precipitation

$Q_{9,GR}$, $Q_{1,GR}$: Outputs of UH1 and UH2

F: Potential exchange with groundwater

$Q_{r,GR}$: Slow flow

$Q_{d,GR}$: Quick flow

Q: Total flow

IHACRES parameters

$X_{1,IHAC}$: Forcing parameter (P) (mm)

$X_{1,IHAC} * X_{6,IHAC}$: Forcing parameter (E) (mm)

$X_{2,IHAC}$: Quick flow–slow flow partitioning (-)

$X_{3,IHAC}$: Quick flow linear reservoir parameter (-)

$X_{3,IHAC} * X_{4,IHAC}$: Slow flow linear reservoir parameter (-)

$X_{5,IHAC}$: Delay parameter (h)

IHACRES states

S_{IHAC} : Humidity index (-)

T_{IHAC} , R_{IHAC} : Reservoir states (mm)

IHACRES internal fluxes and outputs (mm)

AE_{IHAC} : Actual evapotranspiration

PR_{IHAC} : Net precipitation

$Q_{R,IHAC}$: Slow flow

$Q_{T,IHAC}$: Quick flow

Q: Total flow

Figure 4.3. (a) GR4H and (b) IHACRES model structures. A detailed description of the GR4H model equations can be found in Ficchi, Perrin, and Andréassian (2019) and Appendix C. The IHACRES model structure used here was derived from a rewriting of the original equations in a reservoir style by Perrin (2000), which was tested later at the hourly time step by Mathevet (2005). The numerically calibrated model parameters for GR4H and IHACRES are noted as $\{X_{i,GR}\}_{1 \leq i \leq 4}$ and $\{X_{j,IHAC}\}_{1 \leq j \leq 6}$, respectively.

The first level of comparison concerned the ability of the model to reproduce the observed streamflow. To this end, a split-sample test was used (Klemeš, 1986). The catchment period of recording was divided into two subperiods P1 and P2. For each subperiod, each model was calibrated with respect to the Kling-Gupta efficiency (KGE) coefficient (Gupta *et al.*, 2009), which measures the differences between two signals using their correlation, the ratio of their means, and the ratio of their standard deviations. The square root transformation was applied to the simulated and observed streamflows in order to guarantee the most uniform weighting possible over the different components of the streamflow (i.e., high, medium, and low flows; see Oudin *et al.*, 2006a; Santos, Thirel, and Perrin, 2018). For calibration, an algorithm consisting of a broad inspection of the model parameter hyperspace, followed by a gradient descent algorithm, was used (Edijatno *et al.*, 1999). The calibrated parameters were then tested on the alternative period (i.e., P2 if calibrated on P1 and P1 if calibrated on P2). Performances were analyzed over the whole test period. In more detail, the models were also compared during the summer (June-August) and the winter (December-February) within the test period, in order to see whether the dry/humid contrast attenuated or accentuated the differences between the rural and the urbanized catchments.

The second level concerned differences in the calibrated parameters. The distributions of model parameters were determined for each catchment set (rural and urbanized) and the differences interpreted in terms of model functioning. Figure 4.3 illustrates model structures, main parameters, as well as internal variables. Table 4.1 provides a brief description of the role of each model parameter.

The third level involved the comparison of internal fluxes simulated by the models to investigate whether the two models handled the urbanized catchment behavioral properties in a similar way. Fluxes were compared over the calibration period using the coefficient of determination R^2 and the *Bias* (i.e., ratio of means). The main structural differences between GR4H and IHACRES can be summarized as follows: **(1)** production is represented in GR4H using a soil moisture accounting reservoir, where percolation is quantified ($Perc$ in Figure 4.3), while it is characterized in IHACRES based on the catchment moisture deficit parametrized as an exponential decay function. In GR4H, the percolation flux $Perc$ sustains the net precipitation PR_{GR} in the absence of events, which fortifies the low-frequency behavior of PR_{GR} . **(2)** The quick flow/slow flow bifurcation is better identified in IHACRES than in GR4H, first through an optimized partitioning parameter ($X_{2,IHAC}$ vs. the fixed 0.1-0.9 partitioning in GR4H), and second by removing any possible exchange between both branches in IHACRES. In GR4H, a groundwater exchange function interacts with both branches, which eventually influences the simulated water budget. Table 4.2 summarizes the comparison between GR4H and IHACRES fluxes and states (see Figure 4.3 for the location of each flux within the model structure).

Table 4.1. Description of GR4H and IHACRES model parameters and their corresponding roles.

Model	Parameter notation and unit	Role
GR4H	$X_{1,GR}$ (mm)	The maximum capacity of the soil moisture accounting reservoir. This can be interpreted as the capacity of the catchment soil to store water. Higher values are associated with low net precipitation PR_{GR} .
	$X_{2,GR}$ (mm/h)	The potential gain from/loss to groundwater. Higher absolute values indicate higher potential amounts of exchanged water.
	$X_{3,GR}$ (mm)	The maximum capacity of the routing reservoir. This controls the slow flow component $Q_{r,GR}$. Higher values are associated with the long-term memory of the catchment, i.e., important volumes of slow flow $Q_{r,GR}$.
	$X_{4,GR}$ (h)	The base time of the unit hydrographs, which is correlated with the characteristic response time of the catchment.
IHACRES	$X_{1,IHAC}$ (mm)	This controls the increase in the humidity index S_{IHAC} due to precipitation. Higher values lead to smaller increases of S_{IHAC} at the event scale, hence less net precipitation PR_{IHAC} .
	$X_{1,IHAC} \cdot X_{6,IHAC}$ (mm)	This controls the decrease in S_{IHAC} due to evapotranspiration. Higher values indicate a slower decrease of S_{IHAC} .
	$X_{2,IHAC}$ (-)	This influences the slow flow/quick flow routing partitioning. Higher values indicate that the net precipitation PR_{IHAC} passes mainly through the fast flow branch.
	$X_{3,IHAC}$ (-)	This controls the amount of the quick discharge $Q_{T,IHAC}$, for which higher values indicate lower amounts of $Q_{T,IHAC}$.
	$X_{3,IHAC} \cdot X_{4,IHAC}$ (-)	This controls the slow flow $Q_{R,IHAC}$. Higher values of $X_{3,IHAC} \cdot X_{4,IHAC}$ are associated with low quantities of $Q_{R,IHAC}$.
	$X_{5,IHAC}$ (h)	The delay to be applied on $(Q_{T,IHAC} + Q_{R,IHAC})$, the sum of the quick and slow flows. This characterizes the rapidity of the catchment response.

Table 4.2. Comparison of GR4H and IHACRES internal fluxes and states.

Internal fluxes/states	Corresponding GR4H flux/state	Corresponding IHACRES flux/state	Comparison metrics
Catchment humidity state	Production reservoir state $Prod_{GR}$	Humidity index S_{IHAC}	R^2
Net precipitation	PR_{GR}	PR_{IHAC}	$R^2, Bias$
Actual evapotranspiration	Sum of E_i and E_s (AE_{GR})	Sum of decreases in S_{IHAC} multiplied by $X_{1,IHAC}$ (AE_{IHAC})	$Bias$
Total losses	Sum of AE_{GR} and the exports to the groundwater ($LOSS_{GR}$)	AE_{IHAC}	$Bias$
Slow flow routing	Routing reservoir state $Rout_{GR}$	Slow flow reservoir state R_{IHAC}	R^2
Slow flow	Slow branch output $Q_{r,GR}$	Slow flow reservoir output $Q_{R,IHAC}$	$R^2, Bias$
Quick flow	Quick branch output $Q_{d,GR}$	Quick flow reservoir output $Q_{T,IHAC}$	$R^2, Bias$
Total flow	Simulated flow Q ($Q_{sim,GR}$)	Simulated flow Q ($Q_{sim,IHAC}$)	$R^2, Bias$

4.3 Results

4.3.1 Are urbanized catchments hydrologically different from their rural neighbors?

In general, all the hydrological signatures of the urbanized catchments deviated from those of their rural neighbors, as illustrated in Figure 4.4. At low levels of TIA , urbanized catchments exhibited values similar to their rural neighbors, and the greater the TIA , the more the deviation between the two samples increased. These results suggest that urbanization impacted all facets of hydrological behavior. However, in terms of range (or distribution), the hydrological signatures of urbanized catchments were not so different from those of rural catchments. This means that the large diversity encountered in rural catchments encompassed the specificities of the urbanized ones. This was true for all the considered hydrological signatures, except for the flashiness ratio FQ/FP , which showed very high values in urbanized catchments compared with rural catchments.

For a given volume of precipitation, results on Q_{mean}/P_{mean} (Figure 4.4a) indicate that it is more likely to have a higher volume of streamflow in an urbanized catchment than in a

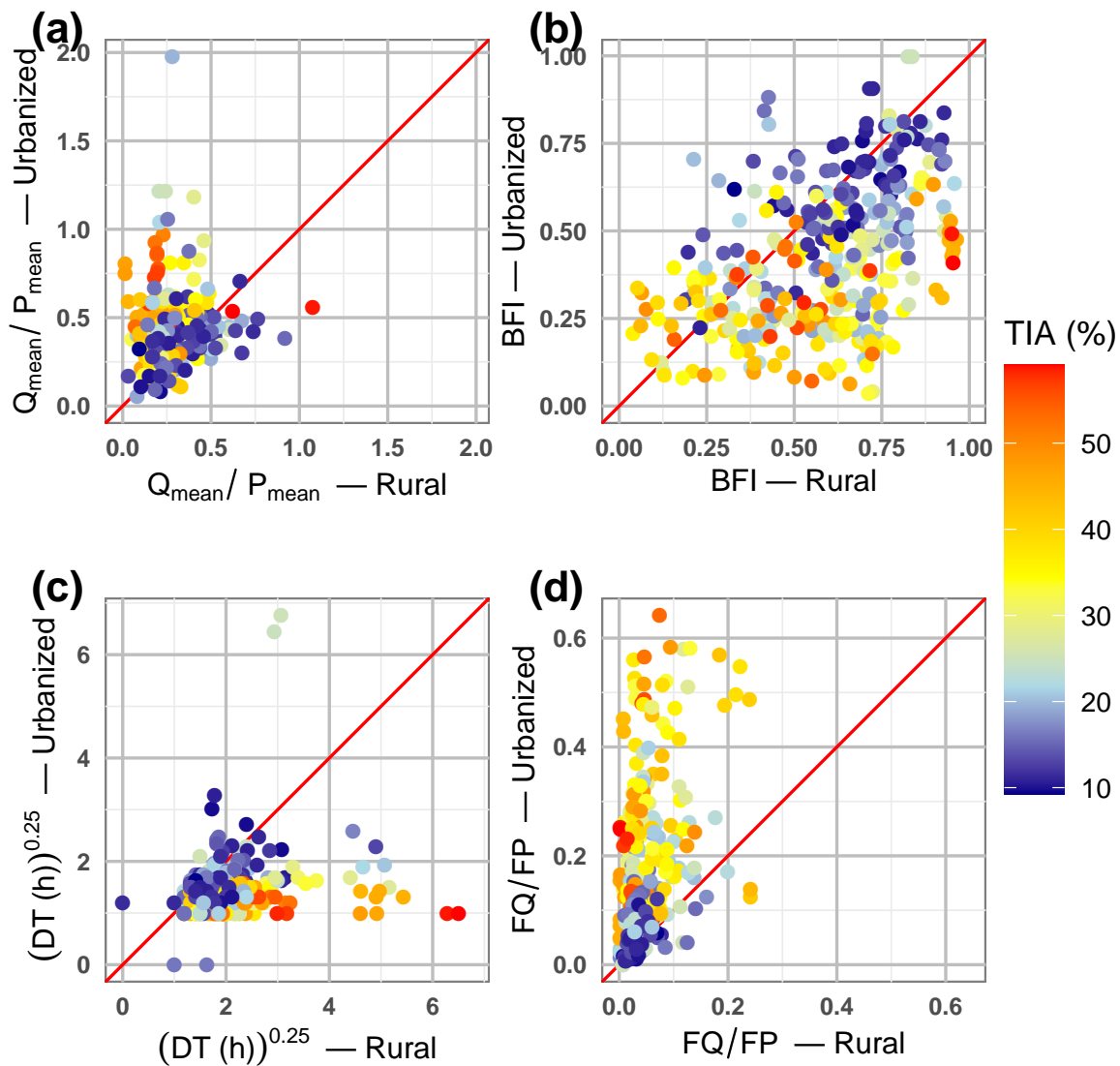


Figure 4.4. Values of the hydrological signatures for urbanized catchments (on y -axes) vs. their rural neighbors (on x -axes): (a) the ratio of mean annual streamflow to mean annual precipitation ($Q_{\text{mean}}/P_{\text{mean}}$), (b) the baseflow index (BFI), (c) the characteristic response time (DT), and (d) the ratio of streamflow flashiness to precipitation flashiness (FQ/FP). Colors indicate the mean total impervious area (TIA) of the urbanized catchment, which ranged between 10% and 59%. The TIA range of their rural neighbors was 0%-5%. Some catchments show abnormal values of $Q_{\text{mean}}/P_{\text{mean}}$ ratio (i.e., higher than 1). These values were exhibited by catchments whose outlets were generally within or close to an urban settlement (these included the USGS station 08116400 on the Dry Creek near Rosenberg, Texas, and the USGS station 02310525 on the Weeki Wachee River near Brooksville, Florida).

rural one: median Q_{mean}/P_{mean} ratio reached 0.39 in urbanized cases, while it was around 0.27 in rural catchments. Conversely, median baseflow index (Figure 4.4b) was higher in rural contexts (0.62 vs. 0.42 in urbanized catchments), indicating that, for the majority of pairs, low flow decreased in urbanized catchments. The characteristic response times DT (Figure 4.4c) were much shorter in urbanized catchments compared with rural catchments (a median DT of 3 h for urbanized catchments vs. 14 h for their rural neighbors), suggesting faster responses for the urbanized cases. Also, FQ/FP ratios (Figure 4.4d) were much higher in high-imperviousness catchments (a median of 0.13 vs. 0.03 in rural catchments), implying that the temporal irregularities of precipitation were more damped in rural catchments than in urbanized ones.

4.3.2 Are there differences between urbanized and rural catchments observable from a rural modeling standpoint?

In terms of model performances, GR4H revealed no clear distinction between urbanized and rural catchments, whereas IHACRES simulated better urbanized catchments, especially during summer (Figure 4.5). The parameter distributions of both models were impacted by urbanization (Figures 4.6 and 4.7), as well as the agreement of GR4H and IHACRES in terms of internal fluxes (Figure 4.8).

Our first level of investigation is model performances, as one would expect that rural models would not be suitable for urbanized catchments given the differences in landscape and hydrological behavior. Concerning calibration (Figure 4.5), both conceptual models showed similar performances over the rural and urbanized catchment sets, meaning that these models were flexible enough to reproduce the rainfall-runoff relationship for the urbanized catchments. IHACRES showed even greater ability in reproducing observed flows in urbanized catchments than in rural ones. Test performances showed a drop in median performances of about 0.07-0.08 of KGE in a similar manner for both models over both catchment samples. When the evaluation concerned only the summer or the winter periods, performances decreased, especially for the summer-based evaluation. However, it is during the summer when IHACRES test performances were significantly better in the urbanized catchments compared with their rural neighbors, as the corresponding difference in median KGE reached 0.09.

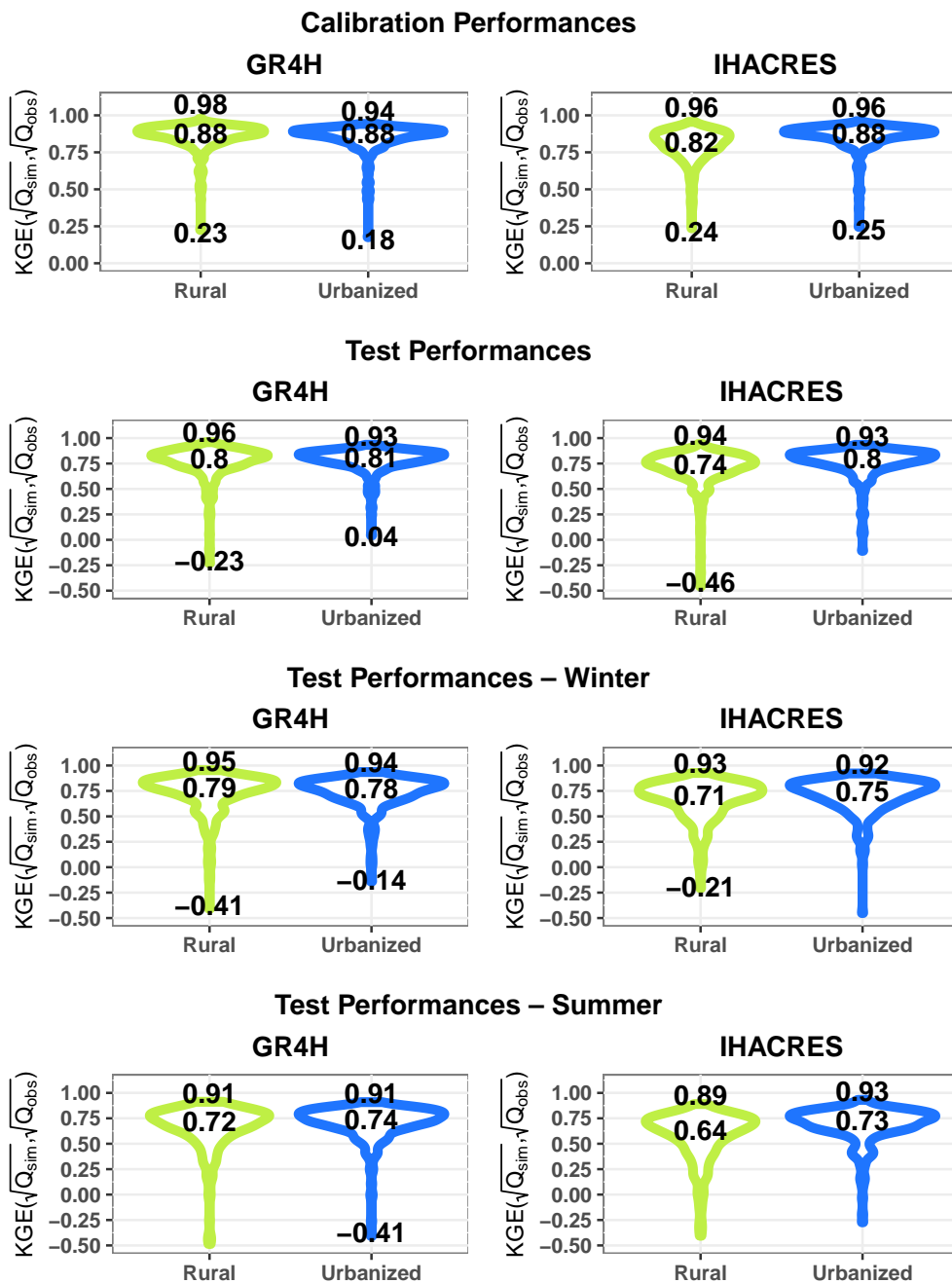


Figure 4.5. Distributions of GR4H (left column) and IHACRES (right column) performances in calibration (first row) and test over the whole test period (second row), with a focus on the winter (third row) and the summer periods (fourth row). Performances are shown for the urbanized catchments (*Urbanized*, in blue) and their rural neighbors (*Rural*, in green). The assessment is shown in terms of the Kling-Gupta Efficiency (*KGE*) score applied on square root values of simulated (Q_{sim}) and observed (Q_{obs}) streamflow time series. Values indicate the minimum, the median, and the maximum. It should be remembered that the ideal value of *KGE* is 1. In many cases, the minimum *KGE* values for rural and urbanized catchments were lower than -0.5 (Test performances over whole period: IHACRES-*Urbanized*: -0.76 ; Test performances, focus on winter: IHACRES-*Urbanized*: -1.24 ; Test performances, focus on summer: GR4H-*Rural*: -11.04 , IHACRES-*Rural*: -22.00 , IHACRES-*Urbanized*: -1.06).

Since GR4H and IHACRES performances were satisfactory for both urbanized and rural catchments, their parameters could be now compared so as to see which ones were impacted by urbanization. We can notice that non-impacted performances of GR4H (Figure 4.5) were countered by different parameter distributions for rural and urbanized catchments, as shown in Figure 4.6. In urbanized catchments, $X_{1,GR}$ showed slightly higher values than in rural ones. On the other hand, $X_{2,GR}$ was more centered on zero in the urbanized set, while it became negative in the majority of rural catchments, meaning that rural catchments exported more water to groundwater compared with urbanized catchments. These two variations for $X_{1,GR}$ and $X_{2,GR}$ yield more runoff for the urbanized catchments, by reducing the losses through evapotranspiration and deep infiltration. The differences were more pronounced concerning $X_{3,GR}$ and $X_{4,GR}$, for which lower values were encountered in the urbanized catchments compared with rural ones. This suggests a short-term memory behavior and more pronounced low flows over urbanized catchments, accompanied by a faster response compared with rural contexts.

Considering IHACRES parameters, the differences can be identified for each parameter, as illustrated in Figure 4.7. The parameter $X_{1,IHAC}$ was significantly lower in urbanized catchments than in rural ones, indicating that the humidity index S_{IHAC} increased more quickly at the event scale in urbanized catchments, thus converting a higher portion of precipitation into runoff compared with rural catchments. However, $X_{1,IHAC} \cdot X_{6,IHAC}$ indicates that S_{IHAC} tended to decrease more quickly in urbanized catchments, especially during dry periods, but the fact that $X_{1,IHAC}$ was lower in this set indicates that the urbanized catchments responded more sensitively to summer events (i.e., dry-period events, due to quick increase in S_{IHAC}) than their rural neighbors did. Therefore, the variance of S_{IHAC} was higher in urbanized catchments. The $X_{2,IHAC}$ parameter was higher in urbanized catchments, which indicates that the excess water during an event passed mostly through the quick flow branch. Conversely, the $X_{3,IHAC}$ values were lower in the urbanized set, whereas the $X_{3,IHAC} \cdot X_{4,IHAC}$ product values were higher. This indicates that the quick flow volumes were higher in urbanized catchments, while slow flow volumes were higher in rural catchments. The low values of $X_{5,IHAC}$ encountered in urbanized catchments suggest that they reacted faster than their rural neighbors.

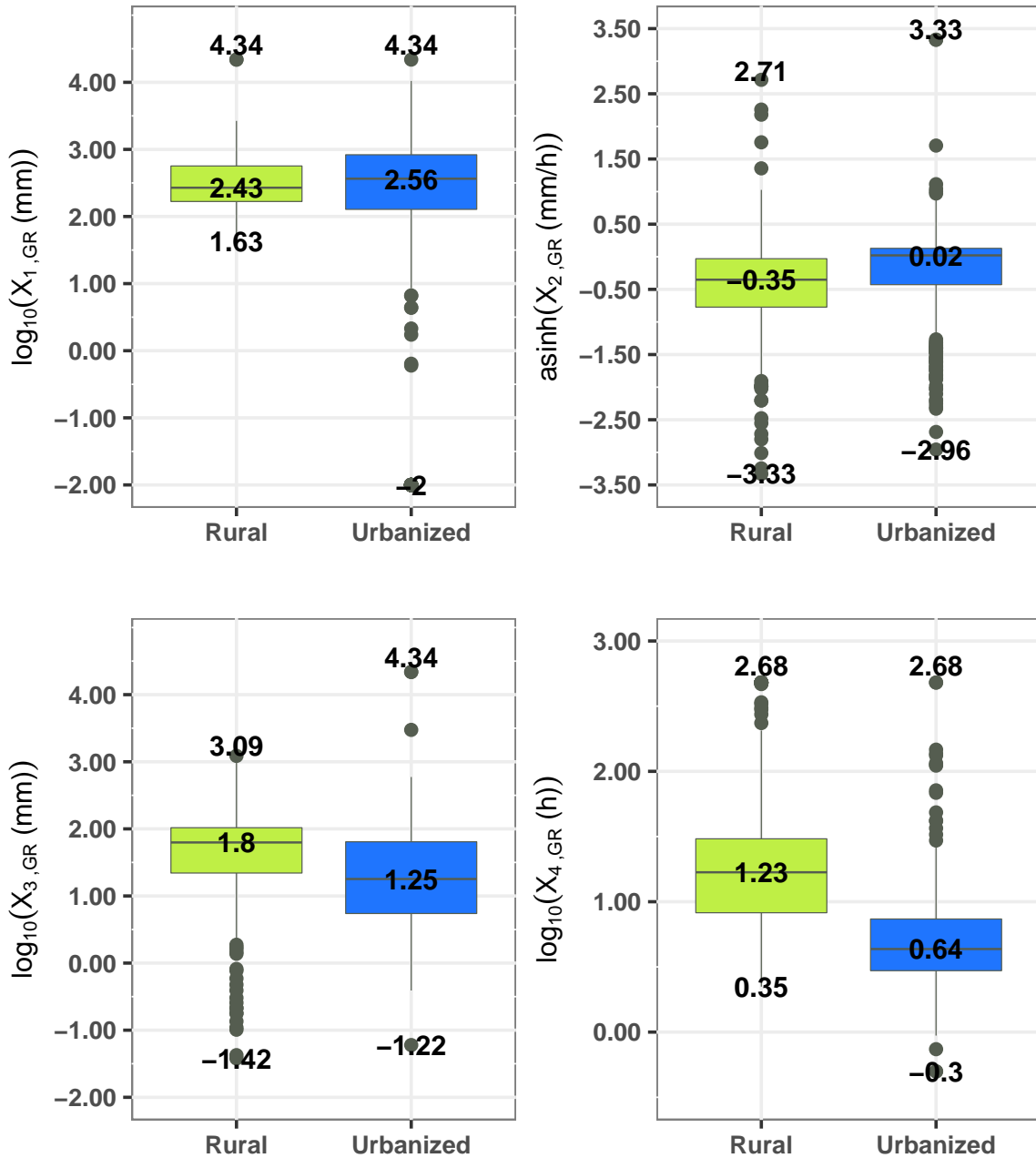


Figure 4.6. Distributions of GR4H parameters over the urbanized catchments (in blue) and their rural neighbors (in green). According to Student’s t -test (Section 4.6), all differences between rural and urbanized samples were statistically significant (p -value < 0.001), except for $X_{1,GR}$. Mann-Whitney-Wilcoxon test (Section 4.6) indicated that all parameter distributions were statistically different (p -value < 0.001). Transformations were applied to better normalize the distributions. Values indicate the minimum, the median, and the maximum. Boxes are delimited by the first and third quartiles.

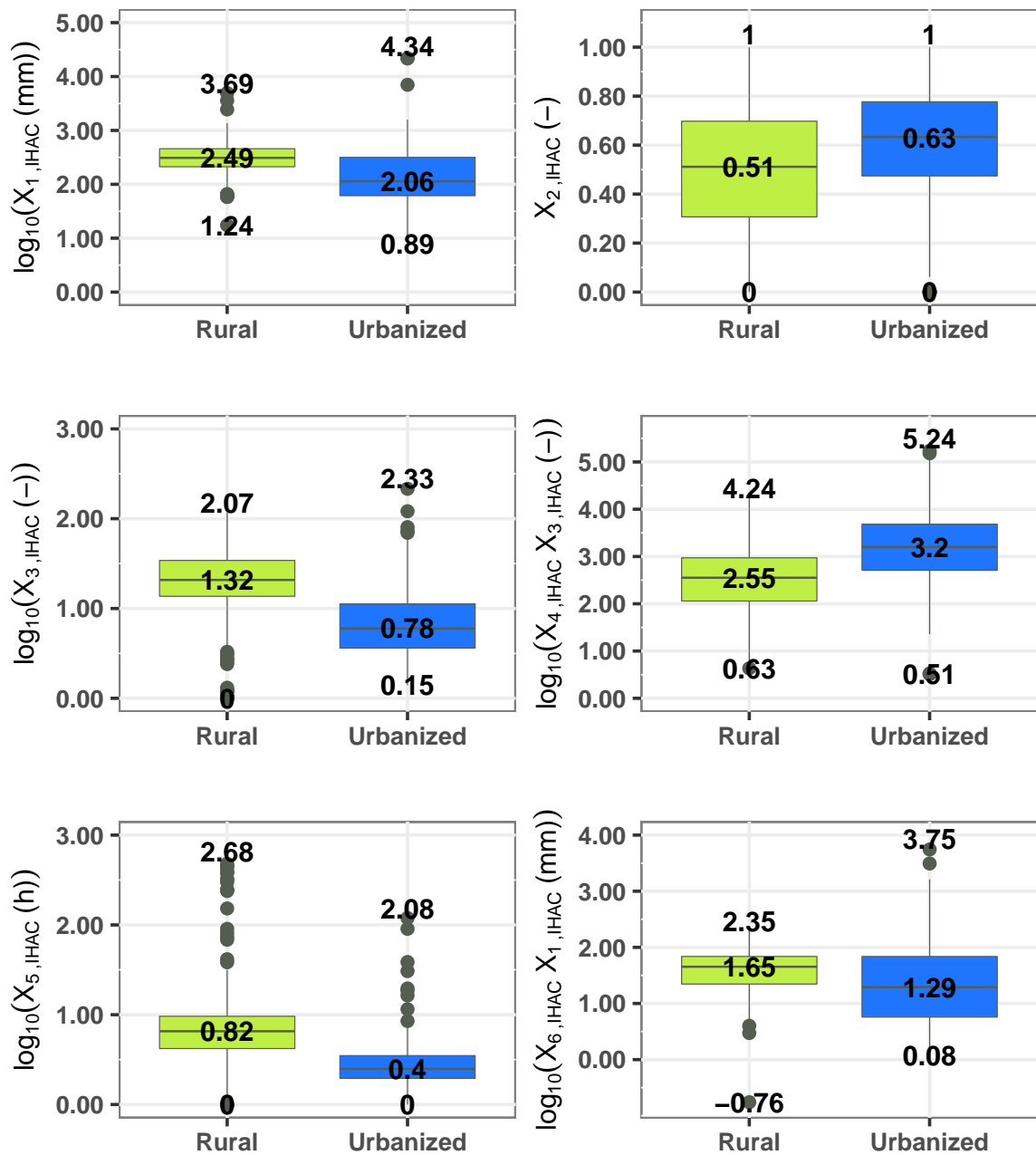


Figure 4.7. Distributions of IHACRES parameters in the urbanized catchments (in blue) and their rural neighbors (in green). According to Student’s *t*-test (Section 4.6), all differences between rural and urbanized samples were statistically significant (*p*-value < 0.001). Mann-Whitney-Wilcoxon test (Section 4.6) indicated that all parameter distributions were statistically different (*p*-value < 0.001). Transformations were applied to better normalize the distributions. Values indicate the minimum, the median, and the maximum. Boxes are delimited by the first and third quartiles.

Finally, it is noteworthy to examine the internal fluxes and outputs of both models, first to see whether GR4H and IHACRES simulated similarly the rural catchments, and if so, whether this agreement was altered by urbanization. In this respect, GR4H and IHACRES internal fluxes were more similar in rural catchments than in urbanized ones, indicating impacted internal model fluxes and states by urbanization, as shown in Figure 4.8 by the median values of R^2 and the *Bias* on the internal model fluxes and states (listed in Table 4.2).

In terms of R^2 , the internal states of the models (i.e., $Prod_{GR}$ vs. S_{IHAC} and $Rout_{GR}$ vs. R_{IHAC}) showed higher correlations when simulating the rural catchments ($R^2(Prod_{GR}, S_{IHAC}) = 0.69$ and $R^2(Rout_{GR}, R_{IHAC}) = 0.53$, median values) than the urbanized catchments. Correlations in reproducing the net precipitation PR were comparatively high, with median R^2 values of approximately 0.9 in rural catchments compared with 0.76 for urbanized catchments. Both models disagreed significantly in simulating slow flow, whereas they agreed in terms of the quick flow, with a slightly higher median correlation in urbanized catchments than in rural ones. Finally, although not as strongly correlated as when simulating net precipitation, the models simulated total flow in a relatively similar manner, with slightly higher correlations in rural catchments than in urbanized ones (rural median $R^2(Q_{sim,GR}, Q_{sim,IHAC})$ attaining 0.78 vs. $R^2(Q_{sim,GR}, Q_{sim,IHAC})$ median value of 0.73 for the urbanized catchments).

The *Bias* terms showed weaker agreement in reproducing the volumes of the analyzed fluxes. With the exception of the simulated streamflows, the values of $Bias(PR_{GR}, PR_{IHAC})$ were the closest to 1, with slightly higher amounts of PR_{GR} produced in rural catchments compared with PR_{IHAC} (i.e., $Bias(PR_{GR}, PR_{IHAC}) > 1.0$). The total losses estimated in GR4H were generally lower than the ones estimated in IHACRES. Slow flow volumes were significantly lower in IHACRES than in GR4H, while the inverse was the case for quick flow volumes. Ultimately, the total simulated flows were virtually the same, leading to a median *Bbias* very close to 1.

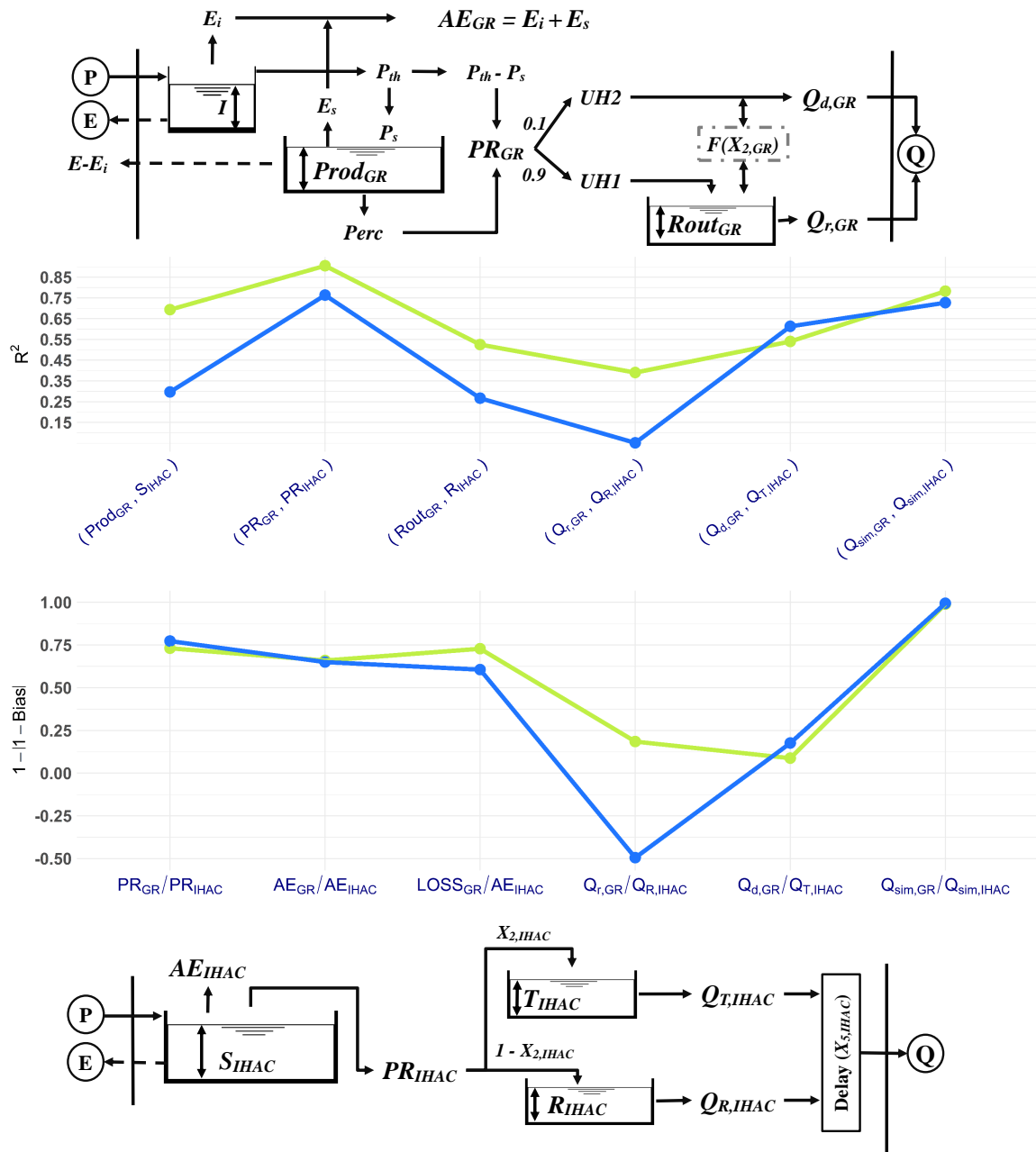


Figure 4.8. Comparison between different GR4H and IHACRES internal fluxes and states (listed in Table 4.2) in terms of median values of R^2 and the median distance of the bias (ratio of means) to 1, in the rural (green lines) and urbanized catchments (blue lines). This comparison concerns the calibration period only. The meaning of acronyms of internal fluxes and states can be found in Figure 4.3. $LOSS_{GR}$ is computed as the total losses via actual evapotranspiration and groundwater exchange.

4.4 Discussion

4.4.1 Impact of urbanization on the hydrological behavior of catchments

Our purpose of comparing the hydrological behavior of urbanized catchments with that of their rural neighbors was to check whether landscape differences (i.e., urbanized vs. non-urbanized) impacted hydrological characteristics. Overall, the obtained results corroborate the literature findings concerning the impact of urbanization, which has been a widely investigated question for decades (Leopold, 1968; Hollis, 1977; Oudin *et al.*, 2018). In this regard, the original facets of our findings are twofold: (1) The examination of hydrological behavior by computing the signature metrics over continuous and relatively long hourly time series, in contrast with event- and season-based metrics (Miller and Hess, 2017; Zhou *et al.*, 2017; Sillanpää and Koivusalo, 2015); (2) The use of a large sample of urbanized catchments in two different countries, which is a key strength in the context of abundant single-catchment experiments in urbanization studies (Salvadore, Bronders, and Batelaan, 2015).

In more detail, the high values Q_{mean}/P_{mean} in urbanized catchments is in line with the large number of empirical studies showing that mean annual flows increase with urban development (DeWalle *et al.*, 2000; Oudin *et al.*, 2018), due to either increased imperviousness or the effluents of water treatment facilities (Diem, Hill, and Milligan, 2018; Zhou *et al.*, 2017).

Conversely, the impact of urbanization on low flows is still difficult to assess. On one hand, the *BFI* could be higher in some urbanized catchments owing to water treatment facility releases during the dry period (Oudin *et al.*, 2018), reduction in vegetation cover, or leakage in high-pressure potable water systems (Göbel *et al.*, 2004; Lancia *et al.*, 2019; Sanzana *et al.*, 2019). On the other hand, *BFI* could be lower in some urbanized catchments in which groundwater pumping is intensive and the river-groundwater connection is cut by sealing the riverbanks (Braud *et al.*, 2013; Brun and Band, 2000; Zektser and Everett, 2004). These two possible changes are observed in our catchment set, but for the majority, *BFI* decreased.

Flow velocities on sealed surfaces are higher than on natural ones (i.e., bare or vegetated soils), due to the reduction of surface roughness. This becomes pronounced with the higher amounts of runoff during an event, ultimately leading to an acceleration of the catchment response (Burns *et al.*, 2005; Miller and Hess, 2017).

To conclude, we found that urbanization modified many facets of hydrological behavior, from water budget to high and low flow characteristics. However, of the signatures considered here, only *FQ/FP* showed significantly different ranges for the two catchment sets (rural and urbanized). This means that the impact of urbanization on hydrological behavior was

not abrupt (Mejía *et al.*, 2015), hence, the diversity of behaviors in rural catchments still encompasses the behaviors of the urbanized ones. Therefore, the large flashiness of the streamflow signal (FQ/FP) is a very specific characteristic to distinguish urbanized from rural catchments.

4.4.2 Differences between urbanized and rural catchments from a modeling standpoint

4.4.2.1 Detection of differences in hydrological behavior by model parameters

The differences shown above in terms of GR4H and IHACRES parameters (Figures 4.6 and 4.7) are consistent with the obtained differences in the hydrological signatures (Figure 4.4). By shifting the production parameters (low mean $X_{1,GR}$, close-to-zero values of $X_{2,GR}$, and low $X_{1,IHAC}$), the models simulated higher Q_{mean}/P_{mean} in the urbanized catchments, whereas lower values of BFI (i.e., much smaller slow-flow component compared with quick flow) were confirmed by the routing parameters (lower values of $X_{3,GR}$, higher values of $X_{2,IHAC}$, lower values of $X_{3,IHAC}$, and higher values of $X_{3,IHAC} \cdot X_{4,IHAC}$). The rapid response and the high flashiness of the urbanized catchments, manifested by short DT and high FQ/FP , were reflected by the timing parameters ($X_{4,GR}$ and $X_{5,IHAC}$).

It may not be surprising that the parameters corroborate what we learn from the hydrological signatures, since the parameters were calibrated in such a way that the simulated outflows of the models match as closely as possible the observed streamflow time series in the urbanized catchments. This helped to infer the impact of urbanization from change in model parameters (McIntyre and Marshall, 2010; Pathiraja *et al.*, 2018; Seibert and McDonnell, 2010). However, this is not usually straightforward, as the model parameters are not only land-use dependent, but also morphology-, geopedology-, and climate-dependent (Merz, Parajka, and Blöschl, 2011), which might explain some of the parameter dispersion. Through a paired-catchment approach, some effects (e.g., related to climate and geopedology) are supposed to be limited, but not to the extent that we would be able to witness differences in parameter distributions based only on the imperviousness level. Besides, the use of a parsimonious model structure is a key factor in limiting the equifinality and the non-identifiability of the parameters (Beven, 2006; Ebel and Loague, 2006), without which interpretation might have been more difficult.

4.4.2.2 Flexibility of conceptual models with regard to urbanized catchments

The performances of GR4H and IHACRES were not negatively impacted by urbanization (Figure 4.5). IHACRES and GR4H had similar performances over the urbanized catchment set, which were even better than their performances in rural catchments. This is an indication that conceptual models can be flexible enough to account for the specificities of urbanized catchments, which is likely due to the fact that the range of behaviors in rural catchments can encompass a wide range of those of urbanized catchments, as discussed above. Test performances registered a similar drop in performances regardless of urbanization, especially for the summer-based evaluation, a result that has been encountered in many model evaluation studies (Brigode, Oudin, and Perrin, 2013; Boer-Euser *et al.*, 2017).

Although GR4H and IHACRES performances in urbanized catchments were more similar than in rural ones, they handled urbanized catchments more differently (Figure 4.8). Despite the fact that both models simulated runoff differently, their generated net precipitation fluxes (PR_{GR} and PR_{IHAC}) were highly correlated in rural catchments, but this agreement was less significant in urbanized catchments. It is reasonable to say that this good agreement is merely a response to the same precipitation forcing. However, the very significant drop in correlations encountered between $Prod_{GR}$ and S_{IHAC} states suggests that the models handle urbanization differently. Moreover, less agreement was found in the routing processes, perhaps because more freedom (i.e., free parameters) is left to IHACRES to better represent quick/slow flow partitioning than GR4H. This can be observed through the slow flow branch, for which the two models scored the lowest values of R^2 . In addition, the *Bias* of $Q_{r,GR}$ and $Q_{r,IHAC}$ confirmed this disagreement in representing slow flows. The relatively high proportion of slow flow of GR4H was also perceived to be a structural characteristic (Boer-Euser *et al.*, 2017). These differences in slow flow volumes were compensated for thanks to the quick flow branches reaching better concordance in simulated total flow volumes (*Bias* in Q_{sim} around 1).

4.5 Conclusion

As a key finding of this study, conceptual models GR4H and IHACRES were flexible enough to reproduce the rainfall-runoff relationship of 175 US and French urbanized catchments. Also, the two models were able to handle the behavioral specificities of urbanized catchments, which we found to be deviated from the specificities of their rural neighbor catchments using four hydrological signature metrics.

Over a large sample of 175 urbanized catchments with $TIA > 10\%$, the hydrological characteristics, i.e., the Q_{mean}/P_{mean} ratio (to analyze the catchment yield), the *BFI* (to characterize the

baseflow), the lag time DT (to characterize the catchment response time), and the FQ/FP ratio (to analyze the capacity of the catchment to damp precipitation flashiness), have confirmed the specificities of the hydrological behavior of urbanized catchments, which can be summarized as more runoff in a shorter span of time, with a reduction in the baseflow contribution. These specificities were demonstrated thanks to the paired-catchment style experiment, but it should be noted that, for the majority of characteristics (Q_{mean}/P_{mean} , BFI and DT), the urbanized catchment sample fell into the range of the rural catchments. Conversely, the flashiness ratio (FQ/FP) in some urbanized catchments exceeded the range observed in rural catchments, making the flashiness of the response a marker of urbanized catchment behavior.

Despite the absence of explicit urban-specific components within their structures, GR4H and IHACRES have shown satisfactory results in both rural and urbanized catchment sets. Specifically, GR4H handled both samples similarly, whereas IHACRES demonstrated greater ability to reproduce the rainfall-runoff relationship in urbanized catchments than in rural ones. The capacity of the models to reproduce the observed runoff in different states of urbanization (TIA between 10% and 59%) has led to different parameter distributions for urbanized and rural catchments. These two points, i.e., good model performances and distinguishable parameter distributions, are promising for model-based evidence of the impact of urbanization on catchment behavior. This also corroborates the view that the urbanized catchment behavior is nothing more than a particular hydrological behavior, which is satisfactorily reproducible by a rural model and for which rural model parameters take different distributions when compared with a predominantly rural-landscaped catchment. However, we showed that GR4H and IHACRES handle urbanized catchments differently, owing to their structural differences and degrees of freedom, but both models were flexible enough to perform well in urbanized catchments.

This study provides the community with elements that could justify the application of rural models to urban-impacted catchments. For perspective studies, a sensitivity analysis with respect to different TIA thresholds would clarify the urbanization level above which changes in hydrological behavior become significant. The fact that the models do not explicitly represent urban features makes their use for future scenarios impractical. Attempts at relating the parameters of a rural model to an urbanization measure have shown the weak sensitivity of the model parameters to changes in the urban measure (Saadi, Oudin, and Ribstein, 2019), meaning that an adaptation of their structure is needed in order to overcome these limitations.

4.6 Supporting information

We investigated the statistical significance of the differences between rural and urbanized samples using two tests: the Student's t -test which we applied on transformed values to

analyze how significant the differences in terms of means were, and the Mann-Whitney-Wilcoxon test which we used to examine how significantly different the distributions were. We have considered the differences in hydrological signatures, model performances, and model parameter values. Results are summarized in Table 4.3.

Table 4.3. Results of statistical tests investigating distribution (Mann-Whitney-Wilcoxon) and mean (Student's *t*-test) equality between the urbanized and rural samples in terms of hydrological characteristics, model performances, and model parameters.

Characteristic/ performance/ parameter	Transformation ^a	Rural median	Urbanized median	Are the distributions equal? ^c	Rural mean ^b	Urbanized mean ^b	Are the means equal? ^c
$Q_{mean}/P_{mean} (-)$	Identity	0.27	0.39	No (***)	0.29	0.42	No (***)
$BFI (-)$	Identity	0.62	0.42	No (***)	0.58	0.43	No (***)
$DT (h)$	\log_{10}	14.00	3.00	No (***)	15.74	3.74	No (***)
$FQ/FP (-)$	Identity	0.03	0.13	No (***)	0.05	0.17	No (***)
$KG_{GR,CAL} (-)$	Identity	0.88	0.88	No (*)	0.86	0.85	Yes
$KG_{GR,VAL} (-)$	Identity	0.80	0.81	Yes	0.77	0.79	Yes
$KG_{GR,VAL,Summer} (-)$	Identity	0.72	0.74	No (**)	0.60	0.71	No (*)
$KG_{GR,VAL,Winter} (-)$	Identity	0.79	0.78	Yes	0.73	0.74	Yes
$X_{1,GR} (mm)$	\log_{10}	268.55	365.59	Yes	308.82	61.21	No (***)
$X_{2,GR} (mm/h)$	$\operatorname{arcsinh}$	-0.36	0.02	No (***)	-0.44	-0.22	No (***)
$X_{3,GR} (mm)$	\log_{10}	62.72	17.90	No (***)	39.88	18.05	No (***)
$X_{4,GR} (h)$	\log_{10}	16.82	4.34	No (***)	19.66	4.97	No (***)
$KG_{IHAC,CAL} (-)$	Identity	0.82	0.88	No (***)	0.80	0.86	No (***)
$KG_{IHAC,VAL} (-)$	Identity	0.74	0.80	No (***)	0.71	0.76	No (***)
$KG_{IHAC,VAL,Summer} (-)$	Identity	0.64	0.73	No (***)	0.45	0.68	No (**)

Continued on next page...

Table 4.3 continued: Results of statistical tests investigating the differences between rural and urbanized samples.

Characteristic/ performance/ parameter	Transformation ^a	Rural median	Urbanized median	Are the distributions equal? ^c	Rural mean ^b	Urbanized mean ^b	Are the means equal? ^c
$KGE_{IHAC,VAL,Winter} (-)$	Identity	0.71	0.75	No (***)	0.66	0.70	No (*)
$X_{1,IHAC}$ (mm)	\log_{10}	309.60	113.89	No (***)	304.88	140.01	No (***)
$X_{2,IHAC} (-)$	Identity	0.51	0.63	No (***)	0.48	0.60	No (***)
$X_{3,IHAC} (-)$	\log_{10}	20.78	5.98	No (***)	18.49	6.89	No (***)
$X_{3,IHAC} \cdot X_{4,IHAC} (-)$	\log_{10}	355.30	1589.78	No (***)	337.03	1624.36	No (***)
$X_{5,IHAC}$ (h)	\log_{10}	6.55	2.49	No (***)	7.29	2.72	No (***)
$X_{1,IHAC} \cdot X_{6,IHAC}$ (mm)	\log_{10}	45.11	19.56	No (***)	37.52	20.38	No (***)

^a Transformations were applied on the variable to better normalize it. Only the Student's t -test is applied on transformed values.

^b The mean of a sample was computed as the transformation's inverse function of the mean of transformed values.

^c Yes: $p \geq 0.05$; No(*): $0.01 \leq p < 0.05$; No(**): $0.001 \leq p < 0.01$; No(***): $p < 0.001$.

End of Table 4.3.

Synthèse du chapitre

Dans ce chapitre, nous avons analysé les structures des modèles GR4H et IHACRES pour déterminer si la présence de l'urbanisation (que les deux modèles ne prennent pas en compte explicitement) conduisait à des performances moindres sur des bassins fortement urbanisés situés en France et aux États-Unis. Comme résultat, nous avons trouvé qu'un modèle hydrologique conceptuel comme GR4H est capable de reproduire de manière satisfaisante le comportement hydrologique dans des situations fortement urbanisées, même sans prise en compte explicite des spécificités de l'urbanisation. Ceci a été rendu possible à travers les paramètres du modèle qui ont changé pour s'accommoder aux caractéristiques du comportement hydrologique des bassins fortement urbanisés, ce qui engendre des distributions de paramètres différentes de celles des bassins faiblement urbanisés, et plus ou moins dépendantes du niveau d'urbanisation du bassin. Ces distributions traduisent les différences de comportement hydrologique entre les bassins fortement urbanisés et leurs voisins faiblement urbanisés, confirmant l'impact de l'urbanisation sur les processus hydrologiques à l'échelle du bassin versant. Cette étape d'analyse de structure de départ a éclairé des voies d'amélioration pour une meilleure prise en compte de l'urbanisation par le modèle.

CHAPITRE 5

LES MODÈLES CONCEPTUELS PLUIE-DÉBIT SONT-ILS RÉGIONALISABLES VIS-À-VIS DE L'URBANISATION ?

Sommaire

5.1	Introduction	197
5.2	Data	201
5.3	Methods	206
5.4	Results	211
5.5	Discussion and conclusions	219

Le chapitre précédent a montré que le modèle GR4H a été capable de reproduire le comportement hydrologique de bassins versants fortement urbanisés, même si sa structure ne représente explicitement aucune spécificité liée à l'urbanisation. De plus, ses paramètres ont montré des distributions différentes entre les bassins fortement et faiblement urbanisés, ce qui suggère une dépendance des paramètres au niveau d'urbanisation du bassin versant. Ces résultats nous amènent à réfléchir sur une manière de lier les paramètres aux spécificités de l'urbanisation, en d'autres termes de régionaliser le modèle GR4H. La régionalisation (si elle était performante) permettrait d'obtenir des relations qui feraient le pont entre les scénarios d'urbanisation et les paramètres du modèle, qui à leur tour moduleraient le comportement hydrologique.

Correspondant à un article publié dans *Water*¹⁴, le présent chapitre montre une tentative de régionalisation des paramètres du modèle hydrologique GR4H en fonction de plusieurs descripteurs climatiques, morphologiques, géopédologiques et d'occupation du sol, dont fait partie une mesure de l'urbanisation. Cette régionalisation est effectuée dans un cadre de régression multiple non linéaire à l'aide de l'algorithme des forêts aléatoires (*random forest*). Nous cherchons à répondre à deux questions principales :

1. Étant donnés des bassins fortement urbanisés (pour lesquels la fraction du couvert urbain *CPD* dépasse 20%), est-ce que les relations obtenues à l'aide du *random forest* entre les descripteurs du bassin et les paramètres du modèle GR4H sont aussi performantes que les méthodes classiques de transfert de paramètres (transfert temporel, proximité spatiale, similarité paysagère) ?
2. Si ces relations sont satisfaisantes d'un point de vue performances du modèle, est-ce qu'elles sont exploitables pour des études de scénarios d'urbanisation ?

Pour répondre à ces questions, nous avons employé tout l'échantillon des 2225 bassins versants français et nord-américains.

¹⁴Sous le titre : "Random forest ability in regionalizing hourly hydrological model parameters", doi : [10.3390/w11081540](https://doi.org/10.3390/w11081540).

Abstract

This study investigated the potential of random forest (RF) algorithms for regionalizing the parameters of an hourly hydrological model. The relationships between model parameters and climate/landscape catchment descriptors were multidimensional and exhibited nonlinear features. In this case, machine-learning tools offered the option of efficiently handling such relationships using a large sample of data. The performance of the regionalized model using RF was assessed in comparison with local calibration and two benchmark regionalization approaches. Two catchment sets were considered: **(1)** A target pseudo-ungauged catchment set was composed of 120 urban catchments and **(2)** 2105 gauged American and French catchments were used for constructing the RF. By using pseudo-ungauged urban catchments, we aimed at assessing the potential of the RF to detect the specificities of the urban catchments. Results showed that RF-regionalized models allowed for slightly better streamflow simulations on ungauged sites compared with benchmark regionalization approaches. Yet, constructed RFs were weakly sensitive to the urbanization features of the catchments, which prevents their use in straightforward scenarios of the hydrological impacts of urbanization.

5.1 Introduction

5.1.1 Why and how do we regionalize hydrological model parameters?

Hydrological models are used for various purposes to represent the water cycle and processes in a defined space-time domain. According to their objectives and their level of complexity, the models are either parsimoniously or heavily parametrized. In order to realistically describe the water movement within the spatially delimited domain (e.g., a hydrological catchment), explicit physically-based and distributed approaches are useful to better track the spatial and temporal variability and the non-linearity of the hydrological processes. This usually ends up in dealing with highly parameterized hydrological models compared to the level of data availability and modeling constraints, which results in high degrees of freedom and parameter uncertainty (Beven, 2001). More parsimonious modeling tools are achieved by seeking effective process representation through implicitly describing the modelled domain, which generally comes at the cost of sacrificing exhaustive spatial description and, to a certain extent, losing parameter interpretability.

At the catchment scale, model parameters are assumed to ideally represent the specificities of the catchment, such as the morphology and land-use characteristics. When discharge measurements are available, i.e., gauged mode, model parameters are estimated with the

objective of reproducing as likely as possible the observed discharge time series (Klemeš, 1986). However, several situations involve unknown discharge measurements, i.e., ungauged mode, such as assessment of future land-use change, land surface model parameterization (Beven, 2012), or predicting flows at ungauged locations (Sivapalan *et al.*, 2003). These problems are referred to as *regionalization problems*.

Major scientific initiatives were carried out in order to better estimate hydrological model parameters and replace numerical calibration by better understanding of the hydrological processes (Beven, 2012). The Model Parameters Estimation Experiment (MOPEX) project (Duan *et al.*, 2006) and the Prediction at Ungauged Basins (PUB) initiative by the International Association of Hydrological Sciences (IAHS; Sivapalan *et al.*, 2003; Hrachowitz *et al.*, 2013) have succeeded in approaching the problem. They have proposed and improved techniques that: (1) relate model parameters to climate/landscape attributes (for they are supposed to detain first-order influence in shaping the represented hydrological phenomena), so as to obtain what can be referred to as *functional relationships* (Seibert, 1999); and (2) permit the transfer of the implicit information contained in the parameter values from one catchment to another under proximity assumptions. The former aims at *a priori* estimating the model parameters through regression-established relationships (Anderson, Koren, and Reed, 2006; Boughton and Chiew, 2007; Drogue *et al.*, 2002; Hundecha, Ouara, and Bárdossy, 2008; Merz and Blöschl, 2004; Oudin *et al.*, 2008; Oudin *et al.*, 2006b; Samaniego, Kumar, and Attinger, 2010; Sefton and Howarth, 1998; Seibert, 1999). The latter aims at efficiently transferring the estimated parameters from gauged locations (i.e., donors) to ungauged locations (i.e., receivers; Kim and Kaluarachchi, 2008; Vandewiele and Elias, 1995; Oudin *et al.*, 2010; Oudin *et al.*, 2008; Parajka, Merz, and Blöschl, 2005). Since the transferred parameters are only numerically optimal and might not be hydrologically optimal (for several reasons, as discussed by Andréassian *et al.*, 2012), other techniques seek to improve identification of the transferred parameters and reduce predictive uncertainty such as regional calibration (Bourgin *et al.*, 2015; Fernandez, Vogel, and Sankarasubramanian, 2000; Castiglioni *et al.*, 2010; Parajka, Blöschl, and Merz, 2007), where groups of sets of parameters are calibrated simultaneously with the functional relationships.

In terms of performance, none of the proposed techniques prevailed over ordinary numerical calibration exercises. The reasons why calibration outperforms regionalization approaches are manifold: It is difficult to determine a relevant similarity measure between the donors and the eventual receivers, while in the case of the *a priori* parameter estimation techniques, a high predictive uncertainty persists. This can be constrained by providing further knowledge about the flow regime in the process (Olden and Poff, 2003; Yadav, Wagener, and Gupta, 2007; Zhang *et al.*, 2008). Also, the nonlinear multidimensional aspect of such problems could be handled by machine-learning (ML) techniques, which have achieved successful applications in several scientific domains (Shen, 2018; Tyrallis, Papacharalampous, and Langousis, 2019).

5.1.2 Random forest: a potentially useful tool for regionalization

The regionalization problem is viewed as a question of how we can predict streamflow in ungauged situations (in space or in time). It can also be conceived as how we can discover the hidden information in the parameters for a better understanding of the multiple factors that shape the catchment response. Unveiling the hidden information in the parameters requires a large number of situations (Merz, Blöschl, and Parajka, 2006b) and *ad hoc* techniques that can efficiently digest a huge number of data. In this sense, ML tools have demonstrated an attractive performance in absorbing and exploring information from a large number of data (Shen, 2018). These tools have gained interest within the hydrological community for various purposes: building data-driven models, supporting the numerical models by relating their parameters to observed data, or replacing certain complicated numerical models by simpler ones without losing prediction performance (Carbajal and Bellos, 2018). Regression tree ensembles such as the RF algorithm (Breiman, 2001a) can be used to represent nonlinear relationships (*unknown target functions* in the ML jargon) between a dependent variable (e.g., a model parameter or a hydrological signature) and multiple independent attributes (e.g., catchment area, topographic index, or mean precipitation). Although the RF algorithm has demonstrated its outperformance in comparison with other ML algorithms, and despite its application to many problems in several environmental sciences (Cutler *et al.*, 2018; Gislason, Benediktsson, and Sveinsson, 2006; Liakos *et al.*, 2018), its use in water sciences is still limited (Tyrallis, Papacharalampous, and Langousis, 2019). Examples of RF applications in hydrology include precipitation downscaling (Diez-Sierra and Jesus, 2019; He *et al.*, 2016), flood prediction and risk assessment (Muñoz *et al.*, 2018; Sultana *et al.*, 2018; Wang *et al.*, 2015), estimating runoff modes or hydrological signatures on a continental scale and predicting flow regimes (Addor *et al.*, 2018; Booker and Woods, 2014; Buchanan *et al.*, 2018; Carlisle *et al.*, 2010) as well as predicting flow characteristics at ungauged locations (Brunner, Seibert, and Favre, 2018; Prieto *et al.*, 2019; Zhang *et al.*, 2018). RFs are constructed by growing a number of regression and classification trees. In this manner, they acquire the ability to reduce overfitting risk and allow for the classification of the predictors in terms of their importance in determining the dependent variable.

5.1.3 Application of RF for model regionalization in multiple land-use environments

The catchment land use is a major driver for how it will respond to a specific event. Understanding how the interactions between different land-use types reshape the response of hydrological catchments is a central issue. For this matter, regionalization with RF can be a handy technique with which to: (1) relate certain well-chosen model parameters to one or more land-use

descriptors (Boulesteix *et al.*, 2012; Tyralis, Papacharalampous, and Langousis, 2019); (2) reveal the relative weight of a specific land-use type vis-à-vis the other attributes; and then (3) make way for impact assessment (or prediction) of past (or future) land-use change scenarios. In particular, urbanization—as a land-use change—has been a widely investigated subject owing to the serious effects that it has on the catchment water cycle. Toward this aim, attempts at expressing model parameters in terms of urbanization measures exist in the literature. For example, Cheng, Lee, and Lee (2010) and Chen, Chuang, and Cheng (2014) aimed to relate Nash model parameters to basic urban indicators, such as imperviousness and population. Kjeldsen, Miller, and Packman (2013) related the four parameters of a basically nonurban event-based and conceptual model to the urbanized proportion of seven catchments in the United Kingdom. Except for a few studies (Oudin *et al.*, 2018; Salavati *et al.*, 2016), where a significant number of catchments were used, such applications were constrained to small samples of catchments. This draws into question the validity of such relationships in a continuous framework and the generalizability of such an approach for a wide range of catchments.

5.1.4 Context and scope of the study

This study is in continuity of a number of regionalization works conducted principally over France using the same model structure. Oudin *et al.* (2006b) have investigated the correlations between the GR4J (Perrin, Michel, and Andréassian, 2003) calibrated parameters and the catchment land-use characteristics using a large sample of 221 French catchments. The catchments sample was enlarged in a later work by Oudin *et al.* (2008), which aimed at comparing different regionalization approaches (regression, spatial proximity, and physical resemblance). Using French and British catchments, Oudin *et al.* (2010) have investigated whether the physical similarity approach is equivalent to hydrological similarity. Finally, Bourgin *et al.* (2015) have dealt with transferring uncertainty from gauged to ungauged catchments.

Another aspect of this study is investigating how urbanization reshapes the catchment water cycle. In this sense, Salavati *et al.* (2016) and Oudin *et al.* (2018) have tried to link the urbanization-induced change in the hydrological behavior of the catchment with some of the catchment landscape metrics. They have presented an approach that combines statistical analysis and hydrological modeling besides the use of the paired-catchments approach, where spatial proximity is involved.

The novelties of the present study can be summarized in the following points: (1) the application of the same model structure but at a finer time step (hourly instead of daily time steps) and (2) the use of a larger set of 2225 catchments containing urban and rural cases, located in France and the United States and (3) the application of a ML algorithm, which is the main

novelty of this study. In contrast to Oudin *et al.* (2018), only one urbanization metric was used herein. Our main objective was to explore how RF can contribute to relating calibrated hourly conceptual model parameters to certain catchment attributes including an urbanization measure. To this aim, the following questions were addressed: (1) To what extent does RF help in estimating certain calibrated model parameters knowing specific catchment climate and landscape descriptors? (2) What benefits can RF yield in terms of estimating calibrated model parameters using a very large dataset? (3) Is the RF sensitive to the urbanization measure used? Does it help predict the change in parameter values by knowing the change in urbanization stage?

5.2 Data

5.2.1 Sample selection

The sample of 2225 catchments was selected with respect to having limited snow effect, limited influence of dams and of storage facilities, and sufficiently available data, namely, at least 8 years with no gaps in hourly precipitations and potential evapotranspiration data and a rate of less than 30% hourly discharge gaps. Of the 2225 selected catchments, 870 are located in the United States (Figure 5.1) and the remaining 1355 are located in France (Figure 5.2). Their sizes range from 0.5 km² to 110,000 km², with a median value of 222 km². Characterization in terms of urbanization was carried out using the Catchment Percent Developed (*CPD*) measure (Oudin *et al.*, 2018), which informs about the fraction of the catchment that is occupied by an urban landscape. Hence, *CPD* varies from 0% for a completely rural catchment to 100% for a completely urbanized catchment.

The U.S. catchments sample was selected among 9322 GAGES-II stream gauges maintained by the U.S. Geological Survey (USGS) (Falcone, 2011). Instantaneous flow measurement extraction was done by using the `dataRetrieval` package (Cicco *et al.*, 2018) in an R development environment (R Core Team, 2019). The availability of hourly hydroclimatic time series was checked between 2002 and 2017. We used the National Land Cover Database (NLCD), available for the years 2001, 2006, and 2011 (Homer *et al.*, 2007; Homer *et al.*, 2015; Homer *et al.*, 2004) to compute *CPD* as the percentage of urban pixels in the catchments, i.e., sum of classes 21 (developed: open space), 22 (developed: low intensity), 23 (developed: medium intensity), and 24 (developed: high intensity).

The French catchments sample was selected from the Banque HYDRO database (<http://www.hydro.eaufrance.fr>; Leleu *et al.*, 2014), where flow measurements are available for over 4000 stations across France. The availability of hourly streamflow and precipitation

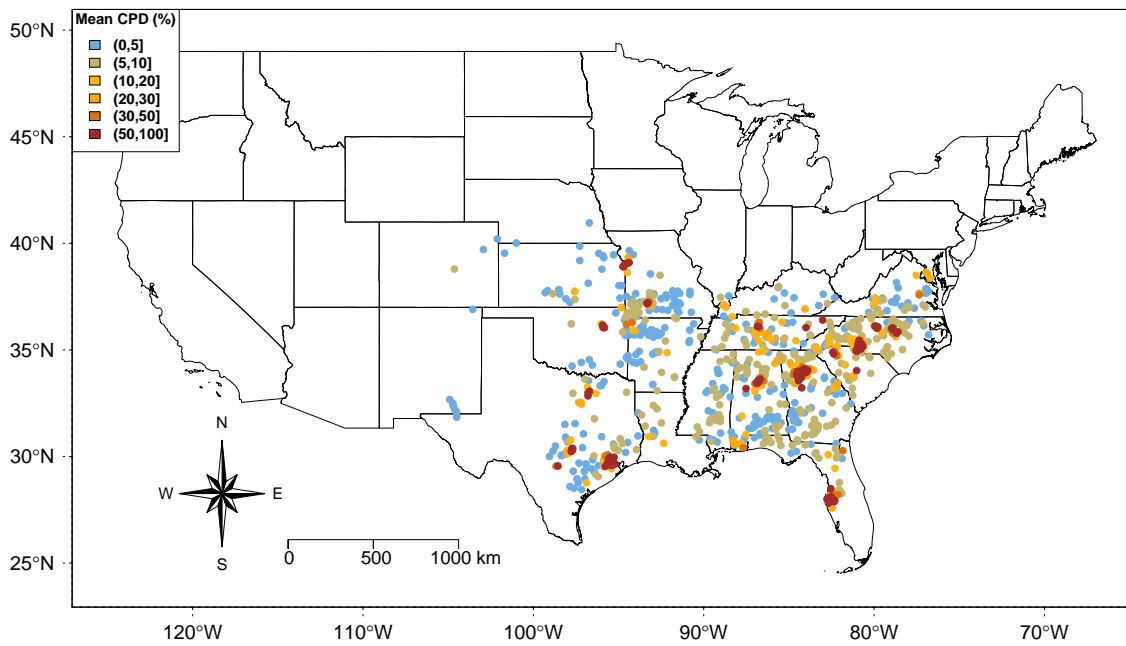


Figure 5.1. The location and mean Catchment Percent Developed (CPD) of the 870 U.S. catchments.

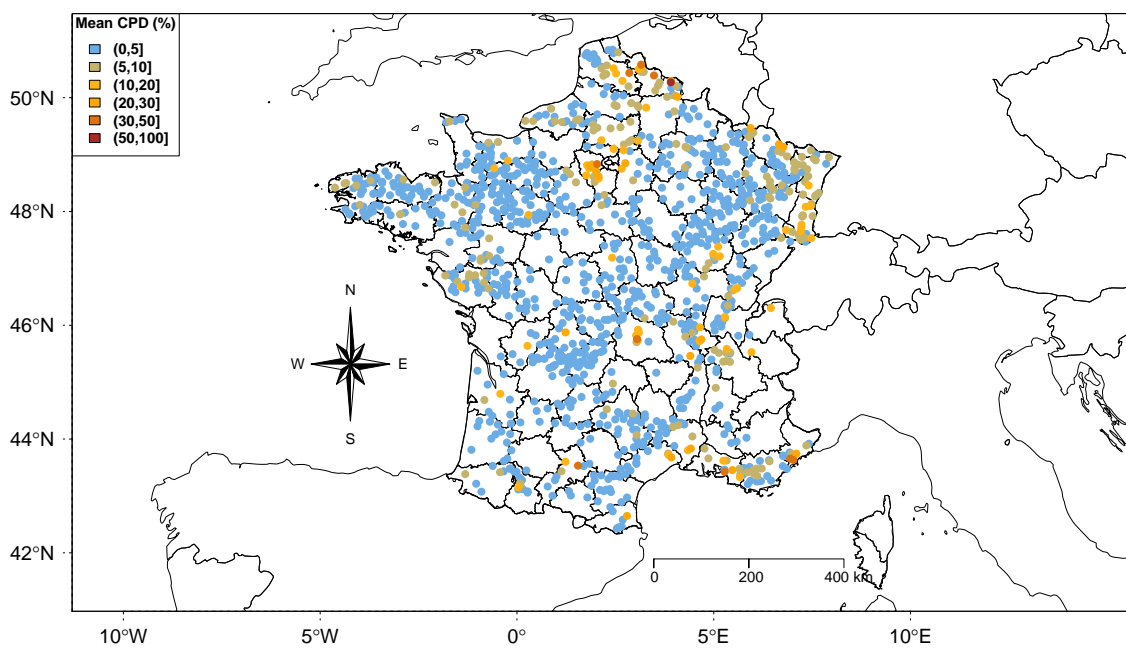


Figure 5.2. The location and mean Catchment Percent Developed (CPD) of the 1355 French catchments.

data was then checked for the period 1997-2012. *CPD* was computed using the CORINE Land Cover (CLC) product of the Copernicus Land Monitoring Service available for the years 1990, 2000, 2006, and 2012 (Büttner *et al.*, 2012). The classes considered to be urban were classes 11 (containing 111: continuous urban fabric, 112: discontinuous urban fabric) and 12 (containing 121: industrial or commercial units, 122: road and rail networks and associated land, 123: port areas, and 124: airports).

5.2.2 Catchment descriptors

The catchment's response is a function of two major drivers: climatic forcing and the landscape characteristics of the catchment. As summarized by Poncelet (2016), the most relevant catchment descriptors can be binned into four categories:

1. **Climate:** The catchment's response inherits most of its variability from the catchment's climate (Budyko, 1974). Many climate characteristics were computed over each catchment's recorded period in order to limit their dependency on the recorded period. As climate descriptors, we considered mean hourly precipitation P (mm/h), mean hourly potential evapotranspiration PE (mm/h), humidity index HI (–), and flashiness of precipitations FP (–).
2. **Morphology:** The catchment's morphology is essential in predicting the catchment's response timing and the repartition of precipitations into infiltration and runoff. For this reason, we used the catchment drained area A (km²), drainage density DD (km/km²), and the median compound topographic index CTI (–) as morphological descriptors.
3. **Land use:** The catchment's water yield and evapotranspiration losses depend on the catchment land use. Also, it is in our case of a central interest as we are dealing with the catchment's level of urbanization. Thus, three land-use metrics were assessed: the *CPD* (%), the fraction of forest $fFOR$ (%), and the fraction of open water fW (%).
4. **Geopedology:** The catchment's water transfers to and from the subjacent aquifers are modulated by the catchment's geological and pedological characteristics. Hence, mean porosity $POROSITY$ (–), mean of log-transformed values of intrinsic permeability PER (m²), mean soil and subsoil content of gravel M_GRAVEL (%), silt M_SILT (%), and clay M_CLAY (%) were considered as geopedological characteristics.

Note that some catchment descriptors were preliminary considered but found to be highly correlated with the aforementioned descriptors, such as flashiness of potential evapotranspiration (highly correlated with mean potential evapotranspiration), mean slope and mean elevation (highly anti-correlated with the compound topographic index), and mean soil and subsoil

content of sand (highly anti-correlated with mean content of silt and clay). To avoid redundancy within each category of descriptors, we kept only the variables that did not exhibit high or significant correlation values with each other. This would prevent having an influenced variable importance by the presence of highly correlated variables. Table 5.1 provides a complete list of the descriptors, the computation method, and the data sources.

Table 5.1. Description of the different climatic, topographic, land-use, and geopedological characteristics estimated for each catchment.

Notation	Index name	Computation	Unit	Data source
P_m	Mean hourly precipitation	Total depth of precipitations over the recorded period (8-16 years) divided by the number of hours, aggregated spatially to the catchment scale	mm/h	COMEPHORE product of Météo France, 1-km resolution (Tabary <i>et al.</i> , 2013) and NEXRAD Stage IV dataset, 4-km resolution, extracted using the <code>geoknife</code> R package (Hardegree <i>et al.</i> , 2008; Horvat <i>et al.</i> , 2011; Read <i>et al.</i> , 2015)
PE_m	Mean hourly potential evapotranspiration	Total depth of potential evapotranspiration over the recorded period (8-16 years) divided by the number of hours, aggregated spatially to the catchment scale	mm/h	Evaluated using temperature-based formula (Oudin <i>et al.</i> , 2005). Daily temperature was extracted from SAFRAN product of Météo France, 8-km resolution (Vidal <i>et al.</i> , 2010) and Daymet dataset, 1-km resolution (Thornton <i>et al.</i> , 2016)
HI	Humidity index	$HI = \frac{P_m}{PE_m}$	—	P_m and PE_m data sources
FP	Flashiness of precipitation	$FP = \frac{\sum_i P_i - P_{i-1} }{\sum_i P_i}$, with P_i the precipitation depth (mm) at hour i (Baker <i>et al.</i> , 2004; Holko <i>et al.</i> , 2011)	—	P_i data source
A	Catchment area	—	km ²	(Bourgin <i>et al.</i> , 2010; Falcone, 2011)

Continued on next page...

Table 5.1 continued: Description of the different climatic, topographic, land-use, and geopedological characteristics estimated for each catchment.

Notation	Index name	Computation	Unit	Data source
<i>DD</i>	Drainage density	$DD = \frac{\sum_i L_i}{A}$, with L_i length of stream i (km) and A the catchment area (km ²)	km/km ²	The hydrographic networks were extracted from the BD Carthage® dataset (France) and the National Hydrography Dataset NHD (USA) using the FedData R Package (Bocinsky, Beaudette, and Chamberlain, 2019)
<i>CTI</i>	Median compound topographic index	$CTI = \text{median} \left(\log \left(\frac{A_{s,i}}{\tan(\beta_i)} \right) \right)$, with $A_{s,i}$ the i^{th} cell's specific area and β_i its slope angle	—	(Verdin, 2017)
<i>CPD</i>	Catchment percent developed	Sum of the pixels attributed to urbanization classes divided by the total number of pixels	%	NLCD 2001, 2006, and 2011 (USA) and CLC 1990, 2000, 2006, and 2012 (France)
<i>fW</i>	Fraction of open water	Sum of pixels occupied by open water class divided by the total number of pixels	%	NLCD 2001, 2006, and 2011 (USA) and CLC 1990, 2000, 2006, and 2012 (France)
<i>fFOR</i>	Fraction of forest	Sum of pixels occupied by forest classes divided by the total number of pixels	%	NLCD 2001, 2006, and 2011 (USA) and CLC 1990, 2000, 2006, and 2012 (France)
<i>POROSITY</i>	Mean porosity of the catchment's soil and subsoil geologic units	Volume of voids divided by the total volume	—	GLobal HYdrogeology MaPS (GLHYMPS; Gleeson <i>et al.</i> , 2014)

Continued on next page...

Table 5.1 continued: Description of the different climatic, topographic, land-use, and geopedological characteristics estimated for each catchment.

Notation	Index name	Computation	Unit	Data source
<i>PER</i>	Mean of logarithm values of soil and subsoil permeability	—	log(m ²)	GLobal HYdrogeology MaPS (GLHYMPS; Gleeson <i>et al.</i> , 2014)
<i>M_GRAVEL</i>	Mean gravel content of soil and subsoil geologic units	—	%	Harmonized World Soil Database (HWSD, Version 1.2; FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012; LeBauer, 2019)
<i>M_SILT</i>	Mean silt content of soil and subsoil geologic units	—	%	Harmonized World Soil Database (HWSD, Version 1.2; FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012; LeBauer, 2019)
<i>M_CLAY</i>	Mean clay content of soil and subsoil geologic units	—	%	Harmonized World Soil Database (HWSD, Version 1.2; FAO/IIASA/ISRIC/ISS-CAS/JRC, 2012; LeBauer, 2019)

End of Table 5.1.

5.3 Methods

5.3.1 Model parameters and calibration

Representing the processes in an urbanized catchment requires working at sub-daily time steps since it is the timescale at which most of the urban hydrological processes are taking place (Fletcher, Andrieu, and Hamel, 2013; Salvadore, Bronders, and Batelaan, 2015). Unfortunately, the fineness of the time step is limited by the availability of the hydroclimatic data that

are needed for model forcing and validation. According to these two requirements, the hourly time step was chosen.

In this study, we used the GR4H lumped, conceptual hourly model (Le Moine, 2008; Mathevet, 2005) modified by Ficchi (2017) to account for interception at sub-daily time steps. It was developed using large and worldwide samples of hydrological catchments over which it has achieved good performances in comparison with many other models (Mathevet, 2005; Esse *et al.*, 2013). In addition, its parsimonious structure helps understanding each parameter's role in simulating the catchment's response.

GR4H can be categorized as a soil moisture accounting model (Perrin, Michel, and Andréassian, 2003), which uses reservoir-style components and unit hydrographs to represent the following hydrological processes: interception, actual evapotranspiration, percolation from the soil, surface water–groundwater exchanges, quick and slow runoff. Figure 5.3 represents the model structure.

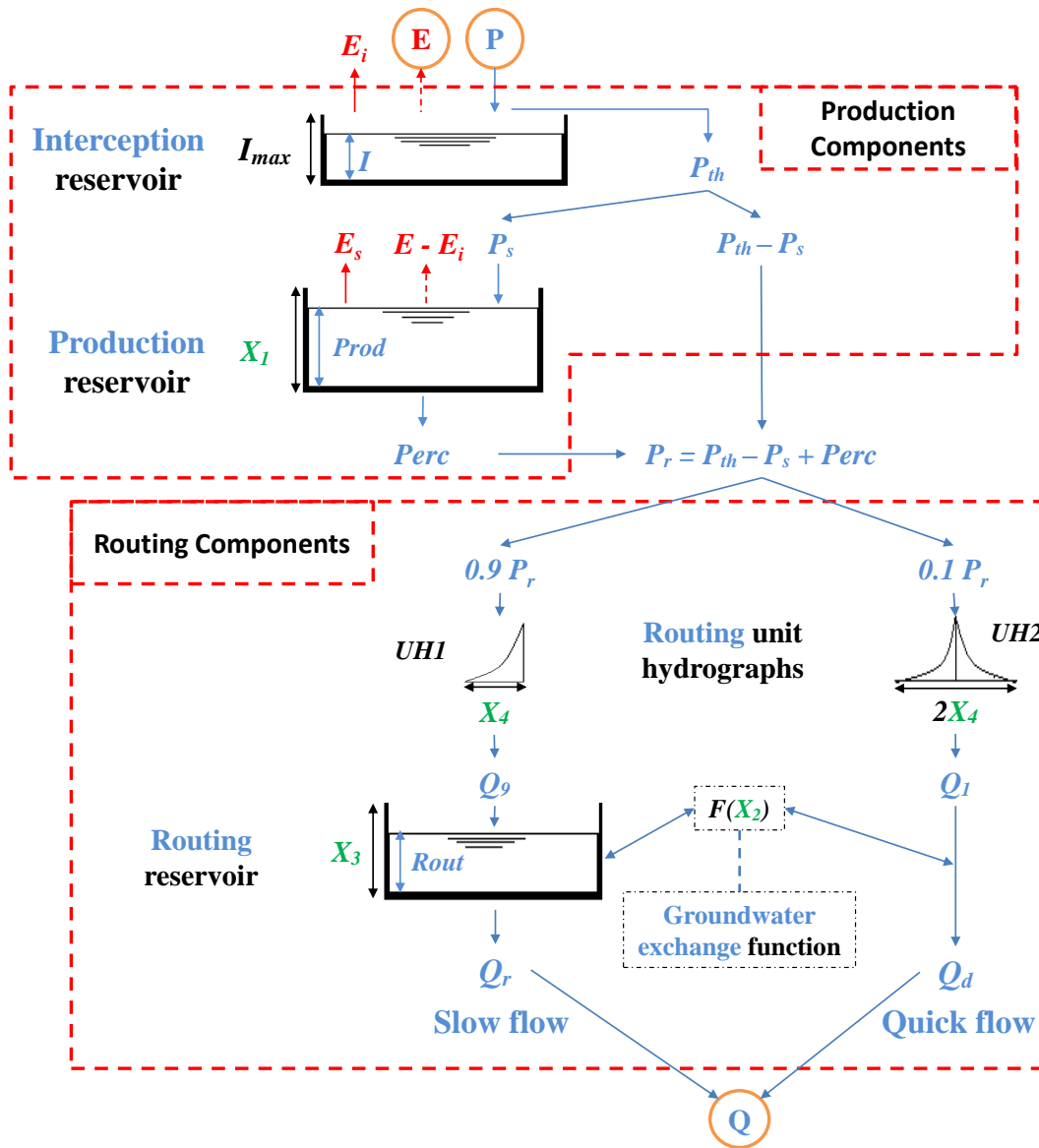


Figure 5.3. GR4H model structure (Ficchi, 2017). The water fluxes are specified in blue and red. The four free model parameters X_1 to X_4 are in green.

The model contains four free parameters. X_1 (mm) is the production store maximum capacity. It controls the amount of water that can be stored and then lost either through actual evapotranspiration or percolation. X_2 (mm/h) is the groundwater–surface water exchange rate, which is positive when importing water to the river stream and negative otherwise. X_3 (mm) is the nonlinear routing store maximum capacity. It plays a role in determining the low-frequency component of the simulated hydrograph. X_4 (h) controls the routing unit hydrograph time base. It represents the characteristic time of the catchment response. Ficchi (2017) and Perrin, Michel, and Andréassian (2003) give a more detailed description of the model equations.

The first year of the recorded period for each catchment was used for warm-up, and the model

was then calibrated on the remaining years of the recorded period. The calibration procedure ran a local search algorithm including a steepest descent variable, as implemented by Edijatno *et al.* (1999). In order to account for possible dependency on the chosen objective function, the model calibration and evaluation were performed with respect to four different metrics: the Kling-Gupta efficiency (*KGE*, Gupta *et al.*, 2009) and the Nash-Sutcliffe efficiency (*NSE*, Nash and Sutcliffe, 1970) scores both applied on non-transformed (i.e., *KGE* and *NSE*) and square root values of observed and simulated runoffs (i.e., *KGESR* and *NSESR*). This was carried out using the `airGR` R package (Coron *et al.*, 2017) in an R development environment (R Core Team, 2019).

5.3.2 Estimating the model parameters at ungauged locations using RF

Following recent studies (Addor *et al.*, 2018; Zhang *et al.*, 2018), we used the RF algorithm (Breiman, 2001a) for classification and prediction implemented in the `randomForest` R package (Liaw and Wiener, 2002) in an R development environment (R Core Team, 2019). RF is a binning supervised ML algorithm that allows a large number of attributes (called *predictors* or *independent variables*) to be classified and used in order to carry out estimations of a response variable (called *target variable* or *dependent variable*) based on a large number of classification and regression trees (CARTs). As advocated by Breiman and Cutler (2019) and presented by Tyrallis, Papacharalampous, and Langousis (2019) in a recent review of literature, RF has the advantages of accuracy (Biau and Scornet, 2016; Díaz-Uriarte and Alvarez de Andrés, 2006), efficiency in large databases (Genuer *et al.*, 2017), estimation of variable importance (Biau and Scornet, 2016; Díaz-Uriarte and Alvarez de Andrés, 2006; Ziegler and König, 2014), and avoidance of overfitting (Díaz-Uriarte and Alvarez de Andrés, 2006).

In this study, RF allowed to represent the multidimensional nonlinear relationships between each model parameter and the 15 catchment descriptors (Table 5.1). First, we divided the sample into two subsamples (Figure 5.4): the first one (120 catchments) served for validation of the regionalized model performance and was not used to train the RFs; it contained only urban catchments ($CPD \geq 20\%$). The second one (2105 catchments) was used to construct two RFs, i.e., as training samples: one with the whole subsample (*RF_ALL*, i.e., with rural and urban catchments) and another with only the urban catchments (*RF_URB*, i.e., with only 119 catchments that had a *CPD* over 20%). The idea was to see whether the RF can efficiently handle the large amount of training samples and produce equivalent model performances using both configurations. Second, the constructed RF was used to estimate model parameters $\{X_i\}$ for the remaining 120 urban catchments (treated as ungauged) using their descriptors $\{D_j\}$. RF construction requires a specification about the number of trees to be grown (here $N_{tree} = 500$) and the number of variables to be selected randomly at each tree growing (here *mtry* was fixed at 5 by using the `tuneRF` function in the `randomForest` R package). While

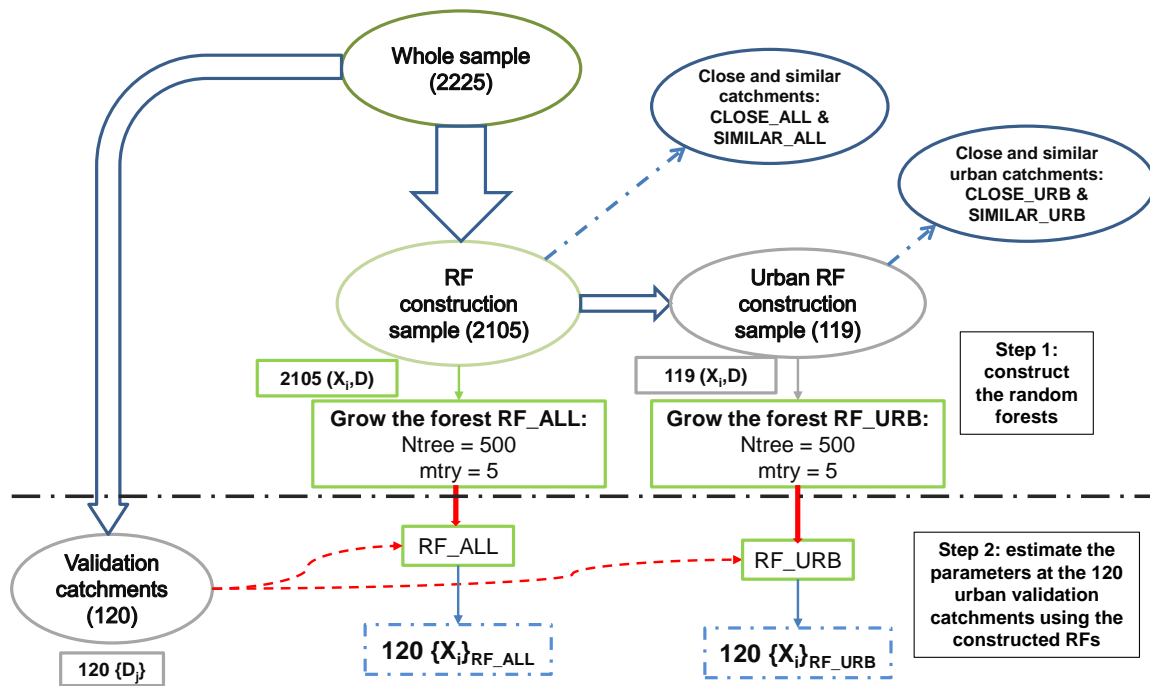


Figure 5.4. Random forest (RF) construction using the 2105 rural and urban catchments. Validation is carried on the remaining set of 120 urban catchments (i.e., with $CPD \geq 20\%$).

growing a tree, about one third of the cases—called out-of-bag (OOB) data—are left out of the sample (Biau and Scornet, 2016; Breiman, 2001a), and then used to compute errors at each grown tree, e.g., the mean squared errors (MSE).

During the process of RF growing, we considered the $IncMSE$ metric to evaluate each descriptor importance in the RF-established relationships (Addor *et al.*, 2018). For a descriptor j (in our case j is between 1 and 15), the values were randomly permuted for the OOB data and a corresponding RF MSE_j was determined. If MSE_0 is the MSE of the RF computed during the construction, i.e., with non-permuted values for the OOB data, then $IncMSE$ represents the difference between MSE_0 and MSE_j , after being computed and then scaled by the standard deviation of the differences over the trees. The greater the $IncMSE$, the more important the predictor variable.

5.3.3 Benchmark regionalization techniques

The validation sample (120 urban catchments) was used to set a benchmark for regionalization methods. For each catchment belonging to this sample, the recorded period was split into two subperiods P1 and P2. For each score SC (KGE , NSE , $KGESR$, and $NSESR$), calibration in P1 with respect to SC yielded a set of parameters to be controlled in P2 using the same

evaluation criterion. We compared the control performances in P2 of this calibrated set with the performances of:

- The RF-estimated parameters using the catchment descriptors.
- The transferred parameters from the closest neighbor catchment. Spatial closeness was computed by weighting the distances between the centroids (80%) and the outlets (20%) of the catchments (Lebecherel, Andréassian, and Perrin, 2013). Close catchments were selected either from the whole 2105 catchments used to train the *RF_ALL* (*CLOSE_ALL*, Figure 5.4) or from the 119 urban catchments used to grow *RF_URB* (*CLOSE_URB*).
- The transferred parameters from the most similar catchment with respect to the descriptors used to construct the RF. For each descriptor, the catchment ranks were determined. Then, the Euclidean distance between ranks was computed in the hyperspace of descriptors (Oudin *et al.*, 2008). Similar catchments were selected either from the whole 2105 catchments used to construct the *RF_ALL* (*SIMILAR_ALL*, Figure 5.4) or from the 119 urban catchments used to construct *RF_URB* (*SIMILAR_URB*).

5.4 Results

5.4.1 Model performances and estimated parameters

As explained before, two RFs were constructed in order to represent the relationships between calibrated model parameters and catchment descriptors. These calibrated model parameters were estimated over the whole period. They corresponded to relatively satisfactory scores at the hourly time step (not shown here). The calibration median scores for the 2105 catchments were about 0.9 in terms of *KGE* and *KGESR*, 0.81 and 0.83 in terms of *NSE* and *NSESR*, respectively. For the 119 urban catchments, a drop of performances was noticed, especially for median *NSE* (0.81 to 0.71), *NSESR* (0.83 to 0.77) and *KGE* (0.89 to 0.83), while the performances remained equivalent with regard to *KGESR* (0.9 to 0.87).

For every validation catchment, six sets of parameters were tested on the second period P2: Two sets were estimated from the RFs (*RF_ALL* and *RF_URB*) and four sets were transferred from the spatially close catchments (*CLOSE_ALL* and *CLOSE_URB*) and the similar catchments (*SIMILAR_ALL* and *SIMILAR_URB*). Their performances were then compared with the control scores of a calibrated set of parameters from P1. Figure 5.5 shows the distribution of the performances evaluated using each of the aforementioned metrics.

The RF-estimated parameters slightly but significantly (different distributions according to

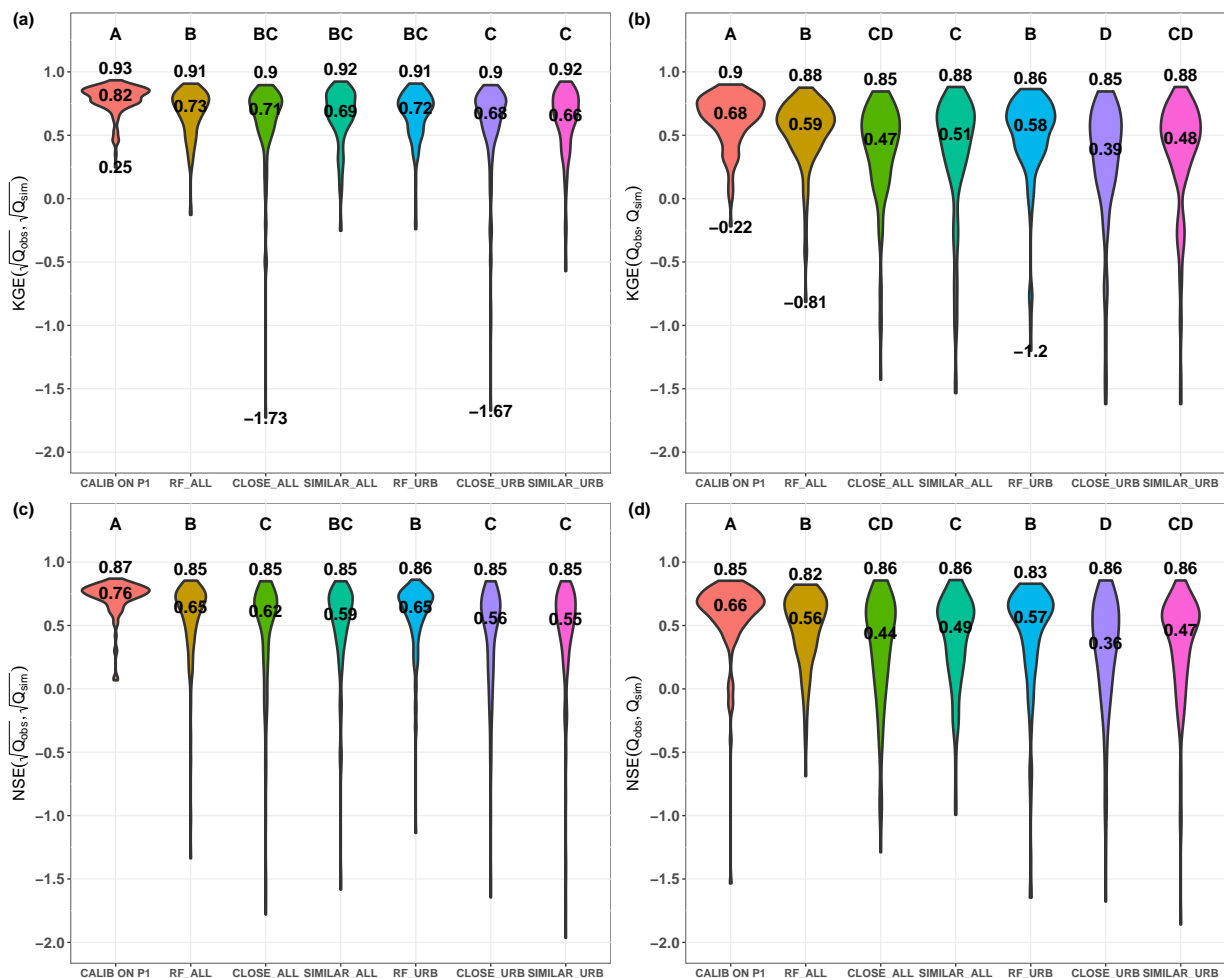


Figure 5.5. Validation scores on P2 in terms of (a) *KGESR* (b) Kling-Gupta efficiency (*KGE*) (c) *NSESR* and (d) Nash-Sutcliffe efficiency (*NSE*) using calibrated parameters on P1 (*CALIB ON P1*), estimated parameters using random forest (*RF_ALL* and *RF_URB*), transferred parameters from the closest catchments (*CLOSE_ALL* and *CLOSE_URB*) and from the most similar catchments (*SIMILAR_ALL* and *SIMILAR_URB*). The values indicate the minimum, the median, and the maximum scores. Some minimum values were not shown as they were less than -2.0 for the considered metric of evaluation. The letters indicate the statistical equality at 10% risk between the performances, estimated using the Mann-Whitney-Wilcoxon test.

Table 5.2. R^2 computed between the calibrated parameters with respect to different objective functions (*KGESR*, *KGE*, *NSESR*, and *NSE*) and the estimated parameters via RF, transferred parameters from the close and from the similar catchments for the 120-urban-catchments sample. Statistical significance of the correlation is indicated by asterisks, where * is significant at 5%, ** at 1%, and *** at 0.1% risk.

Parameter	Objective function	Using the 2105 catchments			Using the 119 urban catchments		
		RF	CLOSE	SIMILAR	RF	CLOSE	SIMILAR
X_1 (mm)	<i>KGESR</i>	0.476***	0.361***	0.367***	0.448***	0.367***	0.330***
	<i>KGE</i>	0.144***	0.077**	0.066**	0.080**	0.047*	0.085**
	<i>NSESR</i>	0.530***	0.353***	0.332***	0.434***	0.301***	0.250***
	<i>NSE</i>	0.152***	0.084**	0.066**	0.111***	0.034*	0.069**
X_2 (mm/h)	<i>KGESR</i>	0.022	0.037*	0.062**	0.009	0.014	0.064**
	<i>KGE</i>	0.054*	0.053*	0.098***	0.064**	0.011	0.045*
	<i>NSESR</i>	0.109***	0.026	0.056**	0.064**	0.023	0.002
	<i>NSE</i>	0.110***	0.085**	0.047*	0.067**	0.105***	0.036*
X_3 (mm)	<i>KGESR</i>	0.449***	0.275***	0.213***	0.346***	0.194***	0.173***
	<i>KGE</i>	0.222***	0.147***	0.094***	0.245***	0.094***	0.069**
	<i>NSESR</i>	0.408***	0.444***	0.333***	0.442***	0.374***	0.202***
	<i>NSE</i>	0.318***	0.225***	0.277***	0.405***	0.227***	0.245***
X_4 (h)	<i>KGESR</i>	0.287***	0.082**	0.207***	0.438***	0.077**	0.207***
	<i>KGE</i>	0.301***	0.076**	0.201***	0.415***	0.09***	0.167***
	<i>NSESR</i>	0.417***	0.064**	0.355***	0.578***	0.1***	0.396***
	<i>NSE</i>	0.613***	0.121***	0.284***	0.604***	0.096***	0.262***

the statistical test, Figure 5.5) outperformed the benchmark regionalization methods based on spatial proximity or climate and landscape similarity. The differences were larger when using non-transformed flows. This was clearer when only urban catchments were considered for training; however, using only urbanized catchments for regionalization purposes reduced the model performance whatever the method considered, meaning that reducing the gauged catchment set is detrimental to regionalization, even if the reduction is designed to fit more the target ungauged catchments. Yet, the scores of time-transferred parameters (i.e., P1 to P2) remained unbeatable and significantly different (type A regardless of the considered metric).

If the performances of the RF-estimated parameters are significantly better than the transferred ones, how much do they succeed in resembling the calibrated parameters? Table 5.2 shows a comparison, by means of Spearman's rank correlation, between the calibrated parameters using the four objective functions, and the estimated parameters from RF or from the transfer according to spatial closeness or climate and landscape similarity using both the 2105-catchments sample and the 119-urban-catchments sample.

Except few cases, RF offered by far the best option in terms of parameter estimation for X_1 , X_3 , and X_4 , while for X_2 , none of the methods was satisfactory. Re-estimating the calibrated parameters with respect to transformed flows was systematically easier than when non-transformed flows were considered for calibration. The statistical test results indicated that the correlations were highly significant, even in the case of X_2 where the values were hardly above 0.1. However, the R^2 values were very moderate and in line with the values obtained in previous studies (Merz and Blöschl, 2004). These values suggest that the calibrated parameters are difficult to estimate, meaning that there are still some remaining issues concerning the descriptors used and/or concerning model structure/parametrization uncertainties. Also, the clear superiority of the RF in estimating each parameter individually did not translate to model performances, especially in the case when $KGESR$ and $NSESR$ were considered. This may be due to the relative sensitivity of the model to its four free parameters and/or the fact that RF considers each parameter independently. Indeed, an RF was constructed for each model parameter, which means that the estimation is independent from one parameter to another. This might diminish the power of possible interactions between the parameters, i.e., compensation effects, whereas these effects were kept in the transferred sets.

To illustrate the supremacy of the RF-estimated parameters, Figure 5.6 shows the simulated hydrographs for the Suwanee Creek catchment at Suwanee, Georgia (USGS code: 02334885, area: 122.1 km², mean CPD : 59.6%) between 1 March 2013 and 31 August 2013. This period was chosen as it belongs to the catchment's wettest year between 2010 and 2017 (i.e., P2). This catchment was recently investigated by Diem, Hill, and Milligan (2018) who found it to be hydrologically altered due to rapid shifts in the catchment's land cover. For this example, only the case when $KGESR$ was used is shown. Table 5.3 details the values of the five sets of parameters used to compute these hydrographs: estimated from RF_ALL and RF_URB , transferred from calibration over the catchment's first period P1, and transferred from the closest catchment (Dick Creek at Old Atlanta Road, near Suwanee, Georgia, USGS code: 02334620, area: 17.8 km², mean CPD : 55.3%) and the most similar catchment (Reedy River near Greenville, South Carolina, USGS code: 02164000, area: 125.1 km², mean CPD : 68.2%).

We can see that the transferred parameters from P1 and the RF-estimated parameters are the best ones to succeed in reproducing the observed flow. In this particular period, a heavy rainfall event was recorded during the second half of May, which attained a maximum intensity of 48.5 mm/h. The corresponding observed peak flow was slightly underestimated by the time-transferred parameters, slightly overestimated by the RF-estimated parameters but erroneously overestimated by the transferred parameters from the closest or the most similar catchment (Table 5.3).

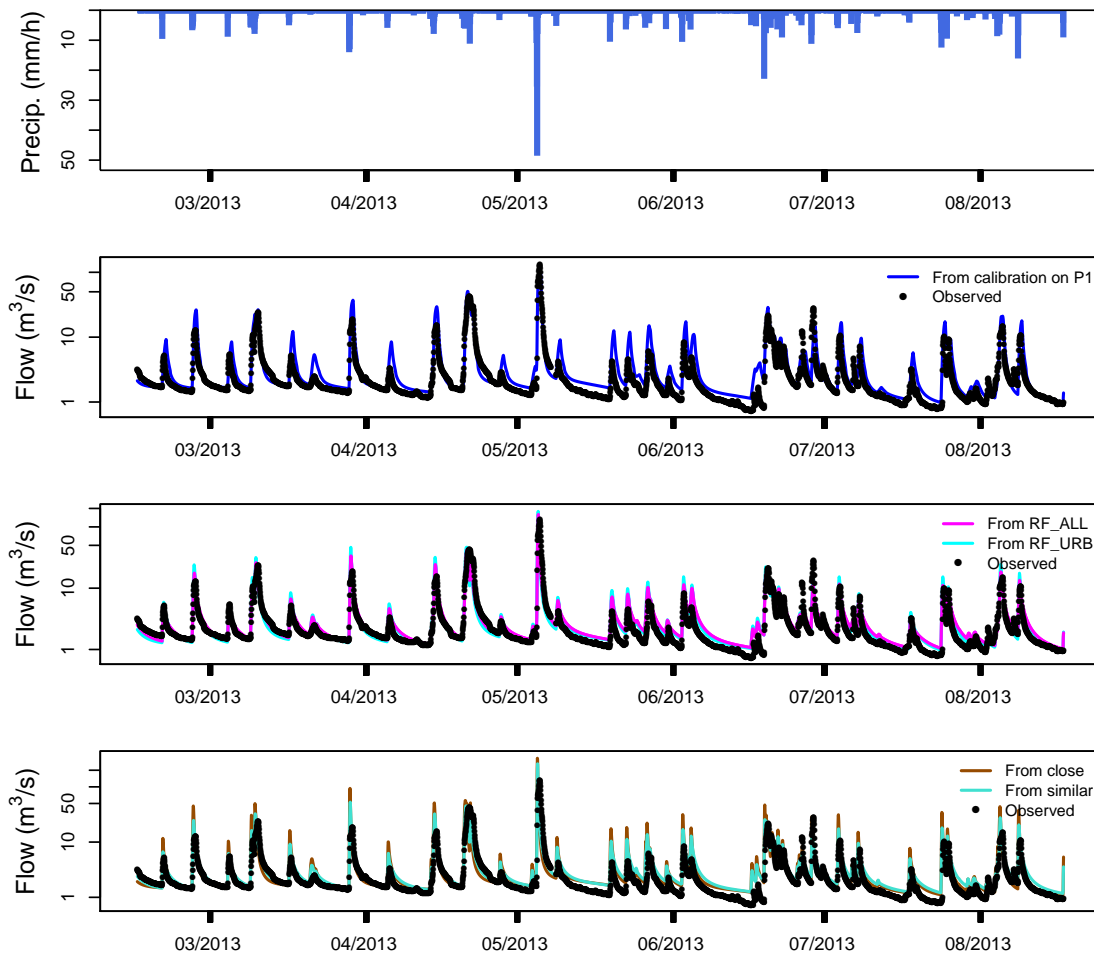


Figure 5.6. Observed hourly flow time series (in m^3/s) and simulated flow using parameters transferred from calibration over P1, estimated parameters using *RF_ALL* and *RF_URB*, and transferred parameters from the closest and the most similar catchment to the Suwanee Creek catchment (U.S. Geological Survey (USGS) code: 02334885), as response to measured precipitations (in mm/h) between 1 March and 31 August 2013. The sets of parameters correspond to *KGESR*.

5.4.2 Descriptor importance

As mentioned earlier, the importance of descriptors in determining each parameter was measured using *IncMSE*, for which the most important variables exhibited the highest scores. The objective was to see which descriptors weighed the most in the regression relationships, and how powerful the *CPD* measure was in deciding the parameter values vis-à-vis the other descriptors. As *RF_ALL* and *RF_URB* performed similarly, we retained for the analysis the *RF_ALL* approach only, since it integrates the largest amount of information and the widest panel of situations. Also, we constrained the analysis to the trained *RF_ALL* using estimated parameters with respect to *KGESR*, as the ranks were not significantly different from one

Table 5.3. Sets of parameters transferred from calibration over P1, estimated using *RF_ALL* and *RF_URB*, and transferred from the closest catchment and the most similar catchment for the Suwanee Creek catchment. **Score Whole** is the *KGESR* over the period 2 (i.e., between 2010 and 2017) and **Score Period** is computed over the 6-month period for which the resulting hydrographs are shown in Figure 5.6. $\frac{\text{MaxSim}}{\text{MaxObs}}$ is the ratio of the simulated and observed peak flows over the same 6-month period.

Source	Parameters				Score Whole	Score Period	$\frac{\text{MaxSim}}{\text{MaxObs}}$
	X ₁ (mm)	X ₂ (mm/h)	X ₃ (mm)	X ₄ (h)			
Calibration over P1	973.92	0.13	9.36	19.7	0.824	0.818	0.72
<i>RF_ALL</i>	1258.52	0.10	32.92	7.71	0.872	0.819	1.21
<i>RF_URB</i>	1250.69	0.08	22.87	7.44	0.876	0.821	1.36
Close	1269.62	0.14	13.27	3.34	0.719	0.600	2.50
Similar	1394.09	0.19	25.53	3.86	0.807	0.722	1.98

objective function to another, especially the top ranks. Figure 5.7 shows the importance of the descriptors for each parameter in terms of *IncMSE*.

Soil characteristics (mean content of gravel, silt, and clay), mean potential evapotranspiration, catchment area, and drainage density exhibited the highest importance scores in determining X_1 . Land-use characteristics, in particular *CPD*, also exhibited moderate weights in deciding X_1 . However, they were not highly decisive compared with the remaining descriptors. This seems coherent, as X_1 modulates the soil-atmosphere interactions, thus relying on a large number of descriptors. In the case of X_2 , the drainage density yielded a remarkable score, followed by mean potential evapotranspiration PE_m and some soil characteristics (mean content of clay, mean porosity, permeability) and other morphology descriptors (area and *CTI*). The influence of PE_m on X_2 may stem from the possible interaction with X_1 , as both play a role in matching the catchment water budget. The drainage density, the mean content of clay, the porosity, and *CTI* give implicit or explicit information about the soil permeability, hence, their influence over X_2 was expected as X_2 characterizes the interactions between the groundwater and the surface water. For X_3 and X_4 , the most influential descriptors were very easily distinguishable. PE_m had the heaviest weight on X_3 , the parameter that shapes the component of slow flow; this may be due to the fact that in recession periods, PE_m plays a major role in conditioning the recession flow curvature. The mean content of clay M_{CLAY} , PE_m , and the drainage density played a major role in determining X_4 . This parameter is coherent with the characteristic time of the catchment response, and drives majorly the correlation at the hourly time step between the observed and simulated flows. Besides, X_4 is correlated with the catchment area, which was revealed to be the fourth most influential descriptor.

The sole measure of urbanization was not found to be useful in determining the calibrated model parameters compared with other descriptors, where it was at best ranked as 7th out

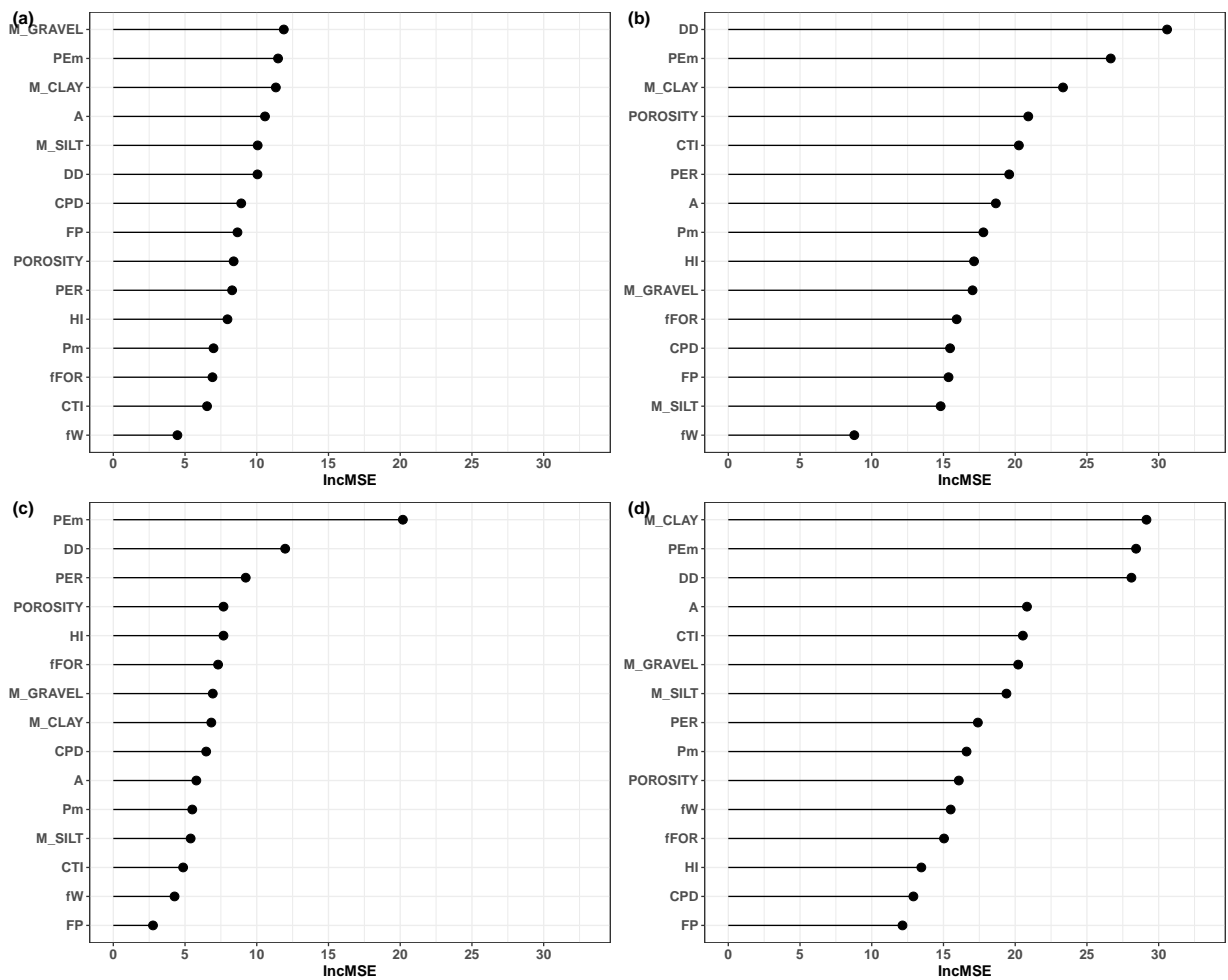


Figure 5.7. *IncMSE* scores of the catchment descriptors for (a) X_1 , (b) X_2 , (c) X_3 , and (d) X_4 . Importance scores were extracted from the *RF_ALL*.

of 15. Therefore, we studied whether the constructed RF was sensitive to this measure. We randomly selected 120 rural catchments from the 2105-catchments sample by applying a *CPD* maximum threshold of 5%. Each rural catchment was then matched with one of the 120-urban-catchments validation sample. We subsequently estimated two sets of parameters for each rural catchment: one without altering the catchment characteristics (*RF_RUR*) and the other by replacing the rural catchment *CPD* by the *CPD* of the corresponding urban catchment (*RF_UPDATED*). These two sets were then compared with the calibrated rural parameters (*CAL_RUR*), the calibrated urban parameters (*CAL_URB*), and the RF-estimated parameters of the 120-urban-catchments sample (*RF_URB*). The tests were conducted using the four objective functions, but only the results corresponding to *KGESR* are discussed hereafter. Figure 5.8 shows the distribution of the parameters.

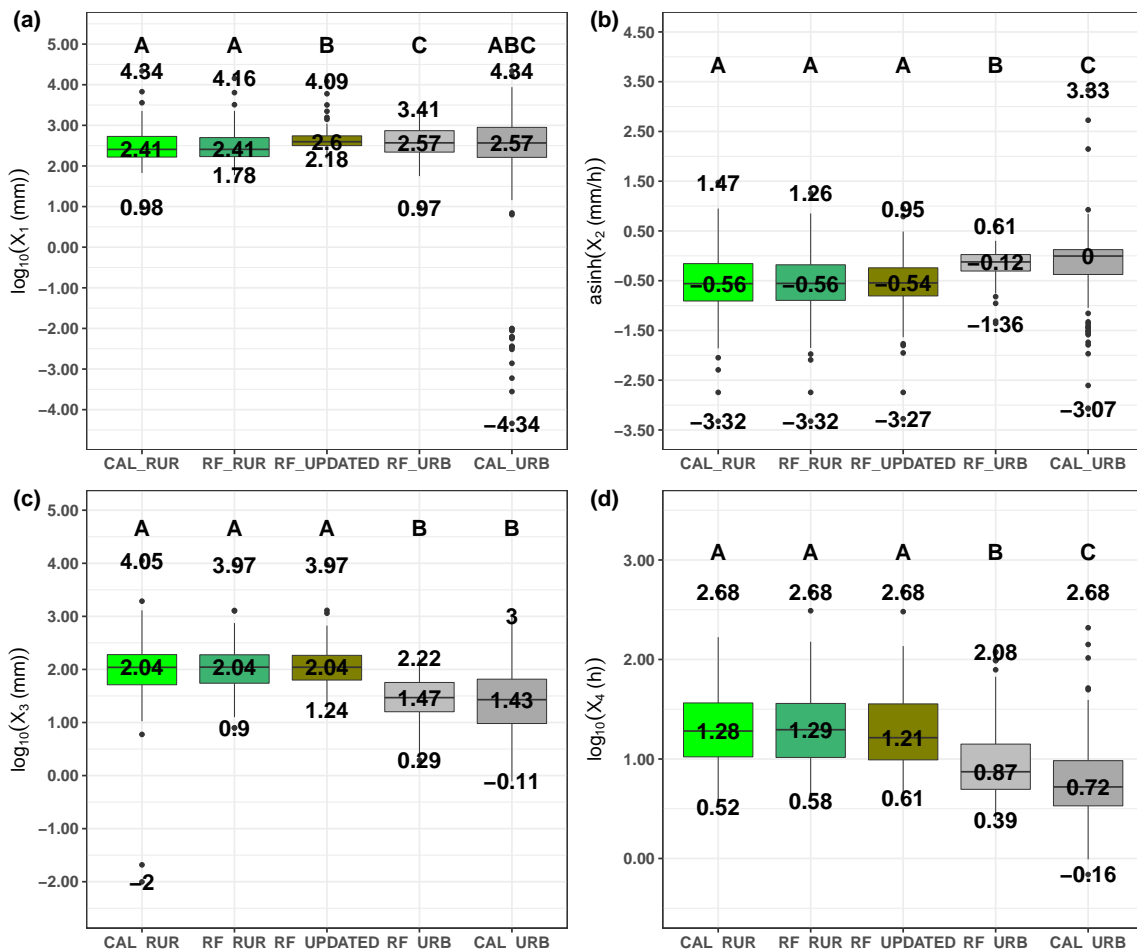


Figure 5.8. Distribution of parameters (a) X_1 , (b) X_2 , (c) X_3 , and (d) X_4 with values indicating the minimum, the median, and the maximum. *CAL_RUR* are the rural calibrated parameters (with respect to *KGESR*), *RF_RUR* are the RF-estimated rural parameters, *RF_UPDATED* are the RF-estimated parameters over the rural sample with transferred *CPD* value from the corresponding urban catchment, *RF_URB* are the RF-estimated parameters over the urban sample, and *CAL_URB* are the calibrated parameters (with respect to *KGESR*). The letters indicate statistical equality at 5% risk between the different parameter distributions, estimated using the Mann-Whitney-Wilcoxon test.

At first, the sets of calibrated parameters were different between the rural and the urban samples. The main impacts were observed on X_3 and X_4 , which represents the transfer function of the model. Their values were lower for the urban catchments, meaning that the time response of the catchments was relatively shortened. The results of the Mann-Whitney-Wilcoxon (MWW) statistical equality test indicated that there were also differences in the budget parameters (i.e., X_2).

The RF_RUR and RF_URB (RF-estimated parameters) presented similar distributions to respectively CAL_RUR and CAL_URB (calibrated parameters), particularly for parameters X_1 and X_3 , meaning that the RF approach succeeded in adapting the model parameters depending on the urbanization stage, if the training sample was adapted to the target catchments. Exceptions were made for X_2 and X_4 , where the statistical test results indicated that RF-estimated parameters for the urban catchments were different from the calibrated ones (RF_URB vs. CAL_URB). For these two parameters, CPD had the lowest ranks (12th and 14th out of 15), which means that for urban catchments with relatively low CPD values (i.e., close to 20%), the estimation of these two parameters was driven more by other descriptors than by CPD .

The shifts in CPD over the rural catchments were between +17.0% and +100.0%, with a median shift of +58.3%. However, none of the parameters responded to this shift, except for parameter X_1 where the parameters exhibited a trend toward higher values as witnessed in the urban sample (confirmed by the MWW test), which suggests an overall weak sensitivity to CPD of the RF_ALL .

5.5 Discussion and conclusions

5.5.1 Regionalization with RF: what is appreciated and what is depreciated?

In this case study, we used the RF algorithm to represent the nonlinear relationships between calibrated model parameters and catchment descriptors. Our goal was to estimate these parameters at (assumed) ungauged intensively urbanized catchments, either by giving exhaustive information about a large number of training catchments or by restraining the focus to urban catchments only. In a regionalization framework, the problem was treated via *a priori* parameter estimation. A comparison was made with transfer-based regionalization approaches. Besides, the study was performed using a large sample of catchments at hourly time steps, which represents two features of originality of this work.

With regard to the chosen spatial representation, the very heterogeneous aspect of urban

catchments' landscape requires using distributed approaches, since the spatial configuration of the rural and urban surfaces can intensify or mitigate the impact of the urban surfaces on the catchment's response (Fletcher, Andrieu, and Hamel, 2013; Mejía and Moglen, 2010; Salvadore, Bronders, and Batelaan, 2015; Singh and Woolhiser, 2002). However, it could have added more parameters in comparison with the lumped approach adopted herein, which we believe that is somehow sufficient for our regionalization objectives, by keeping a number of parameters that are more easily interpreted. Moreover, space is left for efforts to convey this heterogeneity up to the catchment scale from finer scales, via hydrologically relevant landscape descriptors (Ebrahimian, Wilson, and Gulliver, 2016; Oudin *et al.*, 2018; Salavati *et al.*, 2016).

Concerning the regionalized model simulations, the results showed a relatively good performance of the RF for parameter estimation (Figure 5.5). The supremacy of the RF algorithm over the transfer-based regionalization methods was not sensitive to the chosen evaluation metric, and it was clearer when using non-transformed flows in calibration and evaluation. Moreover, the results were not altered by the size of the construction data, which confirms the robustness of the algorithm. Also, RF demonstrated better performance compared with transfer-based regionalization methods with regard to parameter estimation (Table 5.2).

The loss in the regionalized model performance compared with the time-transferred parameters remained large and comparable to previous regionalization studies with large-sample experiments. This can be explained by (1) the fact that the calibrated parameters were not easily estimable, since they may only represent a numerical optimum, although they were estimated over the entire recorded period (from 8 to 16 years at the hourly time step) in order to guarantee a maximum of stability. Note that the RFs are incapable of producing predictions outside the training range, as mentioned by Tyrallis, Papacharalampous, and Langousis (2019) and Hengl *et al.* (2018). Furthermore, the effects of compensation were not handled efficiently; (2) part of the difficulty might lie behind the fineness of the time step at which the study was performed. Regionalization at an hourly time step is rare or even nonexistent in comparison with daily and monthly time steps (Besaw *et al.*, 2010; Razavi and Coulibaly, 2013). If the water budget estimation at hourly time steps is as difficult as at daily time steps, then the challenge is to guarantee good correlation scores (i.e., synchronization) between the observed and the simulated streamflows, i.e., good estimation of the routing parameters. With regard to X_3 and X_4 , *RF_ALL* had a moderate maximum score in estimating X_3 ($R^2 = 0.45$, objective function: *KGESR*, Table 5.2) but failed to maintain the same score for X_4 ($R^2 = 0.29$, Table 5.2), and when the estimation of X_4 was satisfactory ($R^2 = 0.61$, objective function: *NSE*, Table 5.2), the corresponding X_3 estimation score was low ($R^2 = 0.32$, Table 5.2). *RF_URB* appeared to be globally more efficient in estimating these routing parameters ($R^2 = 0.41$ and 0.6 for X_4), perhaps due to the homogeneity in catchment size between the two urban-catchments samples; knowing that X_4 is, to first order, dependent on the catchment area; and (3) the descriptors might not be relevant in regionalizing the model parameters, or the methods of their determination were not suitable. For the time scale, the descriptors were averaged over

the whole simulation period, and also aggregated spatially up to the catchment scale, under the assumption of stable catchment characteristics over the study period. Accounting for intra-temporal and intra-catchment variability could yield more information to explain the variance of the parameters.

RF allows one to assess variable importance. However, interpreting the relative importance of the descriptors was not straightforward owing to (1) the integrative nature of both the parameters and the descriptors and (2) their inter-correlations. Generally, the interpretability of the RF algorithm is not as obvious as its skillfulness (Breiman, 2001b; Iorgulescu and Beven, 2004; Tyralis, Papacharalampous, and Langousis, 2019), which makes it an efficient prediction algorithm but an arduously interpretable one.

5.5.2 Weak sensitivity of the RF-derived relationships with the urbanization measure

Since the variable importance results showed non-significant scores for *CPD*, we attempted to examine whether shifts in this measure would have an impact on the parameter values.

As can be seen in Figure 5.8, X_1 was the only parameter that responded to the *CPD* shifts, displaying values similar to the ones estimated for the urban catchments. The reasons why this shift was not visible on the remaining parameters could be (1) the insensitivity of the model parameters to the *CPD* measure; (2) the multidimensional aspect of the parameter-descriptor relationship, which makes the parameters less sensitive to one descriptor itself but wholly dependent on all the descriptors, as shown in the variable importance, where almost no variable was remarkably influential. Perhaps this parameter insensitivity was amplified by the rural catchment insensitivity to the *CPD* shifts, as the natural descriptors (i.e., all the descriptors except *CPD*) predominate; and (3) the unsuitability (or insufficiency) of the *CPD* measure to describe the urbanization features, which may be viewed as a need for other urbanization descriptors to be included.

5.5.3 Conclusions and perspectives

As the ML algorithms are increasingly applied for various purposes in hydrology, we aimed at exploring the capacity of the RF algorithm in regionalizing hourly hydrological model parameters. The RF algorithm was applied using a large sample of American and French catchments, for which several climate and landscape descriptors were estimated. Two RFs were constructed for different sizes of training samples of catchments. The objective was to represent the relationships between the calibrated model parameters and climate/landscape descriptors

and then use these relationships to carry out estimates on urban catchments. The RF permitted us to: **(1)** achieve relatively good regionalized model performances compared with transfer-based regionalization methods, regardless of the construction sample size or the calibration objective function; **(2)** attain good performances concerning the model parameter estimations, but even better with regard to the transferring methods; and **(3)** perceive how increasing rates of urbanization influence model parameters. This large-sample hourly regionalization work at ungauged urban catchments has confirmed the need for ML algorithms such as RF in treating high-dimensional problems, uncovering nonlinear functions and, accounting for the uncertainty in parameter prediction (not discussed in this paper) with high efficiency and robustness.

Looking ahead, the sets of targeted parameter values should be refined more by looking for multi-objective suitable sets or even generic ones (Andréassian *et al.*, 2014; Andréassian *et al.*, 2012). Moreover, other urbanization descriptors should be included in order to refine the relationships. Compensation effects should be handled by estimating all the parameters at once using a unique RF for all the parameters or an iterative procedure where the sets of parameters are considered to be dependent on each other.

Synthèse du chapitre

L'objectif de ce chapitre est de montrer une méthode de régionalisation des paramètres du modèle GR4H pour les exprimer en fonction de plusieurs caractéristiques physiques du bassin, notamment le niveau d'urbanisation. Cette régionalisation a permis d'obtenir des relations de régression non linéaires (par le biais de l'algorithme du *random forest*) plus performantes que les méthodes classiques de transfert spatial des paramètres (proximité spatiale et similarité paysagère). En revanche, ces relations sont peu sensibles à la fraction du couvert urbain dans le bassin, ce qui décourage leur emploi pour étudier l'impact de schémas d'urbanisation future sur le comportement hydrologique.

CHAPITRE 6

UN MODÈLE PLUIE-DÉBIT CONCEPTUEL ET PHYSIQUEMENT COHÉRENT POUR LES BASSINS VERSANTS URBANISÉS

Sommaire

6.1	Introduction	227
6.2	Catchment sample	230
6.3	Tested model modifications	232
6.4	Results	241
6.5	Discussion and conclusion	248

Le modèle de départ GR4H a montré des résultats satisfaisants sur l'échantillon des bassins versants fortement urbanisés, avec des différences dans les distributions de paramètres traduisant des différences de comportement hydrologique entre les bassins fortement urbanisés et les bassins faiblement urbanisés. Au chapitre précédent, nous avons tenté de lier les paramètres du modèle GR4H avec une mesure du couvert urbain. Cette tentative a donné de bonnes performances, mais les relations entre les paramètres et la mesure du couvert urbain n'étaient pas encourageantes pour établir des scénarios hydrologiques à partir des scénarios d'urbanisation.

Nous suspectons que le niveau de paramétrisation *a minima* du modèle GR4H laisse aux paramètres trop de processus à gérer en même temps (par exemple, gérer à la fois le ruissellement des surfaces naturelles et des surfaces urbaines imperméables par un seul réservoir), ce qui n'impacte pas forcément les performances du modèle calé mais empêche d'avoir des relations de régionalisation satisfaisantes pour les études de scénarios. Une voie d'amélioration consiste à introduire des modifications dans la structure du modèle hydrologique en rajoutant des paramètres auxquels on imposera de traduire les spécificités hydrologiques du couvert urbain.

Correspondant à un article en révision soumis à *Journal of Hydrology*¹⁵, ce chapitre illustre une démarche de modification pas-à-pas du modèle GR4H, par laquelle on cherche à **(1)** mieux lier la structure du modèle aux spécificités de l'urbanisation, et **(2)** améliorer la simulation de la relation pluie-débit sur 273 bassins fortement urbanisés (ayant un taux moyen d'imperméabilité totale *TIA* supérieur à 5%). Les modifications sont inspirées d'une part des constats de la littérature, et d'autre part des résultats de cette thèse sur le comportement hydrologique des bassins fortement urbanisés (voir les chapitres 3 et 4). La pertinence de chacune des modifications est jugée à l'aide de 6 métriques continues et 3 métriques événementielles. La variabilité des paramètres rajoutés en fonction de *TIA* est également analysée, dans la perspective de régionaliser ces paramètres pour utiliser le modèle dans le cadre d'étude de scénarios d'urbanisation.

¹⁵Sous le titre : "Physically consistent conceptual rainfall-runoff model for urbanized catchments".

Abstract

Hydrological models should be tested and evaluated for a wide variety of urban sprawl before they are used to predict the impact of urbanization on catchment behavior. In this study, we illustrate a top-down approach of adapting step by step an hourly conceptual model structure (GR4H) to account for urbanization features. Adapting the model included accounting explicitly for runoff from sealed surfaces by bypassing the soil moisture reservoir and varying the partitioning between quick flow and slow flow. Tested model adaptations were chosen based on the reported specificities of urbanized catchments, namely, decreasing infiltration, increasing runoff, and fast runoff dynamics. Using a split-sample test, the relevance of each modification with regard to the reproduction of catchment response (i.e., observed streamflow) was assessed for a large sample of 273 urbanized catchments, located in France and the United States, for which mean total impervious area (TIA) varied between 0.05 and 0.59. Six continuous and three event-based criteria were used, and a binomial test was applied to assess the statistical significance of improvements. Results showed the following: **(i)** The tested modifications improved the ability of GR4H to reproduce the catchment response, especially high flows and observed streamflow amid dry conditions. **(ii)** Event-based evaluation using more than 45,000 events showed an improvement in predicting the event peak flow and event runoff volume, whereas no significant improvements were obtained in predicting the timing of peak flow. **(iii)** Newly added parameters were moderately to highly correlated with TIA , especially the calibrated proportion of sealed surfaces, which is promising as a hydrological validation of estimated urbanization measures from land cover. The tested modifications have improved both the representation of urbanization processes and the reproduction of the observed streamflow, yielding a simple and credible model for predicting the impact of future urbanization scenarios on catchment response.

6.1 Introduction

6.1.1 Hydrological models as valuable tools for the assessment of urbanization impact

There is strong evidence that urbanization modifies the hydrological behavior of catchments (Braud *et al.*, 2013; Diem, Hill, and Milligan, 2018; Fletcher, Andrieu, and Hamel, 2013; McGrane, 2016; Miller and Hess, 2017). Nonetheless, predicting and quantifying the impact of urbanization on the rainfall-runoff relationship is still a thorny issue. To this end, two types of approaches are generally applied (Braud *et al.*, 2013; Salavati *et al.*, 2016): a statistical approach and a modeling approach. The statistical approach seeks to **(i)** either identify

temporal trends in the behavior of a catchment provided that long periods of hydroclimatic data are available across the urban sprawl over the catchment area (Haase, 2009), or (ii) compare the behavior of an urbanized catchment with another “similar” but non-urbanized one, as in paired-catchment experiments (Bonneau *et al.*, 2018; Prosdocimi, Kjeldsen, and Miller, 2015). These approaches are useful to detect and quantify past changes, but they do not provide straightforward information about how the catchment behavior would be altered under future urbanization scenarios. In this respect, the modeling approach is advantageous because it backs the statistical method by providing a synthesis of the catchment behavior via the model parameters. This helps not only to detect change in catchment behavior (Pathiraja *et al.*, 2018; Saadi, Oudin, and Ribstein, 2020), but also to create hydrological scenarios that correspond to urban sprawl scenarios (De Niel *et al.*, 2020; Niehoff, Fritsch, and Bronstert, 2002; McIntyre *et al.*, 2014; Sanzana *et al.*, 2019).

6.1.2 Overview of hydrological models for urbanized catchments

There is a spectrum of modeling tools along which a compromise is made between detailed representation of the spatial variability of hydrological processes and model simplicity (Hrachowitz and Clark, 2017; Salvatore, Bronders, and Batelaan, 2015; McIntyre *et al.*, 2014). Distributed models that use small-scale equations to represent the main hydrological processes are applied to account for the highly heterogeneous nature of hydrological processes in urban areas, in addition to the rapid dynamics of runoff generation on sealed surfaces (Cristiano, Veldhuis, and Giesen, 2017; Ogden *et al.*, 2011; Salvatore, Bronders, and Batelaan, 2015). Their application is also advocated because they explicitly represent catchment properties, which enables a straightforward assessment of the impact of urbanization by changing model parameters and scaling up the impact (Beven, 2002; Bronstert, Niehoff, and Bürger, 2002). This generally leads to heavily parametrized model structures (e.g., Cuo *et al.*, 2008; Jankowsky *et al.*, 2014; Jia *et al.*, 2001; Sanzana *et al.*, 2019; Stavropoulos-Laffaille *et al.*, 2018), which impedes their use to represent the rainfall-runoff relationship for catchments with a mix of rural and urban areas. Also, constraining these models requires a large volume of data, available only for a handful of monitored catchments (Petrucci and Bonhomme, 2014; Rodriguez, Andrieu, and Creutin, 2003).

Conversely, the use of relatively simple, conceptual models is undermined by lacking an explicit link between urbanization features and model structure or parameters (De Niel *et al.*, 2020; Huang *et al.*, 2008a; Lin *et al.*, 2014). Unlike models with detailed representation of urban landscapes, the relative simplicity of conceptual models offers an opportunity for their improvement with regard to linking their structures to the physical properties of the catchments (Euser *et al.*, 2015; Hrachowitz *et al.*, 2014; Gharari *et al.*, 2014; Kirchner, 2006). In addition, they can be tested over a large set of catchments due to their low data demand (Gupta

et al., 2014). An example was shown by Kjeldsen, Miller, and Packman (2013), who modified a non-urban model structure to account for urbanization features. They explicitly linked the model parameters (up to four) to the proportion of urban cover in the catchment, but their test concerned only seven catchments located in the United Kingdom. Another study by Hamel and Fletcher (2014) illustrated an improvement of the representation of subsurface flow by testing different model configurations, based on an analysis of the hydrological signatures of the McMahons Creek catchment in Australia. Their aim was to predict the impact of different stormwater management strategies on baseflow. Both studies illustrate a top-down approach of improving the representation of urban hydrological processes in conceptual models, although their application was constrained to small sets of catchments. Thus, an attempt is required to balance the simplicity of conceptual models with robust model assessment using large samples of catchments (Andréassian *et al.*, 2009; Gupta *et al.*, 2014).

6.1.3 Novelty and scope of the study

One of the major weaknesses in currently available models for urbanized catchments is the lack of intensive testing on large samples of catchments, with diverse situations of urban sprawl. We argue that this is a requirement to **(i)** advance our general understanding of how urbanized catchments behave hydrologically, by building on what we already know from non-urbanized ones, and thus **(ii)** make a credible extrapolation of catchment behavior under future urban planning schemes. Simple conceptual models developed for non-urbanized catchments have been proved to be flexible enough to adapt their model parameters to urbanized contexts (Saadi, Oudin, and Ribstein, 2020). Adapting their structures to urbanized environments should fulfill the following requirements: **(i)** an equal or improved ability to simulate the response of urbanized catchments (Fidal and Kjeldsen, 2020), and **(ii)** more explicit links between the model parameters/structure and the urban characteristics of the catchments.

In this study, we used a large sample of 273 urbanized catchments located in France and the United States, for which the mean total impervious area (*TIA*) ranged between 0.05 and 0.59. We conducted a step-by-step modification of the hourly non-urban GR4H model (Ficchi, Perrin, and Andréassian, 2019). Each modification was evaluated using a set of six continuous and three event-based evaluation criteria, and a binomial test was applied to evaluate the statistical significance of each improvement with regard to the initial non-urban model structure (Fidal and Kjeldsen, 2020). We aimed at proposing physically sound adaptations by **(i)** testing adaptations that were in agreement with the behavioral specificities of urbanized catchments, and **(ii)** comparing the added model parameters with observable urban characteristics.

This paper is organized as follows: In Section 6.2, we present the catchment set. Section 6.3 details the tested model structure modifications, and the framework of their calibration and

evaluation. Results are presented in Section 6.4, followed by a discussion and conclusion in Section 6.5.

6.2 Catchment sample

We selected a large sample of 273 urbanized catchments, located in the United States (US) and France, based on four criteria (Saadi, Oudin, and Ribstein, 2019):

1. Availability of at least 8 years of hourly streamflow, precipitation, and daily temperature between 1997 and 2017. We considered that a minimum of 4 years are required for model calibration (Perrin *et al.*, 2007).
2. Limited snow influence, as snow melting was not addressed in the tested model.
3. Limited impact of large artificial reservoirs, which required streamflow naturalization.
4. A minimum mean total impervious area (*TIA*) of 0.05. This choice aimed at including the lowest level of urban sprawl above which the impact of urbanization on catchment behavior is often considered significant (Arnold and Gibbons, 1996; Salvatore, Bronders, and Batelaan, 2015). Moreover, it offered a wide range of urban sprawl for a more robust model testing.

For the catchments located in the United States, we obtained the hourly precipitation depths from the Stage IV dataset of National Centers for Environmental Prediction (NCEP) via the `geoknifeR` package (Read *et al.*, 2015). We processed the hourly streamflow depths from the instantaneous hydrographs, which we extracted using the `dataRetrievalR` package (Cicco *et al.*, 2018) from the dataset of gauges maintained by the US Geological Survey (Falcone, 2011). We used the Daymet dataset to compute daily temperature time series (Thornton *et al.*, 2016). Land cover was characterized using the National Land Cover Database available for the years 2001, 2006, 2011, and 2016 (Wickham *et al.*, 2014). The location of the 205 US catchments is shown in Figure 6.1.

For the catchments located in France, we processed the hourly precipitation depths from the COMEPHORE product of Météo France (Tabary *et al.*, 2013). We extracted the hourly streamflow depths from the Banque HYDRO dataset (Leleu *et al.*, 2014). Daily temperature was provided by the SAFRAN product of Météo France (Vidal *et al.*, 2010). Land cover was characterized using the Imperviousness Density layers of the Copernicus Land Monitoring Service, available for the years 2006, 2009, 2012, and 2015 (e.g., Congedo *et al.*, 2016). The location of the 68 French catchments is shown in Figure 6.2.

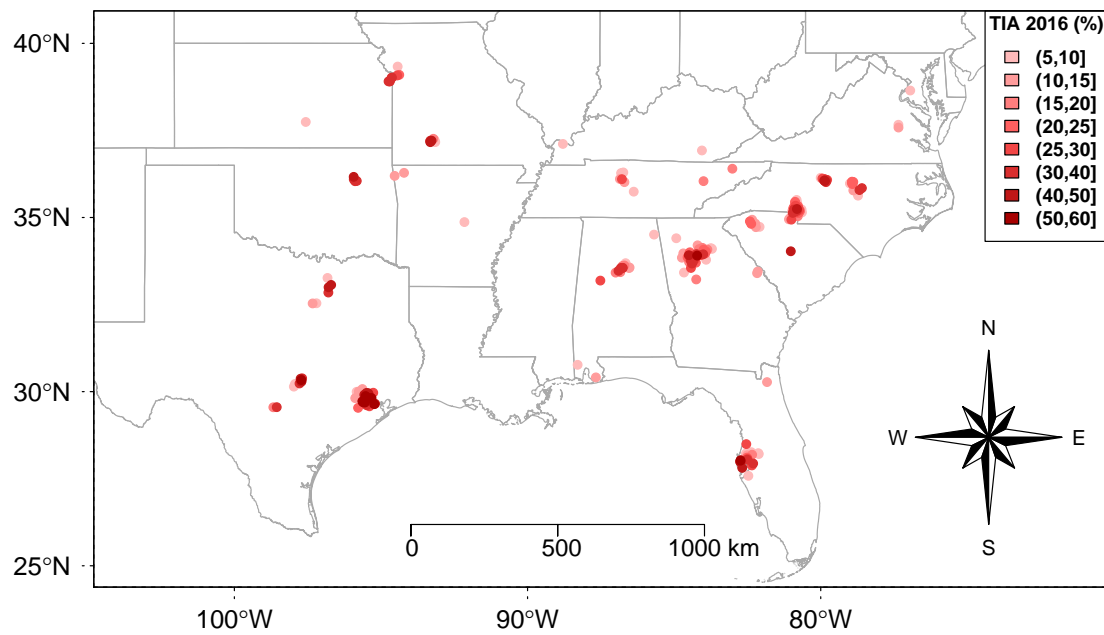


Figure 6.1. Location of the 205 US catchments. A dot represents the location of the catchment centroid. Colors represent the mean total impervious area (*TIA*) of the catchment for the year 2016, estimated from the National Land Cover Database.

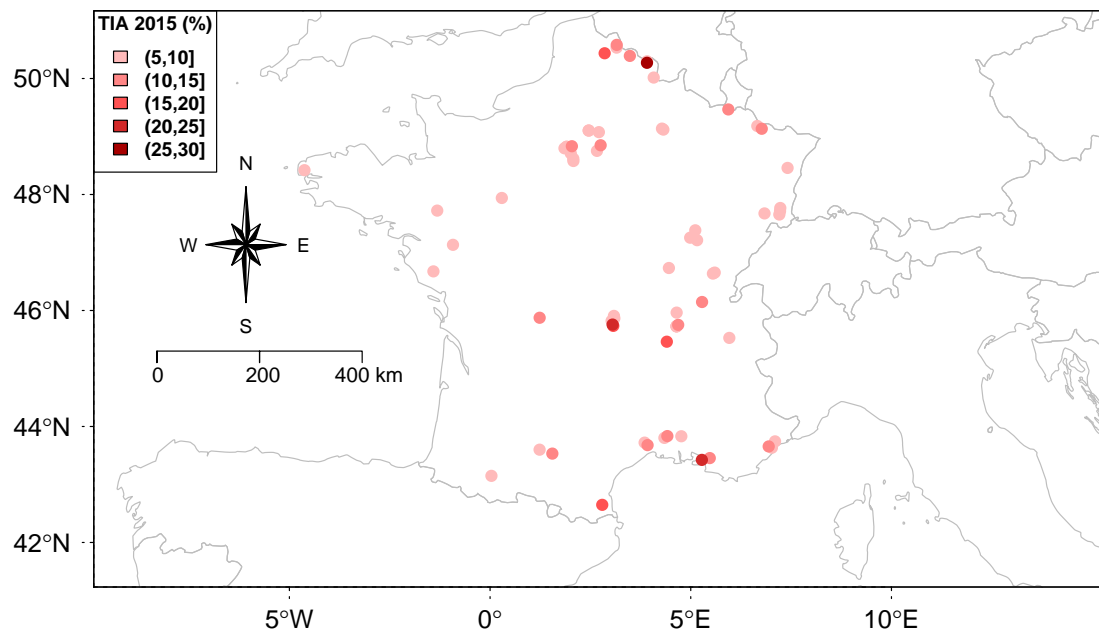


Figure 6.2. Location of the 68 French catchments. A dot represents the location of the catchment centroid. Colors represent the mean total impervious area (*TIA*) of the catchment for the year 2015, extracted from the Imperviousness Density layers of the Copernicus Land Monitoring Service.

In addition to precipitation, potential evapotranspiration was required as climatic forcing. It was estimated at daily time steps using a temperature-based formula (Oudin *et al.*, 2005). Then, an hourly disaggregation using the hourly extraterrestrial radiation helped estimate the hourly values (Allen *et al.*, 1998).

For each year of land cover data, the catchment *TIA* was calculated as the arithmetic mean of the proportion of sealed surface mapped over each 20-m pixel in France and 30-m pixel in the United States. For the missing years, a linear interpolation was used to compute the corresponding *TIA* values.

We extracted a set of events from the precipitation and streamflow time series for event-based assessment, using an empirical method (see Appendix A; Ficchi, Perrin, and Andréassian, 2016). This method was applied to the direct flow time series to distinguish the events originating from surface runoff. Direct flow was determined by applying a numerical filter to extract the baseflow (Blume, Zehe, and Bronstert, 2007; Eckhardt, 2005; Collischonn and Fan, 2013; Mei and Anagnostou, 2015).

Most of the US catchments used in this study are characterized by a temperate humid climate (Cfa type in the Köppen-Geiger classification, see Beck *et al.*, 2018), while the French catchments are characterized by temperate oceanic (Cfb) and Mediterranean climates (Csa). For the whole sample, the catchment area ranged between 1.1 km² and 2100 km², with a median value of 60 km². Mean annual precipitation varied between 510 and 1660 mm/year (median: 1170 mm/year), whereas mean potential evapotranspiration ranged between 620 and 1400 mm/year (median: 1030 mm/year). The ratio of mean precipitation to mean potential evapotranspiration was between 0.6 and 2.2 (median: 1.1). The *TIA* range was 0.05-0.59, with a median of 0.16. The lengths of the recorded periods were 16 years for 190 out of 273 catchments, whereas only 11 catchments had a recorded period of less than 10 years of length. The event separation method yielded a total of 45,025 events, with a median number of events per catchment of approximately 140 events (interquartile range: 83-239).

6.3 Tested model modifications

6.3.1 Original model structure

Urbanized catchments are characterized by fast response dynamics due to surface sealing and direct conveyance through artificial drainage systems (Fletcher, Andrieu, and Hamel, 2013; Rodriguez, Andrieu, and Creutin, 2003; Salvatore, Bronders, and Batelaan, 2015). Hence, the choice of hourly to sub-hourly time steps is necessary in order to accurately describe the

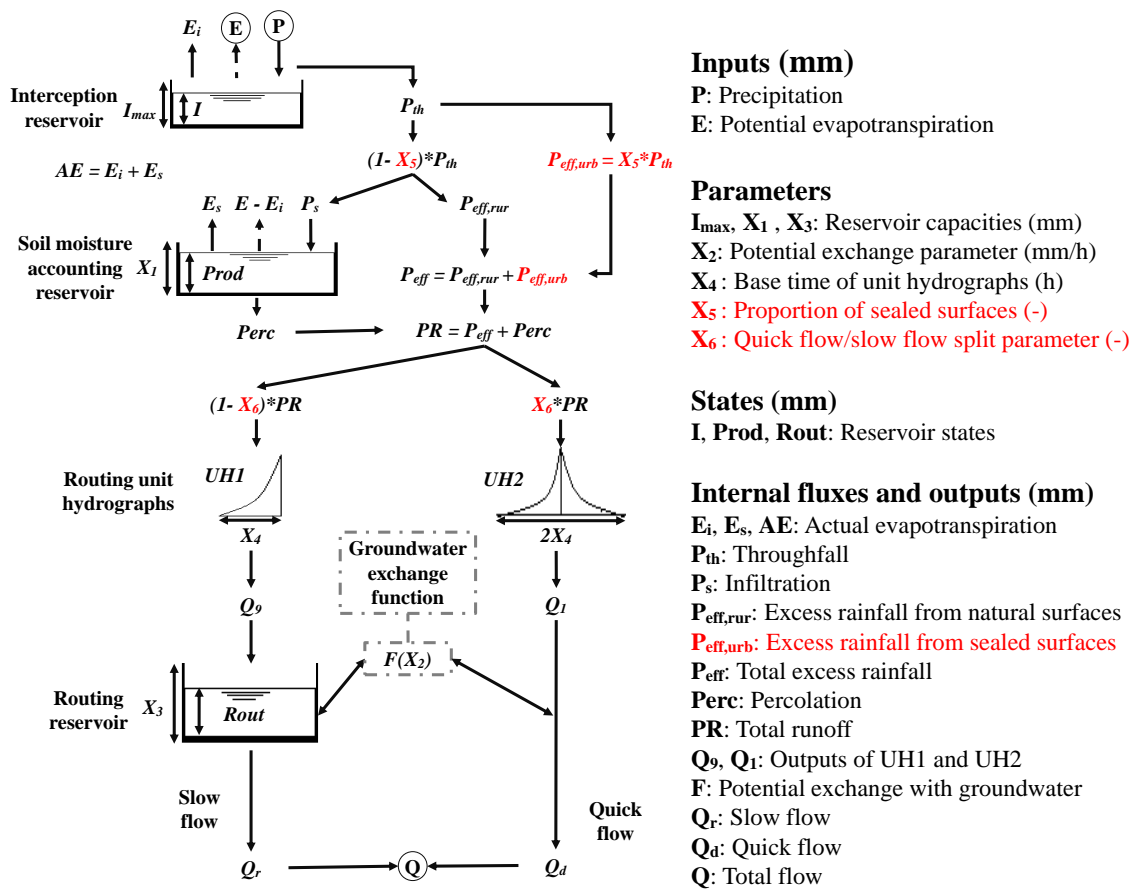


Figure 6.3. Description of GR4H model structure, parameters, states, and internal fluxes. This structure was inherited from the daily version GR4J, for which a detailed description of equations is given by Perrin, Michel, and Andréassian (2003). The model was recently modified by Ficchi, Perrin, and Andréassian (2019) by adding an interception reservoir to guarantee a stable internal functioning across different time steps. In red, we highlight the modifications that we carried out to account for urbanization features. In the original model structure, X_5 is set at 0, and $P_{eff,urb}$ is null. $P_{eff,rur}$ is equal to $(1 - X_5) \cdot P_{th} - P_s = P_{th} - P_s$. Also, X_6 is fixed at 0.1.

catchment response. However, the paucity of hydroclimatic datasets at fine time steps imposes a maximum time resolution. Considering both constraints, we chose the hourly time step.

The GR4H model (Ficchi, Perrin, and Andréassian, 2019) was selected as the starting point for model development. This model was not developed specifically for urbanized catchments, but it has been widely tested in large international sets of catchments and in model comparison studies (Boer-Euser *et al.*, 2017; Esse *et al.*, 2013; Le Moine *et al.*, 2007). In addition, it proved to be comparatively flexible when adapting its calibrated parameters to urbanized catchments, leading to a relatively satisfactory reproduction of observed streamflow for a large spectrum of urbanized contexts (Saadi, Oudin, and Ribstein, 2020). The structure of the model is shown in Figure 6.3.

The interception reservoir is fed by the precipitation P and emptied to satisfy the potential evapotranspiration E . The throughfall P_{th} is generated when the interception reservoir is full. In this case, the soil moisture-accounting reservoir (also referred to as the “production reservoir”) is fed by a depth P_s , function of P_{th} , $Prod$ and X_1 , where $Prod$ is the updated state of the production reservoir and X_1 its capacity. The excess rainfall, i.e., $P_{th} - P_s$, is added to percolation $Perc$ from the production reservoir to constitute the net precipitation PR . The latter is split into $0.1 \cdot PR$ routed through the quick-flow branch (unit hydrograph $UH2$ with a base time $2 \cdot X_4$ hours), and $0.9 \cdot PR$ passed through the slow-flow branch (unit hydrograph $UH1$ with a base time X_4 hours), followed by a non-linear reservoir with updated state $Rout$ and capacity X_3 . Before reaching the simulated total flow, a surface water–groundwater exchange algebraic depth $F(X_2)$, a function of X_2 and the rate $\frac{Rout}{X_3}$, is added to the outflow of $UH1$ and the outflow of $UH2$. $F(X_2)$ is positive (or negative) when water is imported from (or exported to) groundwater. Finally, the total flow is the sum of the quick flow Q_d and the slow flow Q_r , which is the outflow of the routing reservoir, determined using $Rout$ and X_3 .

Only four parameters are numerically calibrated: X_1 (mm) the capacity of the production store, X_2 (mm/h) the exchange function parameter, X_3 (mm) the capacity of the non-linear routing reservoir, and X_4 (h) the base time of the unit hydrographs. I_{max} (mm), the capacity of the interception reservoir, was estimated by reducing the differences between hourly and daily throughfall (Ficchi, Perrin, and Andréassian, 2019), where the latter is estimated at the daily time steps by comparing daily P and E (Perrin, Michel, and Andréassian, 2003). For the set of catchments used here, I_{max} varied between 1.5 and 4.0 mm, with a median value of 2.75 mm.

6.3.2 Tested modifications to account for urbanization features and improve model performances

6.3.2.1 Accounting for the presence of sealed surfaces

Urbanization increases the proportion of sealed surfaces, which limits infiltration and increases runoff (Miller and Hess, 2017; Zhou *et al.*, 2017). This is equivalent to establishing a sort of disconnection between catchment response (i.e., runoff) and the soil moisture state. To account for this effect, the throughfall P_{th} is split into (see Figure 6.3): **(i)** a portion that bypasses the production reservoir and becomes excess rainfall immediately, $X_5 \cdot P_{th}$, which can be interpreted as the excess rainfall from sealed/impervious surfaces, and **(ii)** a portion that falls on the natural surfaces, $(1 - X_5) \cdot P_{th}$, and feeds the production reservoir. The parameter X_5 (–) is intended to represent the proportion of the sealed surfaces, which can be conceptually linked to TIA .

In the original structure, the ratio of instantaneous excess rainfall p_{eff} (mm/h) to instantana-

neous throughfall p_{th} (mm/h) is expressed as:

$$\frac{p_{eff}}{p_{th}} = \left(\frac{Prod}{X_1} \right)^2 \quad (6.3.1)$$

By including the excess rainfall on sealed surfaces, Equation 6.3.1 becomes:

$$\frac{p_{eff}}{p_{th}} = X_5 + (1 - X_5) \cdot \left(\frac{Prod}{X_1} \right)^2 \quad (6.3.2)$$

Under very dry conditions ($Prod \sim 0$), the proportion of throughfall that is converted into excess rainfall is as high as X_5 (Figure 6.4), i.e., as high as the mean catchment imperviousness. Note that with $X_5 = 1$, i.e., a completely sealed catchment, all throughfall bypasses the production reservoir and thus instantaneously reaches the transfer function. On the other hand, with $X_5 = 0$, we go back to the original formulation in Equation 6.3.1, i.e., the pre-sealed era of the catchment, for which all excess rainfall is estimated based on the production reservoir filling rate $\frac{Prod}{X_1}$.

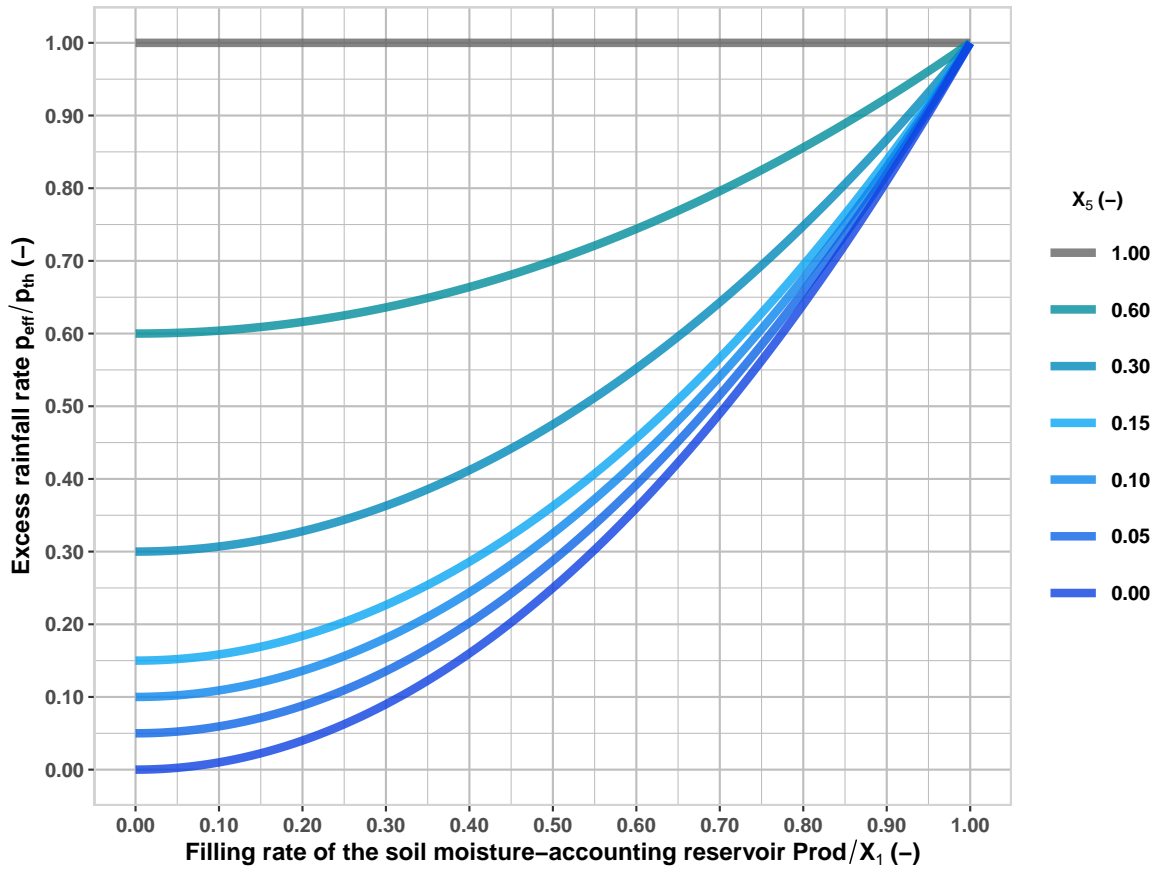


Figure 6.4. Illustration of the effect of different values of X_5 , representing conceptually the proportion of sealed surfaces, on the relationship between the excess rainfall rate $\frac{p_{eff}}{p_{th}}$, and $\frac{Prod}{X_1}$, the filling rate of the soil moisture-accounting reservoir.

Equation 6.3.2 can be also viewed as:

$$\begin{aligned}
 p_{eff} &= X_5 \cdot p_{th} + (1 - X_5) \cdot p_{th} \cdot \left(\frac{Prod}{X_1} \right)^2 \\
 &= p_{eff,urb} + p_{eff,rur}
 \end{aligned} \tag{6.3.3}$$

where $p_{eff,urb} = X_5 \cdot p_{th}$ is the excess rainfall generated in sealed surfaces, and $p_{eff,rur} = (1 - X_5) \cdot p_{th} \cdot \left(\frac{Prod}{X_1} \right)^2$ is the excess rainfall coming from natural surfaces.

By integrating over a time step of 1 hour, the depth P_s (mm) that feeds the soil moisture reservoir becomes:

$$P_s = \frac{X_1 \cdot \left(1 - \left(\frac{Prod}{X_1} \right)^2 \right) \cdot \tanh \left(\frac{(1 - X_5) \cdot P_{th}}{X_1} \right)}{1 + \frac{Prod}{X_1} \cdot \tanh \left(\frac{(1 - X_5) \cdot P_{th}}{X_1} \right)} \tag{6.3.4}$$

which is now dependent on X_5 , in addition to X_1 , P_{th} (p_{th} integrated over 1 hour), and $\frac{Prod}{X_1}$.

6.3.2.2 Accounting for the impact of urbanization on quick-flow/slow-flow partitioning

With increased urbanization, catchment dynamics are intensified (Diem, Hill, and Milligan, 2018; Miller and Hess, 2017; Saadi, Oudin, and Ribstein, 2020), due to reduced surface roughness and artificial drainage systems. This results in higher amounts of net precipitation PR reaching the catchment outlet in a shorter span of time. In the original structure, the quick-flow/slow-flow split parameter X_6 was fixed at 0.1 (Figure 6.3), i.e., 10% and 90% of runoff are constituted by quick flow and slow flow, respectively. Fixing this model parameter was chosen in the early model developments as its calibration on predominantly non-urbanized catchments did not result in significantly improved model performances (Perrin, Michel, and Andréassian, 2003). In urbanized catchments, this parameter would reach higher values since quick flow is expected to be quantitatively increased. Thus, we tested a model version for which this parameter is free and calibrated it for each catchment.

6.3.2.3 Summary of tested models

Six model versions were compared over the 273 catchments, and their configurations are summarized in Table 6.1. Three possibilities were tested for the proportion of sealed surfaces X_5 : (1) $X_5 = 0$ as in the original version, (2) X_5 is updated every year with *a priori* estimated TIA from land cover datasets, and (3) X_5 is set free and calibrated for each catchment. Moreover, two possibilities were tested for X_6 , the quick-flow/slow-flow split parameter: (1) X_6 fixed at 0.1, as in the original version, and (2) X_6 calibrated.

6.3.3 Evaluation of tested modifications

Retaining any of the model modifications should be warranted by a significant improvement of model performances on a wide set of catchments (Andréassian *et al.*, 2009). The five model modifications (listed in Table 6.1) were assessed in comparison with the GR4H original model structure both on continuous and event bases, using a split-sample test (Klemeš, 1986). In this section, we detail (i) the model calibration procedure, (ii) the continuous and event-based assessment criteria, and (iii) the statistical test that we applied to assess the significance of differences in model performances.

First, the recorded period of each catchment is split into two subperiods P1 and P2 with equivalent lengths. For each subperiod, each model is calibrated by maximizing the objective

Table 6.1. Summary of tested model modifications. The number of calibrated parameters is given in each model name. The letter **R** refers to “rural”, i.e., no sealed surfaces are accounted for ($X_5 = 0$), whereas the letter **U** refers to “urbanized”. **G** refers to “Génie”, **M** refers to “model”, and **H** refers to “hourly”.

Model name	Urban production parameter X_5	Quick-flow/slow-flow portioning parameter X_6
GR4H (Original structure)	Fixed, $X_5 = 0$	Fixed, $X_6 = 0.1$
MR5H	Fixed, $X_5 = 0$	Calibrated
MU4H	Fixed, $X_5 = TIA$ estimated from land cover dataset	Fixed, $X_6 = 0.1$
MU5H	Fixed, $X_5 = TIA$ estimated from land cover dataset	Calibrated
MUOpt5H	Calibrated	Fixed, $X_6 = 0.1$
MU6H	Calibrated	Calibrated

function, chosen as the Kling-Gupta efficiency KGE (Gupta *et al.*, 2009) applied to the square root values of observed and simulated streamflow depths ($KGESQ$), in order to guarantee a uniform emphasis on all streamflow components (Oudin *et al.*, 2006a; Santos, Thirel, and Perrin, 2018). $KGESQ$ is computed as:

$$KGESQ = 1 - \sqrt{(1-r)^2 + (1-\alpha)^2 + (1-\beta)^2} \quad (6.3.5)$$

where r refers to the correlation of the observed and simulated time series, α the ratio of their standard deviations, and β the ratio of their means. Ideally, we would like $KGESQ$ to be as close as possible to 1 (i.e., $r = \alpha = \beta = 1$). To this end, we broadly inspected the parameter hyperspace to look for a starting point, from which a gradient descent algorithm was followed to reach the optimal set of parameters that maximizes $KGESQ$ (Coron *et al.*, 2017; Edijatno *et al.*, 1999).

Second, the generated parameter set is tested on the alternative period (P1 if calibration on P2, and *vice versa*). Continuous streamflow evaluation is conducted using two criteria: $KGESQ$ and the Nash-Sutcliffe Efficiency (NSE) applied to observed and simulated streamflow depths, according to (Nash and Sutcliffe, 1970):

$$NSE = 1 - \frac{\sum_h (Q_h^{obs} - Q_h^{sim})^2}{\sum_h (Q_h^{obs} - \overline{Q^{obs}})^2} \quad (6.3.6)$$

where Q_h^{sim} , Q_h^{obs} are simulated and observed discharges at hour h , respectively, and $\overline{Q^{obs}}$ is the mean observed discharge. NSE is equal to 1 with identical simulated and observed streamflow time series.

Contrasts between urbanized and non-urbanized catchments are reported to be more significant during dry periods (Sillanpää and Koivusalo, 2015; Zhou *et al.*, 2017). Thus, $KGESQ$ and NSE were also computed with emphasis on wet ($KGESQ_{wet}$ and NSE_{wet}) and dry ($KGESQ_{dry}$ and NSE_{dry}) conditions of the test period. To do this, we weighted the observed and simulated streamflow time series at each hour h by the weights w_{wet_h} and w_{dry_h} (Oudin *et al.*, 2006a), computed as:

$$\begin{cases} w_{wet_h} &= \frac{1}{1 + \exp\left(-25 \cdot \left(\frac{Prod_h}{X_1} - 0.5\right)\right)} \\ w_{dry_h} &= \frac{1}{1 + \exp\left(+25 \cdot \left(\frac{Prod_h}{X_1} - 0.5\right)\right)} \end{cases} \quad (6.3.7)$$

with $Prod_h$ estimated at hour h using the interception and production reservoirs of GR4H (Figure 6.3), which varies between 0 and X_1 . When $\frac{Prod_h}{X_1}$ approaches 1 (i.e., very wet conditions), $w_{wet_h} \sim 1$ and $w_{dry_h} \sim 0$, and *vice versa*. Note that w_{wet_h} and w_{dry_h} are always between 0 and 1. Different values of X_1 between 10 mm and 750 mm were tested to see which configuration dampens the observed streamflow best in winter (i.e., December to February) with w_{dry_h} , and in summer (i.e., June to August) with w_{wet_h} (not shown here). A value of $X_1 = 200$ mm yielded the best trade-off and was chosen to compute w_{dry_h} and w_{wet_h} at each hour h for every catchment.

We are much more interested in the performances of the tested models relative to the original structure. Hence, for each continuous criterion $Crit$, a relative index R_{Crit} was used to assess the reduction in error due to each modification in comparison with the original model structure. R_{Crit} is computed as (Lerat *et al.*, 2012):

$$R_{Crit}(Modif, Ref) = \frac{Crit(Modif) - Crit(Ref)}{2 - Crit(Modif) - Crit(Ref)} = \frac{\text{Error}(Ref) - \text{Error}(Modif)}{\text{Error}(Ref) + \text{Error}(Modif)} \quad (6.3.8)$$

with $Crit \in \{KGESQ, NSE, KGESQ_{wet}, NSE_{wet}, KGESQ_{dry}, NSE_{dry}\}$, $Modif$ is the modified model, and Ref is the reference model, in our case GR4H. R_{Crit} varies between -1 and 1 , with positive values indicating an improvement and negative values indicating a degradation relatively to the reference model structure Ref (i.e., GR4H). Note that $R_{Crit}(Ref, Ref) = 0$. A better interpretation of R_{Crit} can be obtained by computing the corresponding reduction in

error:

$$\left\{ \begin{array}{l} \text{Error} \\ \text{Reduction in error} \end{array} \right. = 1 - \text{Crit} = 1 - \frac{\text{Error}(\text{Modif})}{\text{Error}(\text{Ref})} = \frac{2 \cdot R_{\text{Crit}}(\text{Modif}, \text{Ref})}{1 + R_{\text{Crit}}(\text{Modif}, \text{Ref})} \quad (6.3.9)$$

In addition, the ability of models to reproduce the hydrographs of events belonging to the test subperiods was assessed using the error in estimating event peak flow e_{Qp} , the error in predicting the timing of peak flow e_{tp} , and the volumetric efficiency VE (Criss and Winston, 2008), defined as:

$$\left\{ \begin{array}{l} e_{Qp} = \frac{Q_p^{\text{sim}} - Q_p^{\text{obs}}}{Q_p^{\text{obs}}} \\ e_{tp} = \frac{h_p^{\text{obs}} - h_p^{\text{sim}}}{h_p^{\text{obs}} - h_{st}^{\text{obs}}} \\ VE = 1 - \frac{\sum_{h_{st}^{\text{obs}} \leq h \leq h_{end}^{\text{obs}}} |Q_h^{\text{obs}} - Q_h^{\text{sim}}|}{\sum_{h_{st}^{\text{obs}} \leq h \leq h_{end}^{\text{obs}}} Q_h^{\text{obs}}} \end{array} \right. \quad (6.3.10)$$

with Q_p^{sim} , Q_p^{obs} the simulated and observed peak flows, and h_p^{sim} , h_p^{obs} the hours of their occurrence. h_{st}^{obs} , h_{end}^{obs} are the starting and ending hours of each event. The ideal value for e_{Qp} and e_{tp} is 0, whereas the ideal value of VE is 1. Positive e_{Qp} indicates an overestimation of the peak flow, whereas positive e_{tp} indicates that the predicted peak flow occurs before the observed peak flow.

Third, a binomial statistical test was used to accept/reject each modification (Clopper and Pearson, 1934; Fidal and Kjeldsen, 2020). The probability that the modified model version M' outperforms the original version M on a number of subperiods or events is given by the binomial distribution. Outperformance of M' with respect to M is achieved with a positive relative index R_{Crit} (Equation 6.3.8), decreased absolute values of e_{Qp} , decreased absolute values of e_{tp} , and increased VE . The null hypothesis is set such that M performs equally to or better than M' . The alternative hypothesis is thus M' outperforms M , i.e., the model modification leads to a statistically significant improvement. The number of trials, i.e., how many times M and M' were compared, is 546 (the number of subperiods) in the case of continuous assessment and 45,025 (the number of events) in event-based assessment. The statistical test was accomplished for each criterion. The null hypothesis is rejected when the confidence interval of the proportion of cases when M' outperforms M (given by the binomial distribution) is greater than 0.5 (i.e., the proportion of cases corresponding to equal performances of M' and M).

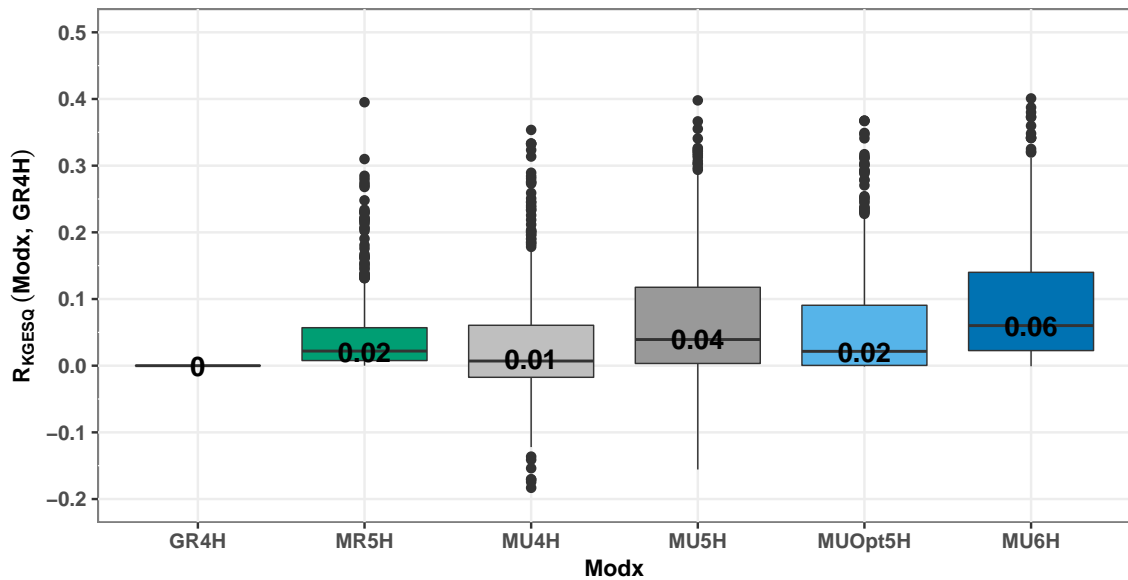


Figure 6.5. Distributions of improvements in calibration performances obtained for the six tested model structures, expressed in terms of relative index R_{KGESQ} . Boxes are delimited by the first and third quartiles. Values indicate the median of each distribution.

6.4 Results

6.4.1 Calibration performances and distributions of calibrated parameters

Optimizing more parameters resulted in improved calibration performances. Moreover, accounting for TIA resulted in better calibration performances for the majority of catchments, even with no additional optimized parameters (MU4H), as shown by the distributions of R_{KGESQ} in Figure 6.5. The median reduction of error ranged between 2% ($R_{KGESQ} = 0.01$) and 11% ($R_{KGESQ} = 0.06$), and improvements were observed for the majority of cases. When adding one numerically calibrated parameter, the improvement by optimizing the proportion of sealed surfaces X_5 was higher than by optimizing the quick-flow/slow-flow split parameter X_6 (MUOpt5H vs. MR5H). However, optimizing X_6 while fixing $X_5 = TIA$ (MU5H) resulted in a better improvement. As expected, the best improvements were obtained by optimizing all possible model parameters (MU6H).

Interestingly, values of calibrated X_5 were related to TIA , and a greater improvement in calibration performances corresponded to a higher level of urban sprawl, as illustrated in Figure 6.6. Note that X_5 was calibrated only for models MUOpt5H and MU6H. The median

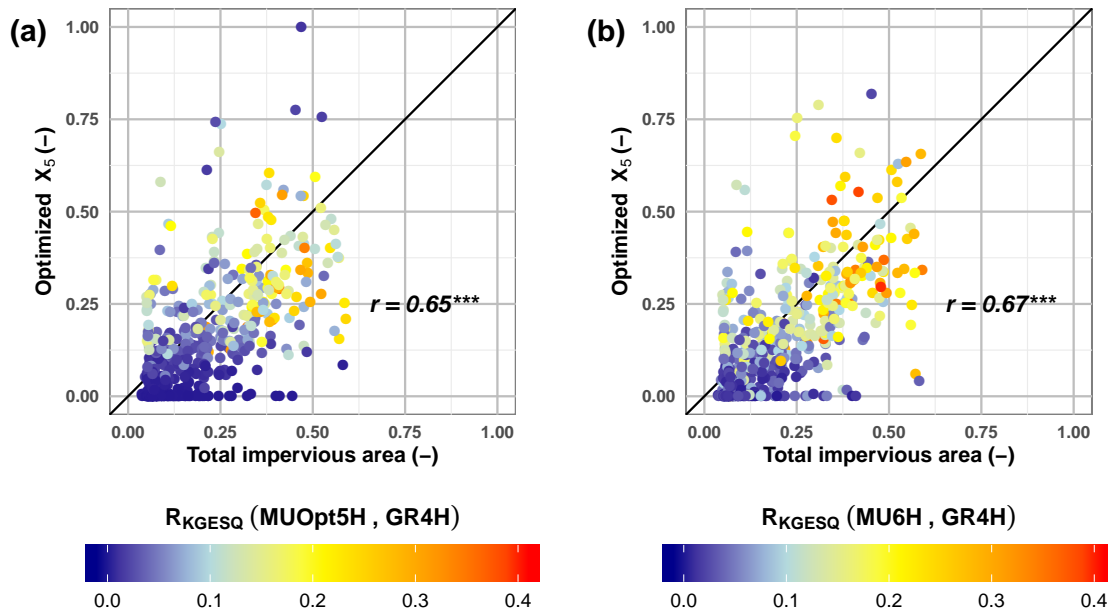


Figure 6.6. Calibrated proportion of sealed surfaces X_5 vs. estimated total impervious area from land cover databases. X_5 was calibrated for the models (a) MUOpt5H and (b) MU6H, as indicated in Table 6.1. Colors indicate the relative index R_{KGESQ} , i.e., the improvement in calibration performances with respect to the original GR4H structure. Asterisks indicate the statistical significance of the correlation coefficient: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

value of mean TIA over the 546 subperiods was 0.15, which was slightly higher than the median value of X_5 calibrated for models MUOpt5H and MU6H (0.11 and 0.10, respectively). The similarity between calibrated X_5 and mean TIA was confirmed by the correlation coefficients, which reached 0.65–0.67 for models MUOpt5H and MU6H ($p < 0.001$). This suggests that X_5 acts as the catchment mean TIA , which can be viewed as a model-based assessment of TIA . Unsurprisingly, the improvements in calibration performances were lower with very low X_5 , i.e., as we converge to the setting of the original GR4H structure.

The distributions of calibrated X_6 confirmed the rapid dynamics in urbanized catchments, by showing values higher than the originally assigned value of $X_6 = 0.1$ for the majority of cases (Figure 6.7). Correlation coefficients with TIA were moderate but significant, and X_6 tended to be higher with higher values of TIA . This suggests that catchments with a higher proportion of sealed surfaces favored the quick-flow branch to convey water to the catchment outlet.

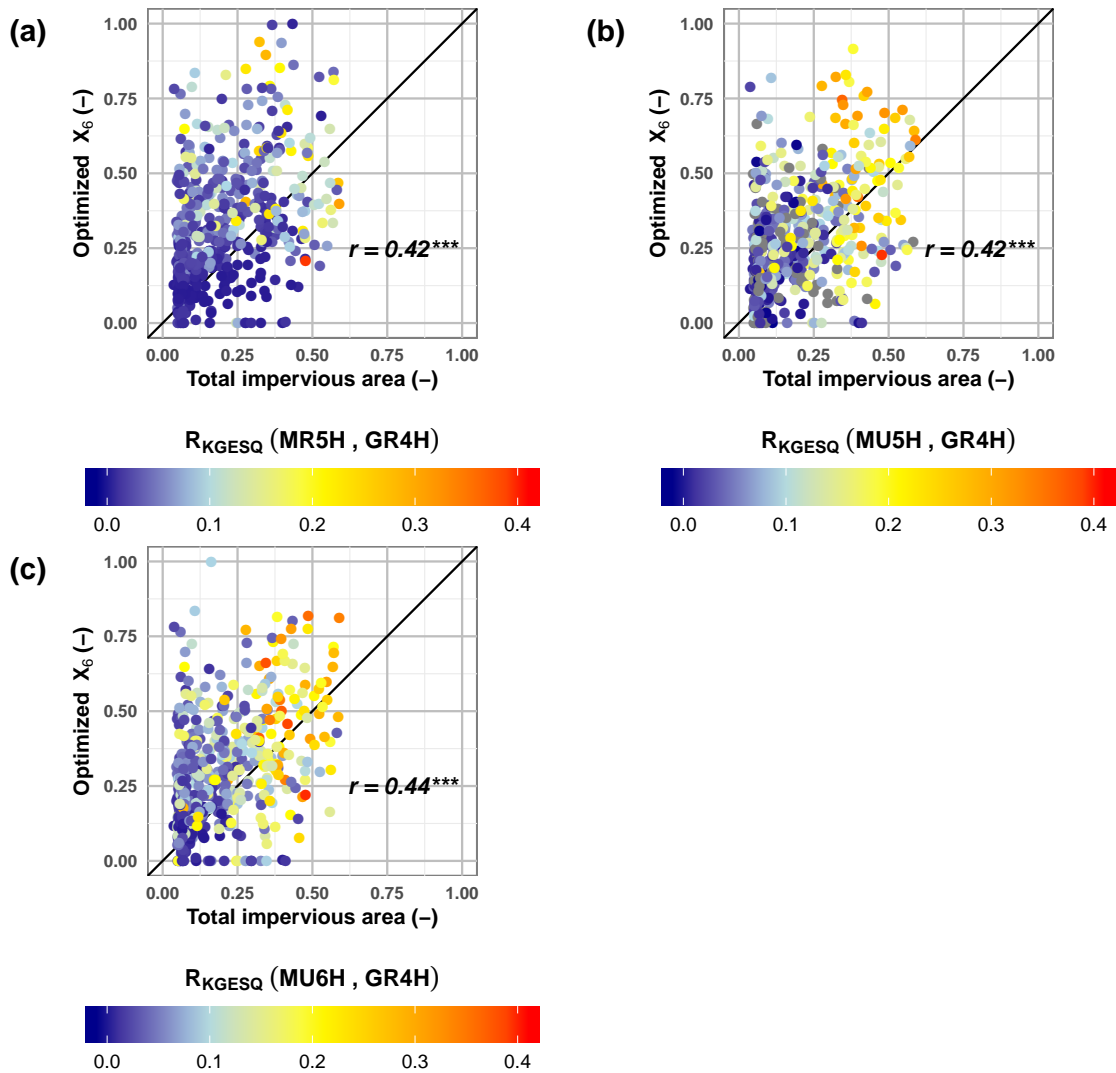


Figure 6.7. Calibrated quick-flow/slow-flow partitioning parameter X_6 vs. estimated total impervious area from land cover databases. X_6 was calibrated for the models (a) MR5H, (b) MU5H, and (c) MU6H, as indicated in Table 6.1. Colors indicate the relative index R_{KGESQ} , i.e., the improvement in calibration performances with respect to the original GR4H structure. Asterisks indicate the statistical significance of the correlation coefficient: $***p < 0.001$, $**p < 0.01$, $*p < 0.05$.

6.4.2 Test performances

The tested modifications resulted in improved performances for the majority of catchments according to all continuous assessment criteria (Figure 6.8). The median relative index varied between 0.00 ($R_{KGESQ_{wet}}$ of MUOpt5H) and 0.08 (R_{NSE} of MU5H), corresponding to a median reduction in error ranging from 0% to 14.8%. The most important improvements were registered for NSE and $KGESQ_{dry}$, which suggests that the tested modifications were more beneficial for reproducing the high flows and the catchment response amid dry conditions. Exploiting the information of TIA to constrain the proportion of sealed surfaces X_5 , with no additional calibrated parameters, gave better improvements than calibrating either X_5 or X_6 (MU4H vs. {MR5H, MUOpt5H}). The best overall improvements were obtained by MU5H and MU6H, whereas MUOpt5H showed fewer improvements, suggesting that a varying proportion of sealed surfaces X_5 should be accompanied by a varying quick-flow/slow-flow partitioning parameter X_6 .

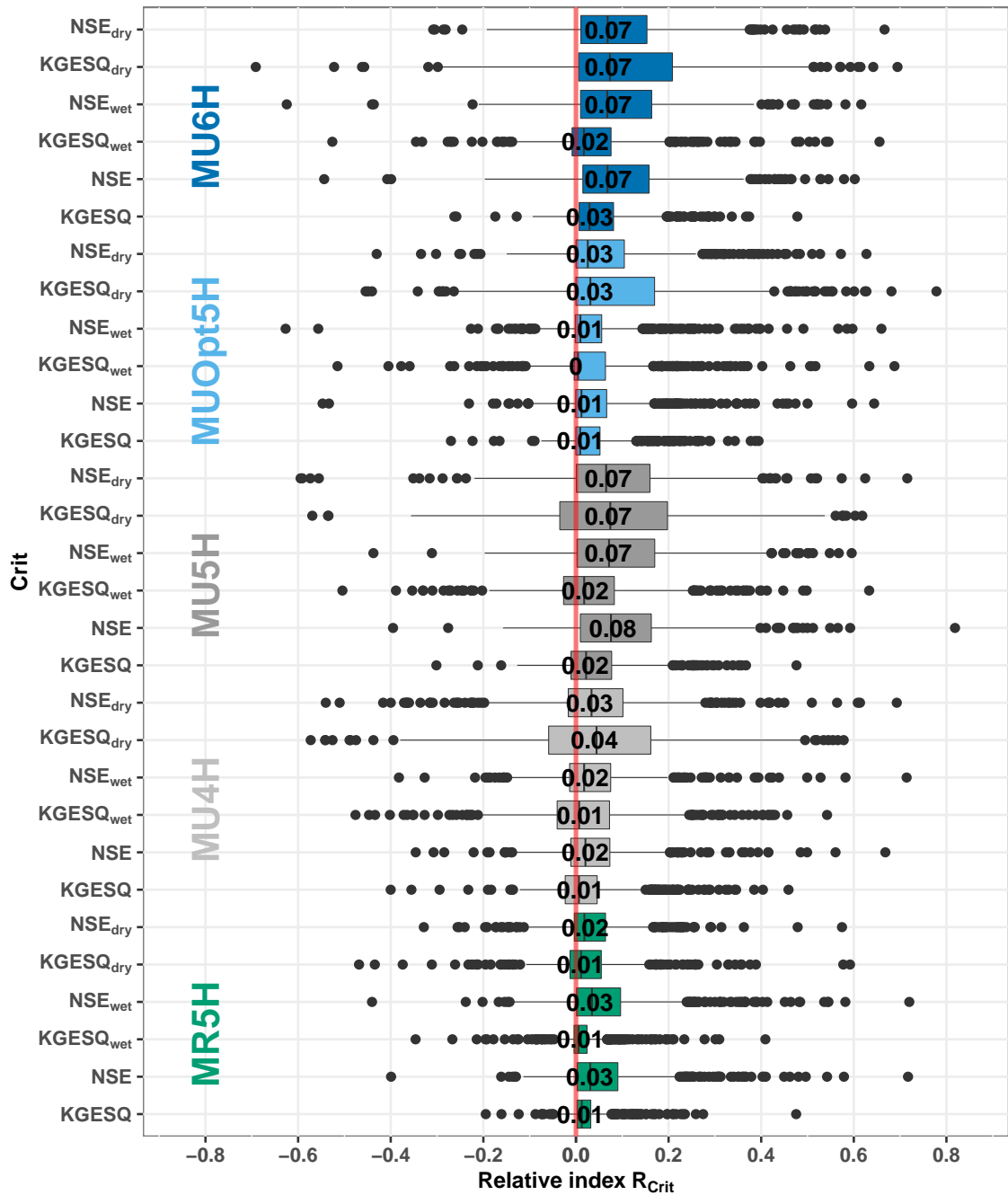


Figure 6.8. Distributions of improvements in test performances for the five tested modifications, using six criteria for continuous assessment. Boxes are delimited by the first and third quartiles. Values indicate the median relative index R_{Crit} for each criterion $Crit$. The red line constitutes the border between improvement (on the right side) and degradation (on the left side) relatively to the original model structure GR4H.

Table 6.2. Median values of event-based assessment criteria for the six tested models over 45,025 events. e_{Qp} indicates the error in estimating the event peak flow (ideal value: 0). e_{tp} indicates the error in estimating the timing of the event peak flow (ideal value: 0). VE is the volumetric efficiency, which is the error in estimating event runoff depths (ideal value: 1). The best performance for each criterion is highlighted in bold.

Criterion		Model					
		GR4H	MR5H	MU4H	MU5H	MUOpt5H	MU6H
e_{Qp}	Signed	-37.6%	-31.1%	-28.8%	-23.6%	-25.5%	-24.2%
	Absolute	51.3%	46.0%	45.1%	39.7%	44.2%	40.2%
e_{tp}	Signed	0.0%	-8.3%	0.0%	0.0%	0.0%	0.0%
	Absolute	17.9%	20.0%	15.8%	16.7%	16.7%	16.7%
	# of cases when $e_{tp} = 0$	11,925	10,721	13,275	12,007	12,580	11,808
VE		0.471	0.479	0.506	0.508	0.508	0.509

Concerning event-based assessment, the majority of modifications improved the performances over the 45,025 events, as suggested by the median event-based performances in Table 6.2. The overall superiority of MU5H and MU6H was confirmed in comparison with GR4H and the remaining tested modifications. In general, models tended to underestimate the event peak flow ($e_{Qp} < 0$). For the original model structure (GR4H), this underestimation reached on average 37.6% of the peak flow, but the error was reduced to roughly one fourth of the peak flow thanks to the inclusion of a sealed surface proportion (fixed at TIA or calibrated). Absolute error in peak flow estimation using GR4H was approximately 51%, and was reduced to around 40%. The error in the timing of peak flow e_{tp} suggested that GR4H behaved as well as the modified versions, meaning there was no significant improvement in predicting the timing of peak flow. The median absolute error was 17.9% of the delay between the beginning and the peak flow moments for GR4H, and these errors were slightly reduced to a median of 15.8% when the proportion of sealed surfaces was included. GR4H successfully predicted the timing of peak flow for 11,925 events (26.5% of events). This number was increased by MU4H (13,275), MU5H (12,007), and MUOpt5H (12,580), but they were all less than 30% of the total number of events. Finally, the median volumetric efficiency VE was improved by including the proportion of sealed surfaces, reaching a median value of 0.51 for all tested modifications, except for MR5H. This result confirms the usefulness of bypassing the soil moisture-accounting reservoir in order to better estimate event hydrographs.

The binomial test indicated that all alternative versions were statistically better than the original model structure, except for the estimation of the timing of peak flow (Figure 6.9). As shown above (Figure 6.8), the greatest improvements concerned $KGESQ$ in dry periods and NSE over the whole test period as well as in wet and dry conditions. Although the improvements in median $KGESQ$ were somewhat low, the number of cases at which this improvement was

obtained was high, especially for MU6H and MR5H (85% and 83% of cases, respectively). Conversely, improvements in median NSE_{dry} were important, attaining an increase of +0.13, with increased NSE for more than two thirds of the cases and for all tested modifications. A reduction in the error of peak flow estimation was obtained for nearly 59%–66% of events, whereas the volumetric efficiency was improved for 55%–60% of events. On the other hand, the tested modifications resulted in a similar estimation of the timing of the peak flow for the majority of events (range: 40%–59%), and improvements were registered only for approximately 26%–33% of events. The overall continuous and event-based performances are in favor of MU5H as an alternative model for urbanized catchments.

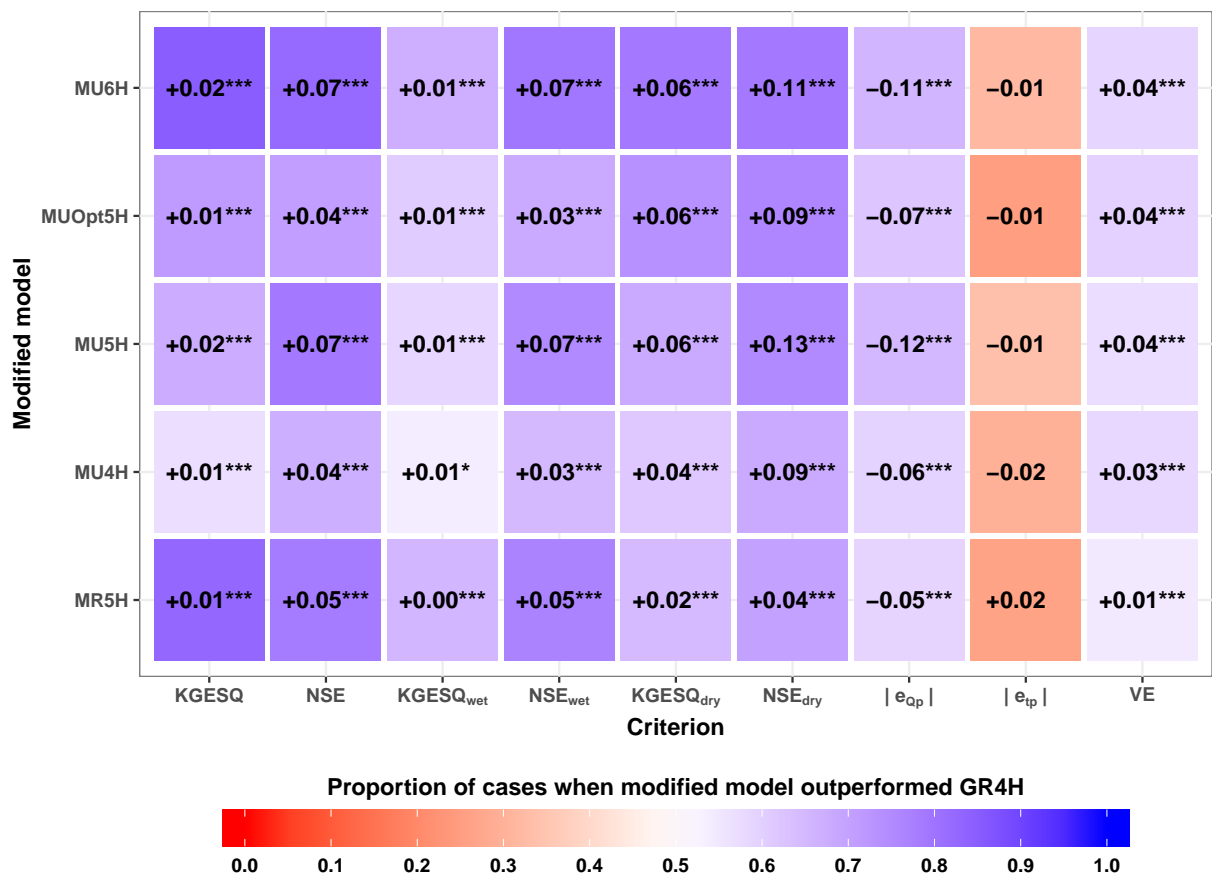


Figure 6.9. Summary of the improvements in median assessment criteria. Asterisks indicate the statistical significance of the resulting improvements using the binomial test, where *** indicates a statistical significance for a p -value threshold of 0.001 and * for a p -value threshold of 0.05. Colors indicate the proportion of cases when the modified model outperformed the original model GR4H. Note that improvements are obtained with positive differences in median values for all criteria except the absolute values of e_{Qp} and e_{tp} .

6.4.3 Were the improvements greater with increasing urbanization level? Investigation of the continuous assessment performances of MU5H

The higher the urbanization level, the better the improvements, as suggested by the distributions of the relative index R_{Crit} across TIA for the continuous assessment criteria in Figure 6.10. Improvements in $KGESQ$ were important only for TIA above 30%, whereas NSE was improved even for lower TIA levels and more importantly for higher TIA . Focusing on dry and wet conditions, improvements were less homogeneous, especially during dry conditions for all TIA levels. Overall, these results indicate that reproducing the catchment response was improved mostly for intensively urbanized catchments, and that during dry conditions, these improvements concerned also (but to a lesser extent) less urbanized cases.

6.5 Discussion and conclusion

6.5.1 Comparison with existing model testing experiments

In this study, we have shown an example of improving a conceptual model by modifying step by step a non-urban model structure to account for widely recognized urbanization features. Our study was motivated by the need for a robustly tested, simple model structure that enables a straightforward catchment-scale prediction of urbanization impact on catchment hydrological behavior. Simple modeling tools with possible straightforward assessment of urbanization impact exist in the literature, such as the Curve Number method (Huang *et al.*, 2008a; Isik *et al.*, 2013), but its use for continuous applications is impeded by inconsistencies in its formulation (Michel, Andréassian, and Perrin, 2005), and its ability to predict the impact of land-use changes (including urbanization) on catchment response is generally unverified (Ogden *et al.*, 2017). Conversely, little effort was made to robustly develop and test models over a wide variety of urban sprawl situations, which emphasizes the novelty aspect of this study.

Our tests arise in the context of flourishing model comparison studies and model testing frameworks (Clark, Kavetski, and Fenicia, 2011; Boer-Euser *et al.*, 2017; Fenicia, McDonnell, and Savenije, 2008; Knoben *et al.*, 2019; Esse *et al.*, 2013), where a relative assessment of multiple working hypotheses is accomplished to explain the rainfall-runoff relationship of a single or several catchments. Our strategy followed a data-based, top-down approach of model development (Jothityangkoon, Sivapalan, and Farmer, 2001; Massmann, 2020; Young, 2003), where learning from data was accomplished in two directions: (i) attempting physically sound modifications that are in line with the specificities of the behavior of urbanized catchments (Gupta, Wagener, and Liu, 2008; McMillan *et al.*, 2011), and (ii) evaluating the

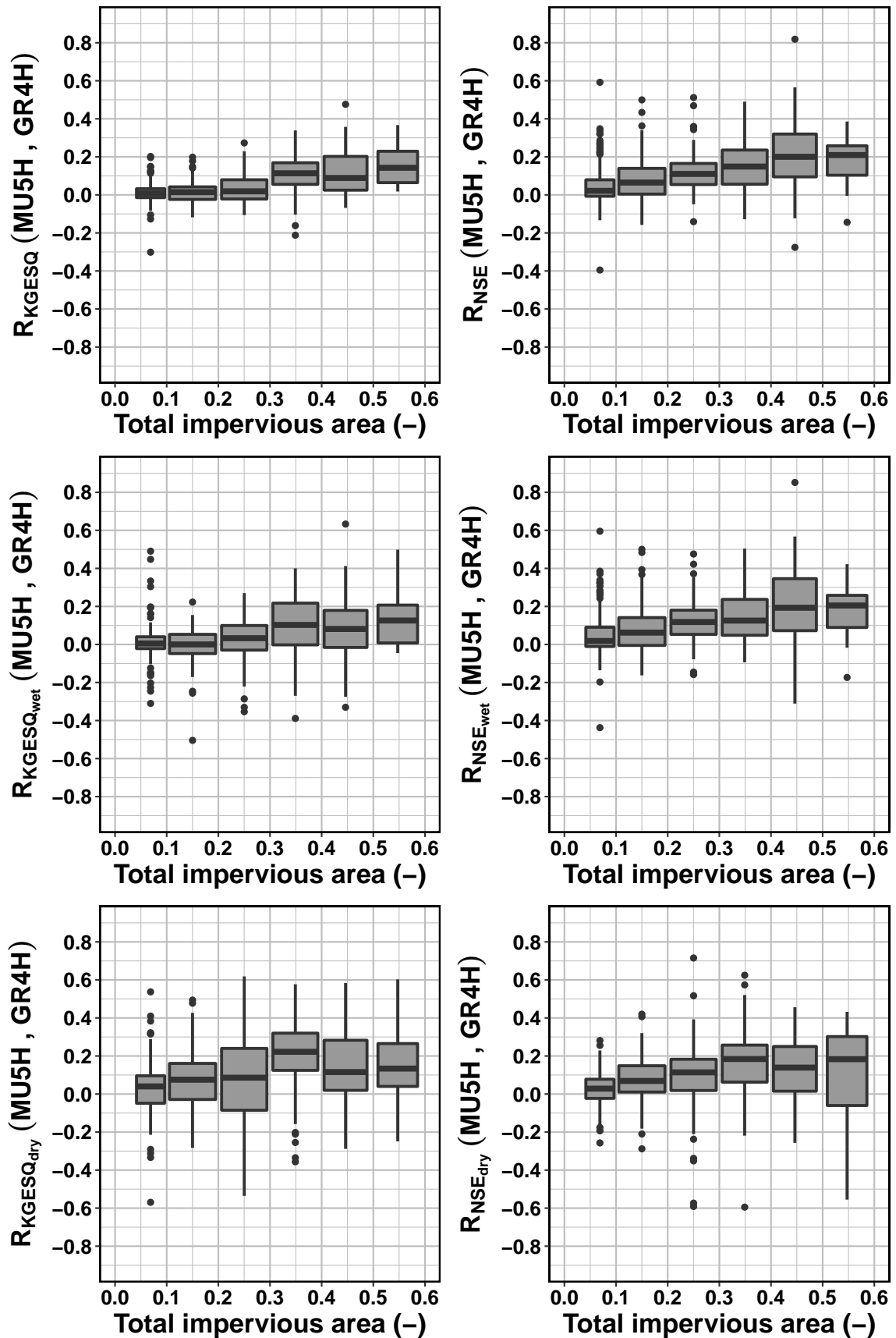


Figure 6.10. Improvements in reproducing the catchment response function of the total impervious area of the catchment during the test subperiod. Improvements are shown for the MU5H model, and are expressed using the relative index R_{Crit} (Equation 6.3.8).

relevance of these modifications in reproducing the observed response of a large sample of urbanized catchments.

We aimed at adapting the model structure by including information about urban sprawl from land cover databases, which is similar to the work by Gharari *et al.* (2014), who attempted to constrain model parameters by topography-derived information, and more particularly the model developments by Kjeldsen, Miller, and Packman (2013). In our case, including a proportion of sealed surfaces and calibrating the quick-flow/slow-flow partitioning parameter yielded significant improvement of the model over the urbanized catchments. The proportion of sealed surfaces is perhaps the most basic feature that is accounted for in urban models (e.g., Bronstert, Niehoff, and Bürger, 2002; Dotto *et al.*, 2011; Furusho, Chancibault, and Andrieu, 2013; Lhomme, Bouvier, and Perrin, 2004). Yet, confronting the hydrologically based, calibrated proportion of sealed surfaces with an estimated proportion from land cover datasets had not been done before, to our knowledge, for catchments with diverse sizes and *TIA* levels.

Another original facet of this study—perhaps the most important one—is the use of a large sample of urbanized catchments to test whether the attempted modifications were essential to better reproduce the rainfall-runoff relationship, which was needed for robust testing of hypotheses to explain the catchment-scale behavior of urbanized catchments. The use of large samples of catchments is advocated for many reasons (Addor *et al.*, 2020; Gupta *et al.*, 2014), among which are: (i) setting hard tests for models (Andréassian *et al.*, 2009), which is beneficial for detecting the weaknesses of models, and increasing their credibility (Klemeš, 1986). Models for urbanized catchments were mostly tested using small samples of catchments, for which detailed monitoring was provided, but this undermines their use for solid scenario studies. (ii) Using a large sample helps to also design a minimal set of hypotheses that could be verified for many places and for many conditions, thus advancing the quest for catchment-scale hydrological laws through a top-down approach (McDonnell *et al.*, 2007). This minimalism was also accentuated by the use of the split-sample test; one can see that adding more and more parameters improves calibration (Figure 6.5), but this is generally not the case for test performances (as suggested by MR5H, MUOpt5H, and MU6H in Figure 6.8).

We implemented our evaluation by using a panel of evaluation criteria. Model evaluation remains a difficult exercise as there is no perfect performance criterion (Krause, Boyle, and Bäse, 2005; Legates and McCabe, 1999), and focusing on only one evaluation criterion could mislead the modeler's choice (Schaeffli and Gupta, 2007). To overcome this issue, we used a set of two different criteria, *NSE* and *KGESQ*, and conducted an assessment during wet and dry conditions, in addition to an event-based assessment. All criteria yielded convergent conclusions.

A persistent issue with every model modification is deciding whether the corresponding improvements were significant or not. Specifically, it might appear tricky to interpret that, e.g.,

an increase in median $KGESQ$ of only 0.01 is a significant improvement. First, a comparison with existing studies can guide the interpretation. Le Moine *et al.* (2007) compared a number of strategies to account for groundwater–surface water exchange at daily time steps within the GR4J model, and obtained an increase of 0.006–0.08 in median NSE applied to square root values of simulated and observed streamflow depths, but for a larger number of 2080 cases of validation over 1040 catchments. Here, at the hourly time step, an improvement of 0.13 in median NSE was obtained for the 546 cases by MU5H and MU6H. Another example is shown by Ficchi, Perrin, and Andréassian (2019), who obtained an improvement of 0.011 in median KGE over a sample of 240 catchments, which is lesser than the median improvements obtained here in terms of $KGESQ$. Second, we followed the binomial test proposed by Fidal and Kjeldsen (2020). Our results suggest that the attempted modifications led to statistically significant improvements, in contrast to Fidal and Kjeldsen (2020), who found that their urban model did not significantly outperform the rural model for a sample of 27 catchments. This could be attributed to the sample size, or the modeling time step, because using a daily time step could have masked model improvements. Note that the majority of catchments used in this study had a characteristic response time of less than 24 h (251 out of 273 catchments).

6.5.2 Potentials of newly added parameters and further improvements

We found that calibrated X_5 was highly correlated with mean TIA for the majority of cases (Figure 6.6). The coefficient of determination R^2 reached 0.45, which is remarkable compared with reported results in regionalization studies (Merz and Blöschl, 2004; Saadi, Oudin, and Ribstein, 2019). This is promising for a hydrologically based evaluation of estimated measures of urbanization from land cover. The quick-flow/slow-flow partitioning parameter X_6 showed significant but moderate correlation values, which suggests that other factors should be included to better explain its variability. By regionalizing X_5 and X_6 using TIA and other urban landscape measures (Oudin *et al.*, 2018), one would be able to evaluate the effect of future urban planning configurations (expressed in terms of changes in TIA and urban landscape measures) on catchment response, using either MU5H with a regionalized X_6 , or MU6H with regionalized X_5 and X_6 .

Our study was limited by the use of only one *a priori* estimated measure of urbanization (i.e., TIA), which has the merits of being simple to compute and relatively homogeneous over both France and the United States. This has led to advances in the production component. Future work should consider refining the routing component, where lie most of the controversies regarding the impact of urbanization (Oudin *et al.*, 2018), particularly on low flows (Braud *et al.*, 2013). Thus, future modifications should address the features related to subsurface flow and groundwater–surface water interactions in urbanized areas (as in Hamel and Fletcher, 2014). An attempt was illustrated by Furusho, Chancibault, and Andrieu (2013), by including

the interaction between groundwater and sewerage systems, but a similar implementation would require a detailed description of the sewerage system of each catchment. Another way of improving the model structure would be to constrain the quick-flow/slow-flow partitioning parameter X_6 using information about the directly connected impervious area (Ebrahimian, Wilson, and Gulliver, 2016). This has the potential of a model-based estimation of the effective impervious area (Alley and Veenhuis, 1983; Boyd, Bufill, and Knee, 1993), which accounts for the routing processes in pervious areas.

Synthèse du chapitre

Nous avons testé des modifications pas-à-pas de la structure du modèle de départ pour (1) prendre plus explicitement en compte la présence de surfaces urbaines dans le bassin versant, et (2) obtenir une meilleure reproduction de la relation pluie-débit. Ces modifications ont été testées sur un large échantillon de 273 bassins versants, pour lesquels le taux moyen d'imperméabilité totale *TIA* a varié entre 5% et 59%. Ce chapitre a montré que la prise en compte du ruissellement sur les surfaces imperméables, accompagnée d'un partage variable entre écoulement lent et écoulement rapide, a nettement amélioré les performances par rapport au modèle de départ. Cette amélioration a été bénéfique surtout pour la reproduction des débits forts et la simulation de la réponse hydrologique en période sèche. Comparés avec 41 modifications testées (Annexe E), les deux modèles MU5H et MU6H ont obtenu les meilleurs résultats en moyenne. Par ailleurs, les nouveaux paramètres introduits dans la structure du modèle ont montré des relations assez fortes avec l'imperméabilité du bassin versant, ce qui montre indirectement que les modifications de structure sont utilisées à bon escient. Ce résultat ouvre des perspectives intéressantes pour l'utilisation du modèle afin d'évaluer l'impact de l'urbanisation sur le comportement hydrologique à partir de schémas d'aménagement du territoire.

CONCLUSION GÉNÉRALE

Les modèles hydrologiques sont des outils incontournables pour quantifier et prédire l'impact de l'urbanisation sur le comportement hydrologique d'un bassin versant. La revue de la littérature a montré une grande diversité de modèles existants pour les bassins urbanisés. Malgré cette diversité, les développements de ces modèles n'ont concerné que de faibles échantillons de bassins versants, ce qui ne permet pas d'établir un diagnostic sur leur robustesse ou d'en faire usage pour prédire l'impact hydrologique de futurs schémas d'urbanisation. Cette thèse a comme objectif de chercher une structure de modèle testée sur un large échantillon de bassins versants présentant différentes situations hydroclimatiques et d'urbanisation. Cette recherche a été entreprise en partant d'une structure de modèle relativement simple pour ensuite la modifier afin d'améliorer la prise en compte des spécificités de l'urbanisation.

Principaux résultats

La revue de la littérature proposée dans cette thèse soulève un paradoxe sur l'état des connaissances de l'impact hydrologique de l'urbanisation et sur la façon d'en tenir compte en modélisation hydrologique : malgré les résultats divergents de certaines études sur les questions scientifiques posées, il existe un « sens commun » important sur l'impact hydrologique de l'urbanisation. Ce « sens commun » s'est traduit tout au long de ce travail par une mise en question de l'originalité de la thèse et des résultats montrés au cours des différents chapitres par rapport à *ce qu'on connaissait déjà*. Les commentaires des relecteurs sur les travaux communiqués (articles soumis, posters dans les conférences) posaient presque systématiquement la question de l'originalité de ces travaux car on savait presque tout sur l'impact de l'urbanisation sur les processus hydrologiques ; il y a une pléthore de modèles pour les milieux urbanisés ; certains de ces modèles existants sont très *sophistiqués* et permettent une meilleure prise en compte des spécificités de l'urbanisation que celle du modèle que nous étions en train de chercher ; et l'on ne peut pas représenter tout simplement un bassin urbanisé par un modèle global ! Ces commentaires nous ont poussés à argumenter sur l'importance de nos travaux, au-delà d'un exercice classique de traitement d'un gros volume de données pour aboutir à des résultats déjà connus. Pour ces raisons, nous allons dédier cette première partie de la conclusion à montrer en quoi les résultats de cette thèse sont importants vis-à-vis des questions que nous avons posées au chapitre bibliographique et vis-à-vis de la méthodologie que nous avons mise en place pour y aboutir.

Collecte d'un large échantillon de données

Nous avons fait le choix de travailler sur un large échantillon de bassins versants urbanisés, ce qui est nécessaire pour pouvoir répondre aux controverses liées à l'impact de l'urbanisa-

tion sur le comportement hydrologique. Ces controverses ont été signalées notamment par les travaux de thèse de Salavati (2015), la revue bibliographique de Redfern *et al.* (2016) et dans l'article d'Oudin *et al.* (2018). En outre, le contexte actuel des travaux de modélisation hydrologique recommande de faire un effort de collecte d'un large volume de données pour pouvoir généraliser les conclusions et pour permettre un développement plus robuste des outils de modélisation (Addor *et al.*, 2020; Gupta *et al.*, 2014). Si ces recommandations sont appliquées par plusieurs travaux à partir des années 2000, nous avons constaté un manque de larges échantillons de bassins versants urbanisés. Soit les bassins urbanisés n'ont jamais fait l'objet d'un travail de collecte d'un large échantillon de données (c'est-à-dire sur différents lieux), soit ils ont été écartés des échantillons existants à cause de l'influence anthropique. Le caractère unique du paysage d'un lieu par rapport à un autre (Beven, 2000), rendu encore plus unique par la présence de formes urbaines plus complexes que le paysage non urbanisé, peut être l'une des raisons de l'absence d'un large échantillon de bassins versants urbanisés. Pour répondre à ce besoin, nous avons préparé des séries hydroclimatiques au pas de temps horaire pour un échantillon de 2225 bassins versants situés en France et aux États-Unis, que nous avons caractérisés d'un point de vue hydroclimatique, morphologique, géopédologique et d'occupation du sol. Nous avons montré que l'échantillon contient une diversité de situations climatiques et d'urbanisation. S'il est principalement composé de bassins versants non urbanisés, il contient néanmoins un bon nombre de bassins urbanisés dont certains ont connu un gradient d'urbanisation relativement important (mais trop limité pour une étude temporelle) sur la durée de disponibilité de leurs données hydroclimatiques (cf. Figure 2.11). L'inclusion d'un échantillon témoin de bassins non urbanisés a été exigée dans un cadre méthodologique comparatif plus robuste permettant (i) de mieux illustrer les différences de comportement hydrologique, et (ii) de mieux juger les améliorations par rapport au modèle hydrologique de départ. Cette inclusion a permis de pallier le manque de profondeur temporelle requise pour disposer d'un gradient d'urbanisation suffisamment large (c'est-à-dire relatant tous les différents stades de l'urbanisation) sur un seul bassin versant : ce gradient a été retrouvé à partir d'un large nombre de situations d'urbanisation mais dans différents lieux (*trading space for time*). Néanmoins, cet échange de l'espace contre le temps présente certaines limites à cause de l'existence d'autres contrastes régionaux (autre le niveau d'urbanisation) qui rentrent en jeu dans la variabilité du comportement hydrologique (éventuellement des performances du modèle), ce qui rend plus difficile l'attribution par exemple des différences hydrologiques aux contrastes d'urbanisation.

Différences hydrologiques entre un bassin urbanisé et un bassin non urbanisé

Si on veut développer un modèle pour prédire l'impact d'une urbanisation future sur le comportement hydrologique d'un bassin versant, on doit d'abord comprendre ce qui fait

la spécificité (d'un point de vue hydrologique) des bassins urbanisés par rapport aux bassins non urbanisés. Ainsi, une des préoccupations principales de cette thèse était de répondre aux questions suivantes : (1) qu'est-ce qui rend un bassin urbanisé hydrologiquement différent d'un bassin non urbanisé? (2) Si des différences existent, sous quelles conditions deviennent-elles significatives? En particulier, existe-il un seuil d'urbanisation au-delà duquel un bassin urbanisé devient hydrologiquement différent d'un bassin non urbanisé?

Par rapport à la première question, nos résultats confirment en grande partie ceux de la littérature, à savoir une augmentation du ruissellement, une diminution des apports de base et une accélération de la réponse hydrologique sur les bassins urbanisés (Fletcher, Andrieu et Hamel, 2013; McGrane, 2016; Jacobson, 2011; Salvadore, Bronders et Batelaan, 2015). Ce que rajoute notre démarche par rapport à ce qui existait déjà est la quête de conclusions les plus générales possibles en utilisant un large échantillon. Un deuxième trait d'originalité est d'identifier, à partir de séries pluie-débit, des lois de comportement hydrologique des bassins en fonction de plusieurs facteurs, notamment l'imperméabilité (facteur prépondérant reflétant le couvert urbain) et l'humidité du sol (facteur prépondérant reflétant le couvert non urbain). Plusieurs études ont souligné que le rôle de l'humidité du sol dans un bassin urbanisé n'est pas élucidé (voir Cristiano, Veldhuis et Giesen, 2017; Smith *et al.*, 2013). D'autres études ont suggéré que c'est au-delà d'un certain seuil d'imperméabilité que le rôle de l'humidité du sol devient sans importance, fait vérifié sur quelques cas seulement (Miller et Hess, 2017; Zhou *et al.*, 2017). Il était donc légitime de se poser la question sur le rôle de l'humidité du sol dans la réponse d'un bassin urbanisé, vu la variabilité de la réponse en milieu urbanisé selon la saison (Sillanpää et Koivusalo, 2015; Zhou *et al.*, 2017). D'ailleurs même à l'échelle locale, les surfaces considérées traditionnellement comme complètement imperméables peuvent avoir des comportements hydrologiques similaires à des surfaces non urbaines (Redfern *et al.*, 2016). Nos résultats ont montré que même à des niveaux d'urbanisation importants (fraction de couvert urbain à 100%), la réponse du bassin (indiquée par le coefficient de ruissellement¹⁶) était influencée par son état d'humidité du sol. Il est vrai que des études précédentes prenaient déjà ce facteur dans leurs modèles hydrologiques des milieux urbanisés, comme la relation empirique proposée par Brun et Band (2000), les tests de sensibilité de Zhang et Shuster (2014), ou encore la méthode du Curve Number... sauf qu'il manquait de vérifier et surtout de valider ces relations sur un grand nombre de cas d'urbanisation!

Un grand débat se fait encore sur la deuxième question à propos de l'existence d'un seuil d'urbanisation au-delà duquel le comportement hydrologique change significativement. Certains auteurs ont discuté de la pertinence de définir un tel seuil, revendiquant plutôt un continuum entre les bassins non urbanisés et les bassins urbanisés sur lequel le changement

¹⁶Nous avons également regardé l'évolution du Curve Number (CN), calé à partir des événements, en fonction de l'imperméabilité et de l'humidité du sol. Cette évolution a montré sa dépendance à l'état d'humidité du sol et à l'imperméabilité du bassin versant. En revanche, CN devient insensible aux conditions d'humidité antérieures quand le taux moyen d'imperméabilité totale dépasse un certain seuil ($TIA \geq 35\%$). De par sa formulation, CN était très dépendant de l'intensité des événements pluviométriques par rapport au coefficient de ruissellement.

hydrologique n'est pas aussi abrupt qu'on le pense (Brabec, Schulte et Richards, 2002). Le fait de travailler sur un nombre limité de bassins versants pose problème pour définir ou discuter d'un tel seuil. Probablement en conséquence de ces problèmes méthodologiques, Salvadore, Bronders et Batelaan (2015) ont constaté que ce seuil peut varier de 4% à 98% en termes de taux d'imperméabilité totale du bassin, selon les études recensées. Bénéficiant d'un large échantillon, nous avons montré qu'au-delà d'un taux moyen d'imperméabilité totale *TIA* de 5% (Figure 3.5, Chapitre 3), l'effet de l'imperméabilité devient significatif sur la réponse hydrologique. En revanche, l'évolution du changement semble plutôt graduelle (c'est-à-dire non abrupte, sans discontinuité importante). Dans le Chapitre 4, nous avons d'abord considéré un seuil de 10% (en se basant sur la littérature, et en effectuant quelques tests à des seuils de 5% et 20%) pour sélectionner les bassins urbanisés et les bassins non urbanisés, en écartant tous les bassins ayant un *TIA* entre 5% et 10%. Bien que nous ayons constaté des différences dans le comportement hydrologique sur la base du taux de l'imperméabilité totale, les résultats nous suggèrent que l'intervalle des comportements hydrologiques des bassins non urbanisés est suffisamment large pour comprendre les comportements hydrologiques de leurs bassins voisins urbanisés, excepté le ratio de la flashiness de débit sur la flashiness de précipitation (Figure 4.4), dont la plage des valeurs obtenues sur les bassins urbanisés dépasse largement les limites des valeurs calculées sur les bassins non urbanisés. Nous confirmons donc les résultats de la littérature dans le sens où on *peut* définir un seuil de *TIA* (en l'occurrence, 5%) au-delà duquel on peut constater un impact significatif de l'urbanisation, mais nous n'avons pas pu généraliser une frontière comportementale (c'est-à-dire en termes de signatures hydrologiques) qui permette de définir *nettement* ce qui fait un bassin urbanisé d'un point de vue hydrologique.

Nous avons également abordé le rôle des caractéristiques de l'événement (intensité et volume de précipitations) et du bassin (à savoir le climat, la morphologie ou la géopédologie) dans la variabilité de la réponse hydrologique sur une large plage de situations d'urbanisation, question qui n'a jamais été traitée rigoureusement dans la littérature. Malheureusement, la hiérarchisation de ces caractéristiques a montré qu'il est difficile de trancher sur celles qui sont les plus déterminantes, soulignant le caractère multi-dépendant de la réponse hydrologique.

Test d'un modèle hydrologique sur un large échantillon de bassins versants urbanisés

Le débat sur les différences hydrologiques entre les bassins urbanisés et non urbanisés montre que l'urbanisation ne supprime pas les processus habituellement rencontrés en milieu non urbanisé mais tend plutôt à les modifier, ce qui est manifesté en partie par le rôle persistant de l'humidité du sol sur les bassins urbanisés. De plus, les concepts et les équations servant à représenter les processus hydrologiques sont similaires entre les modèles utilisés en milieu

urbanisés et ceux appliqués dans les milieux non urbanisés (Fletcher, Andrieu et Hamel, 2013; Salvadore, Bronders et Batelaan, 2015). Le modèle conceptuel de départ GR4H prenant déjà en compte l'humidité du sol dans sa fonction de production, nous nous sommes posés la question de son aptitude à reproduire la relation pluie-débit sur les bassins urbanisés, même s'il ne tient compte (explicitement) d'aucune spécificité de l'urbanisation (Chapitre 4). Pour généraliser les résultats, nous avons élargi le test à IHACRES, une autre structure conceptuelle utilisant différentes équations avec un niveau de paramétrisation différent de GR4H. On s'attendait à obtenir des performances dégradées de GR4H et IHACRES sur les bassins urbanisés. Or, GR4H a conservé le même niveau de performances que sur les bassins non urbanisés et IHACRES a même mieux fonctionné sur les bassins urbanisés que sur les bassins non urbanisés. Cela indique que l'on peut reproduire de manière satisfaisante le comportement hydrologique d'un bassin urbanisé sans représenter explicitement les spécificités de l'urbanisation au sein du modèle. En revanche, les distributions des paramètres ont été impactées et ont traduit les différences comportementales entre les bassins urbanisés et non urbanisés. Notons que ces distributions n'étaient pas nettement différentes et que plusieurs zones de valeurs de paramètres ont été couvertes à la fois par les bassins urbanisés et leurs voisins non urbanisés. Nous avons approfondi l'analyse des deux structures en comparant leurs flux internes, ce qui a montré que les deux modèles divergent significativement dans la représentation de la relation pluie-débit des bassins fortement urbanisés en comparaison avec les bassins faiblement urbanisés.

Ce résultat assez étonnant nous suggère qu'on peut se contenter de la structure de départ et de chercher une manière d'exprimer (régionaliser) ses paramètres en fonction d'une ou plusieurs mesures de l'urbanisation. Ainsi, une méthode de régionalisation des paramètres du modèle GR4H a été entreprise pour exprimer ses paramètres en fonction de plusieurs caractéristiques physiques du bassin, notamment le niveau d'urbanisation (Chapitre 5). Cette régionalisation a permis d'obtenir des relations de régression non linéaires (par le biais de l'algorithme du *random forest*) qui étaient plus performantes que les méthodes de transfert spatial de paramètres. Sans souligner le nombre important des bassins versants utilisés dans ce travail, la régionalisation d'un modèle fonctionnant au pas de temps horaire n'est pas abordée abondamment dans la littérature (Razavi et Coulibaly, 2013), surtout dans un contexte de prédiction de l'impact de changement d'occupation du sol. En outre, nous avons tenté d'inclure les caractéristiques climatiques, morphologiques et géopédologiques dans cette étude pour prendre en compte la particularité paysagère de chaque bassin urbanisé dans la relation qui lie son urbanisation avec son comportement hydrologique, mais le fait que les relations de régression obtenues étaient faiblement sensibles à la métrique utilisée du niveau d'urbanisation nous a empêchés de les considérer pour la suite.

Il s'agit donc d'un long travail de construction pour finalement se rendre compte de la nécessité de modifier la structure du départ pour représenter plus explicitement les spécificités hydrologiques de l'urbanisation. Les résultats du Chapitre 5 nous ont suggéré de considérer

une structure plus contrainte pour que les paramètres puissent véhiculer moins de processus à la fois et améliorer leur régionalisation. Les résultats du Chapitre 3 (Figures 3.3, 3.4 et 3.5) nous ont facilité la compréhension des rôles de l'imperméabilité et de l'état d'humidité du sol, ce que nous avons traduit par une prise en compte du ruissellement sur les surfaces imperméables du bassin dans la structure du modèle. À l'issue des résultats du Chapitre 4 (Figure 4.4), nous avons testé la prise en compte des différences de dynamique entre les surfaces imperméables et les surfaces non urbanisées du bassin, en calant le paramètre de partage entre écoulement lent et écoulement rapide. Chacune de ces modifications devait être justifiée par de meilleures performances par rapport à la structure de départ, selon les principes de l'approche descendante détaillés par Mathevet (2005, pp. 39-44). En outre, ces modifications ont été testées sur un large échantillon de 273 bassins versants, pour lesquels *TIA* a varié entre 5% et 59%, ce qui permettait de maintenir une continuité entre les bassins non urbanisés et les bassins urbanisés tout en focalisant sur les bassins fortement urbanisés. Comme résultat, la prise en compte du ruissellement sur les surfaces imperméables, accompagnée d'un partage variable entre écoulement lent et écoulement rapide, a nettement amélioré les performances par rapport au modèle de départ (Chapitre 6). Cette amélioration a été bénéfique surtout pour la reproduction des débits forts et la simulation de la réponse hydrologique en période sèche. Les deux modèles MU5H et MU6H qui résultent de ces modifications ont donné les meilleurs résultats par rapport à 41 autres structures de modèles testées, comme détaillé dans l'Annexe E. Si ces modifications sont déjà prises en compte par plusieurs modèles existants, l'originalité de ces tests tient à la démarche constructive poursuivie, qui a combiné à la fois (1) un large échantillon de bassins versants, avec divers cas d'urbanisation (Andréassian *et al.*, 2009; Gupta *et al.*, 2014); (2) une analyse au préalable du fonctionnement hydrologique des bassins urbanisés, pour comprendre leurs traits comportementaux par rapport aux bassins non urbanisés (Hamel et Fletcher, 2014); (3) une structure de départ non dédiée aux environnements urbanisés (Furusho, Chancibault et Andrieu, 2013; Kjeldsen, Miller et Packman, 2013; Singh, Maheshwari et Malano, 2009), que nous avons d'abord analysée en profondeur afin d'en cerner les parties susceptibles d'être modifiées; (4) une approche de complexification pas-à-pas justifiée par l'aptitude du modèle à reproduire les écoulements observés (Ficchi, 2017; Le Moine, 2008; Kjeldsen, Miller et Packman, 2013); et (5) une prise en compte de l'évolution au pas de temps annuel du couvert urbain (cf. les modèles MU4H et MU5H). L'amélioration des performances du modèle peut sembler modeste, mais elle est conséquente en comparaison des travaux antérieurs de développement de modèles sur de larges échantillons (Ficchi, 2017; Le Moine, 2008). Les tests ont porté sur plusieurs aspects de la réponse (en continu et au cours des événements, en saison sèche et en saison humide), ce qui a confirmé la pertinence des modifications effectuées.

La structure finale (MU5H ou MU6H) répond aux trois conditions fixées au départ : (1) elle est paramétrée *a minima* et demande un volume relativement faible de données pour estimer ses paramètres; (2) elle contient un lien explicite avec certaines propriétés de l'urbanisation facilitant son utilisation pour l'étude de l'impact de scénarios futurs d'étalement urbain; et (3) elle est testée sur un large échantillon de bassins versants avec diverses caractéristiques

climatiques et d'urbanisation. Cette structure hybride réunit les avantages des approches de modélisation explicite et implicite : ses paramètres peuvent être contraints à partir de mesures du paysage tout en étant applicable sur un large échantillon de données avec des simulations satisfaisantes de débit (Euser *et al.*, 2015 ; Gharari *et al.*, 2014). Mais cela n'empêche pas de poursuivre les développements car il reste des possibilités d'amélioration de la prise en compte de l'urbanisation dans la structure du modèle, ou pour élargir l'éventail des situations climatiques et d'urbanisation dans l'échantillon des bassins versants.

Quelques perspectives

Au début de ce manuscrit, nous avons fait état des questions sociétales et scientifiques qui encadrent les travaux de cette thèse : comment aider les planificateurs urbains à décider des choix d'urbanisation ayant le moindre impact dans un cadre de développement durable ? Comment représenter un bassin versant en transition, que ce soit dans son paysage ou dans le forçage climatique qu'il subit (cf. décennie *Panta Rhei*) ? Le point où nous sommes maintenant, à la fin de ces travaux, reste encore loin des réponses à ces questions. Dans un premier temps, nous faisons un bilan des travaux à mettre en place pour arriver à une structure applicable dans le cadre d'étude de l'impact de scénarios d'urbanisation sur le comportement hydrologique. Dans un second temps, nous élaborons un croquis des différentes applications de cette structure une fois validée.

L'approche statistique suivie dans ces travaux peut être complétée par des analyses sur quelques bassins versants. **(1)** D'abord, on peut examiner les bassins versants sur lesquels les modifications ont diminué les performances du modèle, comme on a pu le constater à partir du Chapitre 6. Comprendre les raisons derrière cette diminution sera bénéfique pour une amélioration éventuelle du modèle. **(2)** Ensuite, on peut regarder l'évolution des performances sur des bassins versants qui sont des observatoires en France (par exemple, les bassins péri-urbains de l'Yzeron à Lyon, des Gohards et de la Chézine à Nantes). Cette analyse permettra de comparer des modèles relativement simples (MU5H ou MU6H) avec les modèles développés sur ces observatoires et de dresser les avantages et les inconvénients de chacune des approches. **(3)** Enfin, on peut focaliser sur des groupes de bassins versants emboîtés situés notamment aux États-Unis pour voir d'une part l'évolution d'amont en aval des caractéristiques et de l'impact de l'urbanisation, et d'autre part l'évolution des performances du modèle. Ceci permettra de traiter plus en profondeur la question de la dépendance de l'impact de l'urbanisation et des performances du modèle à l'échelle spatiale.

Les modèles MU5H et MU6H doivent être validés sur des bassins versants ayant un large gradient temporel d'urbanisation. Cet exercice est indispensable pour pouvoir appliquer ces structures de manière fiable dans l'étude de l'impact des scénarios d'urbanisation. Au

sein de l'échantillon que nous avons collecté, le nombre de bassins qui ont eu un gradient d'urbanisation important ($\Delta CPD \geq 10\%$) est de 45 bassins versants. Cet échantillon peut être complété en travaillant au pas de temps journalier pour permettre plus de profondeur temporelle au niveau des données hydroclimatiques, et en utilisant d'autres indicateurs comme la densité de population pour pallier le manque de données d'occupation du sol.

Les travaux de régionalisation du modèle montrés au Chapitre 5 peuvent être améliorés en partant de la nouvelle structure (MU5H ou MU6H) au lieu de la structure de départ (GR4H) utilisée au Chapitre 4, et en se servant de plusieurs mesures du couvert urbain. Dans ce sens, nous citons un travail qui a été effectué d'estimation de l'imperméabilité effective sur un large échantillon de bassins versants (voir Annexe D). L'objectif était de trouver une meilleure mesure du couvert urbain, hydrologiquement validée à partir des événements, et permettant la prise en compte de la densité de population, le type de surface urbaine, la connectivité et la proximité au réseau de drainage. Nous avons testé plusieurs hypothèses pour pallier le manque de données de connectivité des surfaces urbaines sur un grand nombre de cas. Malheureusement, l'imperméabilité effective estimée par cette méthode n'a pas mieux expliqué la variabilité des coefficients de ruissellement que l'imperméabilité totale *TIA*. Des alternatives existent à travers l'utilisation des mesures de fragmentation du paysage (Oudin *et al.*, 2018).

À moyen terme, on peut s'intéresser à l'amélioration de la représentation spatiale dans le bassin versant pour passer d'un modèle global à un modèle semi-distribué. En effet, les modifications testées sur le modèle GR4H peuvent être regardées comme un premier degré de spatialisation de la structure en distinguant les surfaces naturelles des surfaces imperméables. Un approfondissement dans cette direction pourra considérer des unités de réponse hydrologique (HRU).

Compte tenu de ces développements, le modèle hydrologique peut être finalement utilisé pour répondre aux enjeux scientifiques et sociétaux actuels, en combinant des scénarios d'urbanisation (ou de changement d'occupation du sol) avec des scénarios de changement climatique (Chung, Park et Lee, 2011; Karlsson *et al.*, 2016) à l'échelle d'un bassin versant urbanisé. Ce travail concernera principalement des bassins versants ayant un gradient temporel important de changement d'occupation du sol. De longues séries de données hydroclimatiques et d'occupation du sol doivent alors être mises à disposition. Actuellement, cet exercice se fait avec des modèles qui en général ne sont pas suffisamment testés pour différentes situations climatiques et d'urbanisation, ce qui rend les études de scénarios moins crédibles. Si un large panel de situations d'urbanisation a été parcouru dans cette thèse, la plage des situations climatiques devrait être étendue pour inclure des climats froids (bassins avec influence de neige) et arides. Ce travail amènera sur un modèle permettant d'étudier la combinaison des deux changements (du forçage et du paysage), ce qui serait largement bénéfique aux preneurs de décisions et aux aménageurs du territoire.

Nous finissons par une ouverture sur quelques aspects théoriques et méthodologiques de la modélisation des bassins versants en transition. Si on se rappelle de la définition d'un modèle hydrologique, on s'aperçoit qu'il combine des variables qui changent à chaque pas de temps (états du modèle) et des paramètres invariables censés traduire les caractéristiques *statiques* du bassin versant, en particulier ses propriétés physiques. L'invariabilité des paramètres du modèle censés traduire une certaine invariabilité des propriétés du bassin versant se heurte au caractère variable/en transition des bassins urbanisés. Un changement de paradigme en modélisation hydrologique est donc nécessaire en repensant cette invariabilité des paramètres. Ce changement de paradigme devrait permettre aux paramètres du modèle une élasticité rendant compte du changement dans les propriétés physiques du bassin versant. À ce titre, une méthode de calage proposée par de Vos, Rientjes et Gupta (2010) montre un cas de relaxation de l'hypothèse d'invariabilité des paramètres d'un modèle hydrologique. La structure du modèle MU5H illustre en quelque sorte ce schéma, où les 5 paramètres calés étaient invariables sur toute la période de simulation (calage ou validation), et le paramètre du taux de surfaces urbaines X_5 , fixé à TIA , changeait chaque année. En se basant sur ce schéma, on peut imaginer différents niveaux de paramétrisation des modèles hydrologiques correspondant aux échelles temporelles du changement des caractéristiques du bassin versant, c'est-à-dire **(i)** un niveau bas/invariable regroupant les paramètres statiques, reflétant des caractéristiques qui changent sur un pas de temps très long par rapport à l'échelle temporelle de l'étude, **(ii)** un niveau haut/très variable des états du modèle qui changent à chaque pas de temps, et finalement **(iii)** un niveau intermédiaire des paramètres/états qui varient selon un pas de temps commensurable à l'échelle temporelle caractéristique du changement qui s'opère sur les caractéristiques du bassin versant, comme le pas de temps annuel pour l'urbanisation, ou le pas de temps mensuel pour les cultures. Ce schéma qui semble assez simple à concevoir pourrait être difficile à valider compte tenu des problèmes inhérents à la paramétrisation des modèles hydrologiques.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ADDOR, N., *et al.* 2018. "A ranking of hydrological signatures based on their predictability in space". *Water Resources Research* 54 (11) : 8792-8812. ISSN : 1944-7973. doi :[10.1029/2018WR022606](https://doi.org/10.1029/2018WR022606). Cité pp. 136, 150, 199, 209, 210.
- ADDOR, N., *et al.* 2020. "Large-sample hydrology : recent progress, guidelines for new datasets and grand challenges". *Hydrological Sciences Journal* 65 (5) : 712-725. ISSN : 0262-6667. doi :[10.1080/02626667.2019.1683182](https://doi.org/10.1080/02626667.2019.1683182). Cité pp. 168, 250, 258.
- ADDOR, N., *et al.* 2017. "The CAMELS data set : catchment attributes and meteorology for large-sample studies". *Hydrology and Earth System Sciences* 21 (10) : 5293-5313. ISSN : 1027-5606. doi :[10.5194/hess-21-5293-2017](https://doi.org/10.5194/hess-21-5293-2017). Cité p. 168.
- AHIABLAME, L. M., B. A. ENGEL et I. CHAUBEY. 2012. "Effectiveness of low impact development practices : literature review and suggestions for future research". *Water, Air, & Soil Pollution* 223 (7) : 4253-4273. ISSN : 1573-2932. doi :[10.1007/s11270-012-1189-2](https://doi.org/10.1007/s11270-012-1189-2). Cité p. 21.
- AHN, K.-H., et V. MERWADE. 2014. "Quantifying the relative impact of climate and human activities on streamflow". *Journal of Hydrology* 515 :257-266. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2014.04.062](https://doi.org/10.1016/j.jhydrol.2014.04.062). Cité p. 37.
- ALBERTI, M., *et al.* 2007. "The impact of urban patterns on aquatic ecosystems : An empirical analysis in Puget lowland sub-basins". *Landscape and Urban Planning* 80 (4) : 345-361. ISSN : 0169-2046. doi :[10.1016/j.landurbplan.2006.08.001](https://doi.org/10.1016/j.landurbplan.2006.08.001). Cité p. 32.
- ALLEN, R. G., *et al.* 1998. "Crop evapotranspiration - Guidelines for computing crop water requirements". *FAO Irrigation and drainage paper* 56 :15. Cité pp. 74, 232.
- ALLEY, W. M., et J. E. VEENHUIS. 1983. "Effective impervious area in urban runoff modeling". *Journal of Hydraulic Engineering* 109 (2) : 313-319. doi :[10.1061/\(ASCE\)0733-9429\(1983\)109:2\(313\)](https://doi.org/10.1061/(ASCE)0733-9429(1983)109:2(313)). Cité pp. 160, 161, 252.
- ANCTIL, F., *et al.* 2004. "A soil moisture index as an auxiliary ANN input for stream flow forecasting". *Journal of Hydrology* 286 (1) : 155-167. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2003.09.006](https://doi.org/10.1016/j.jhydrol.2003.09.006). Cité pp. 146, 160.
- ANDERSON, R. M., V. I. KOREN et S. M. REED. 2006. "Using SSURGO data to improve Sacramento Model a priori parameter estimates". *Journal of Hydrology*, The model parameter estimation experiment, 320 (1) : 103-116. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2005.07.020](https://doi.org/10.1016/j.jhydrol.2005.07.020). Cité p. 198.
- ANDO, Y., K. MUSIAKE et Y. TAKAHASI. 1984. "Modelling of hydrologic processes in a small urbanized hillslope basin with comments on the effects of urbanization". *Journal of Hydrology*, Global Water : Science and Engineering The Ven Te Chow Memorial Volume, 68 (1) : 61-83. ISSN : 0022-1694. doi :[10.1016/0022-1694\(84\)90204-X](https://doi.org/10.1016/0022-1694(84)90204-X). Cité p. 136.
- ANDRÉASSIAN, V., *et al.* 2009. "HESS Opinions "Crash tests for a standardized evaluation of hydrological models"". *Hydrology and Earth System Sciences* 13 (10) : 1757-1764. ISSN : 1027-5606. doi :[10.5194/hess-13-1757-2009](https://doi.org/10.5194/hess-13-1757-2009). Cité pp. 66, 229, 237, 250, 262.
- ANDRÉASSIAN, V. 2004. "Waters and forests : from historical controversy to scientific debate". *Journal of Hydrology* 291 (1) : 1-27. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2003.12.015](https://doi.org/10.1016/j.jhydrol.2003.12.015). Cité p. 16.

- ANDRÉASSIAN, V., *et al.* 2012. "All that glitters is not gold : the case of calibrating hydrological models". *Hydrological Processes* 26 (14) : 2206-2210. ISSN : 1099-1085. doi :[10.1002/hyp.9264](https://doi.org/10.1002/hyp.9264). Cité pp. 198, 222.
- ANDRÉASSIAN, V., *et al.* 2014. "Seeking genericity in the selection of parameter sets : Impact on hydrological model efficiency". *Water Resources Research* 50 (10) : 8356-8366. ISSN : 1944-7973. doi :[10.1002/2013WR014761](https://doi.org/10.1002/2013WR014761). Cité p. 222.
- ANIM, D. O., *et al.* 2018. "Effect of urbanization on stream hydraulics". *River Research and Applications* 34 (7) : 661-674. ISSN : 1535-1467. doi :[10.1002/rra.3293](https://doi.org/10.1002/rra.3293). Cité p. 19.
- APPLEYARD, S. 1995. "The impact of urban development on recharge and groundwater quality in a coastal aquifer near Perth, Western Australia". *Hydrogeology Journal* 3 (2) : 65-75. ISSN : 1435-0157. doi :[10.1007/s10040050072](https://doi.org/10.1007/s10040050072). Cité p. 20.
- ARNOLD, C. L. J., *et C. J. GIBBONS.* 1996. "Impervious surface coverage : The emergence of a key environmental indicator". *Journal of the American Planning Association* 62 (2) : 243-258. ISSN : 0194-4363. doi :[10.1080/01944369608975688](https://doi.org/10.1080/01944369608975688). Cité pp. 18, 31, 170, 230.
- ARONICA, G. T., *et M. CANNAROZZO.* 2000. "Studying the hydrological response of urban catchments using a semi-distributed linear non-linear model". *Journal of Hydrology* 238 (1) : 35-43. ISSN : 0022-1694. doi :[10.1016/S0022-1694\(00\)00311-5](https://doi.org/10.1016/S0022-1694(00)00311-5). Cité p. 50.
- ARONICA, G. T., *et L. G. LANZA.* 2005. "Drainage efficiency in urban areas : a case study". *Hydrological Processes* 19 (5) : 1105-1119. ISSN : 1099-1085. doi :[10.1002/hyp.5648](https://doi.org/10.1002/hyp.5648). Cité p. 44.
- AWOL, F. S., P. COULIBALY *et B. A. TOLSON.* 2018. "Event-based model calibration approaches for selecting representative distributed parameters in semi-urban watersheds". *Advances in Water Resources* 118 :12-27. ISSN : 0309-1708. doi :[10.1016/j.advwatres.2018.05.013](https://doi.org/10.1016/j.advwatres.2018.05.013). Cité p. 58.
- BACH, P. M., *et al.* 2014. "A critical review of integrated urban water modelling – Urban drainage and beyond". *Environmental Modelling & Software* 54 :88-107. ISSN : 1364-8152. doi :[10.1016/j.envsoft.2013.12.018](https://doi.org/10.1016/j.envsoft.2013.12.018). Cité p. 43.
- BAKER, D. B., *et al.* 2004. "A new flashiness index : Characteristics and applications to midwestern rivers and streams". *JAWRA Journal of the American Water Resources Association* 40 (2) : 503-522. ISSN : 1752-1688. doi :[10.1111/j.1752-1688.2004.tb01046.x](https://doi.org/10.1111/j.1752-1688.2004.tb01046.x). Cité pp. 102, 141, 142, 168, 172, 173, 204.
- BARTENS, J., *et al.* 2008. "Can urban tree roots improve infiltration through compacted subsoils for stormwater management?" *Journal of Environmental Quality* 37 (6) : 2048-2057. ISSN : 1537-2537. doi :[10.2134/jeq2008.0117](https://doi.org/10.2134/jeq2008.0117). Cité p. 19.
- BAYLISS, A. C., *et al.* 2006. *URBEXT2000—A new FEH catchment descriptor : Calculation, dissemination and application*. R&D Technical Report FD1919/TR. Wallingford : Centre for Ecology & Hydrology Wallingford. Cité p. 31.
- BECK, H. E., *et al.* 2018. "Present and future Köppen-Geiger climate classification maps at 1-km resolution". *Scientific Data* 5 (1) : 1-12. ISSN : 2052-4463. doi :[10.1038/sdata.2018.214](https://doi.org/10.1038/sdata.2018.214). Cité pp. 81, 104, 137, 172, 232.
- BELLAL, M., X. SILLEN *et Y. ZECH.* 1996. "Coupling GIS with a distributed hydrological model for studying the effect of various urban planning options on rainfall-runoff relationship in urbanized watersheds". *IAHS Publication* 235 :99-106. Cité p. 44.
- BEREZOWSKI, T., *et al.* 2012. "Impact of remotely sensed land-cover proportions on urban runoff prediction". *International Journal of Applied Earth Observation and Geoinformation* 16 :54-65. ISSN : 0303-2434. doi :[10.1016/j.jag.2011.11.007](https://doi.org/10.1016/j.jag.2011.11.007). Cité p. 47.
- BERGHUIJS, W. R., *et al.* 2016. "Dominant flood generating mechanisms across the United States". *Geophysical Research Letters* 43 (9) : 4382-4390. ISSN : 1944-8007. doi :[10.1002/2016GL068070](https://doi.org/10.1002/2016GL068070). Cité p. 136.

- BERNE, A., *et al.* 2004. “Temporal and spatial resolution of rainfall measurements required for urban hydrology”. *Journal of Hydrology*, Urban Hydrology, 299 (3) : 166-179. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2004.08.002](https://doi.org/10.1016/j.jhydrol.2004.08.002). Cité p. 55.
- BERTHET, L., *et al.* 2009. “How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments”. *Hydrology and Earth System Sciences* 13 (6) : 819-831. ISSN : 1027-5606. doi :[10.5194/hess-13-819-2009](https://doi.org/10.5194/hess-13-819-2009). Cité pp. 145, 146.
- BERTHIER, E., H. ANDRIEU et J. D. CREUTIN. 2004. “The role of soil in the generation of urban runoff : development and evaluation of a 2D model”. *Journal of Hydrology*, Urban Hydrology, 299 (3) : 252-266. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2004.08.008](https://doi.org/10.1016/j.jhydrol.2004.08.008). Cité pp. 48, 161.
- BERTHIER, E., H. ANDRIEU et F. RODRIGUEZ. 1999. “The Rezé urban catchments database”. *Water Resources Research* 35 (6) : 1915-1919. ISSN : 1944-7973. doi :[10.1029/1999WR900053](https://doi.org/10.1029/1999WR900053). Cité p. 56.
- BESAW, L. E., *et al.* 2010. “Advances in ungauged streamflow prediction using artificial neural networks”. *Journal of Hydrology* 386 (1) : 27-37. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2010.02.037](https://doi.org/10.1016/j.jhydrol.2010.02.037). Cité p. 220.
- BEVEN, K. J. 2006. “A manifesto for the equifinality thesis”. *Journal of Hydrology*, The model parameter estimation experiment, 320 (1) : 18-36. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2005.07.007](https://doi.org/10.1016/j.jhydrol.2005.07.007). Cité pp. 56, 187.
- . 2012. “Beyond the primer : Predictions in ungauged basins”. In *Rainfall-runoff modelling : the primer*, 2. ed, 329-342. Chichester, UK : Wiley-Blackwell. ISBN : 978-0-470-71459-1. Cité p. 198.
- . 2001. “How far can we go in distributed hydrological modelling?” *Hydrology and Earth System Sciences* 5 (1) : 1-12. ISSN : 1027-5606. doi :[10.5194/hess-5-1-2001](https://doi.org/10.5194/hess-5-1-2001). Cité p. 197.
- . 2002. “Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system”. *Hydrological Processes* 16 (2) : 189-206. ISSN : 1099-1085. doi :[10.1002/hyp.343](https://doi.org/10.1002/hyp.343). Cité pp. 56, 228.
- . 2000. “Uniqueness of place and process representations in hydrological modelling”. *Hydrology and Earth System Sciences* 4 (2) : 203-213. ISSN : 1027-5606. doi :[10.5194/hess-4-2-203-2000](https://doi.org/10.5194/hess-4-2-203-2000). Cité p. 258.
- BEVEN, K. J., et M. J. KIRKBY. 1979. “A physically based, variable contributing area model of basin hydrology”. *Hydrological Sciences Bulletin* 24 (1) : 43-69. ISSN : 0303-6936. doi :[10.1080/02626667909491834](https://doi.org/10.1080/02626667909491834). Cité pp. 59, 112.
- BEVEN, K. J., E. F. WOOD et M. SIVAPALAN. 1988. “On hydrological heterogeneity — Catchment morphology and catchment response”. *Journal of Hydrology* 100 (1) : 353-375. ISSN : 0022-1694. doi :[10.1016/0022-1694\(88\)90192-8](https://doi.org/10.1016/0022-1694(88)90192-8). Cité p. 15.
- BHASKAR, A. S., *et al.* 2020. “Hydrologic signals and surprises in U.S. streamflow records during urbanization”. *Water Resources Research* 56 (9) : e2019WR027039. ISSN : 1944-7973. doi :[10.1029/2019WR027039](https://doi.org/10.1029/2019WR027039). Cité p. 97.
- BHASKAR, A. S., D. M. HOGAN et S. A. ARCHFIELD. 2016. “Urban base flow with low impact development”. *Hydrological Processes* 30 (18) : 3156-3171. ISSN : 1099-1085. doi :[10.1002/hyp.10808](https://doi.org/10.1002/hyp.10808). Cité pp. 33, 36.
- BHASKAR, A. S., *et al.* 2015. “Untangling the effects of urban development on subsurface storage in Baltimore”. *Water Resources Research* 51 (2) : 1158-1181. ISSN : 1944-7973. doi :[10.1002/2014WR016039](https://doi.org/10.1002/2014WR016039). Cité p. 37.
- BIAU, G., et E. SCORNET. 2016. “A random forest guided tour”. *TEST* 25 (2) : 197-227. ISSN : 1863-8260. doi :[10.1007/s11749-016-0481-7](https://doi.org/10.1007/s11749-016-0481-7). Cité pp. 209, 210.
- BLÖSCHL, G. 2017. “Debates—Hypothesis testing in hydrology : Introduction”. *Water Resources Research* 53 (3) : 1767-1769. ISSN : 1944-7973. doi :[10.1002/2017WR020584](https://doi.org/10.1002/2017WR020584). Cité pp. 7, 41.
- BLÖSCHL, G., *et al.* 2007. “At what scales do climate variability and land cover change impact on flooding and low flows?” *Hydrological Processes* 21 (9) : 1241-1247. ISSN : 1099-1085. doi :[10.1002/hyp.6669](https://doi.org/10.1002/hyp.6669). Cité p. 27.

- BLÖSCHL, G., *et al.* 2019a. “Changing climate both increases and decreases European river floods”. *Nature* 573 (7772) : 108-111. ISSN : 1476-4687. doi :[10.1038/s41586-019-1495-6](https://doi.org/10.1038/s41586-019-1495-6). Cité p. 16.
- BLÖSCHL, G., *et al.* 2019b. “Twenty-three unsolved problems in hydrology (UPH) – a community perspective”. *Hydrological Sciences Journal* 64 (10) : 1141-1158. ISSN : 0262-6667. doi :[10.1080/02626667.2019.1620507](https://doi.org/10.1080/02626667.2019.1620507). Cité p. 13.
- BLUM, A. G., *et al.* 2020. “Causal effect of impervious cover on annual flood magnitude for the United States”. *Geophysical Research Letters* 47 (5) : e2019GL086480. ISSN : 1944-8007. doi :[10.1029/2019GL086480](https://doi.org/10.1029/2019GL086480). Cité pp. 32, 35, 38.
- BLUME, T., E. ZEHE et A. BRONSTERT. 2007. “Rainfall—runoff response, event-based runoff coefficients and hydrograph separation”. *Hydrological Sciences Journal* 52 (5) : 843-862. ISSN : 0262-6667. doi :[10.1623/hysj.52.5.843](https://doi.org/10.1623/hysj.52.5.843). Cité pp. 103, 136, 145, 161, 172, 232, 298.
- BOCINSKY, R. K., D. BEAUDETTE et S. CHAMBERLAIN. 2019. *FedData : Functions to automate downloading geospatial data available from several federated data sources*. Visité le 21/05/2019. <https://CRAN.R-project.org/package=FedData>. Cité pp. 113, 142, 205.
- BOER-EUSER, T. de, *et al.* 2017. “Looking beyond general metrics for model comparison – lessons from an international model intercomparison study”. *Hydrology and Earth System Sciences* 21 (1) : 423-440. ISSN : 1027-5606. doi :[10.5194/hess-21-423-2017](https://doi.org/10.5194/hess-21-423-2017). Cité pp. 188, 233, 248.
- BONNEAU, J., *et al.* 2018. “The impact of urbanization on subsurface flow paths – A paired-catchment isotopic study”. *Journal of Hydrology* 561 :413-426. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2018.04.022](https://doi.org/10.1016/j.jhydrol.2018.04.022). Cité pp. 38, 228.
- BOOKER, D. J., et R. A. WOODS. 2014. “Comparing and combining physically-based and empirically-based approaches for estimating the hydrology of ungauged catchments”. *Journal of Hydrology* 508 :227-239. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2013.11.007](https://doi.org/10.1016/j.jhydrol.2013.11.007). Cité p. 199.
- BOOTH, D. B., et C. R. JACKSON. 1997. “Urbanization of aquatic systems : Degradation thresholds, stormwater detection, and the limits of mitigation”. *JAWRA Journal of the American Water Resources Association* 33 (5) : 1077-1090. ISSN : 1752-1688. doi :[10.1111/j.1752-1688.1997.tb04126.x](https://doi.org/10.1111/j.1752-1688.1997.tb04126.x). Cité pp. 30, 109.
- BOOTH, D. B., *et al.* 2004. “Reviving urban streams : Land use, hydrology, biology, and human behavior”. *JAWRA Journal of the American Water Resources Association* 40 (5) : 1351-1364. ISSN : 1752-1688. doi :[10.1111/j.1752-1688.2004.tb01591.x](https://doi.org/10.1111/j.1752-1688.2004.tb01591.x). Cité p. 31.
- BOUGHTON, W., et F. CHIEW. 2007. “Estimating runoff in ungauged catchments from rainfall, PET and the AWBM model”. *Environmental Modelling & Software*, Special section : Environmental Risk and Emergency Management, 22 (4) : 476-487. ISSN : 1364-8152. doi :[10.1016/j.envsoft.2006.01.009](https://doi.org/10.1016/j.envsoft.2006.01.009). Cité p. 198.
- BOULESTEIX, A.-L., *et al.* 2012. “Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics”. *Wiley Interdisciplinary Reviews : Data Mining and Knowledge Discovery* 2 (6) : 493-507. ISSN : 1942-4795. doi :[10.1002/widm.1072](https://doi.org/10.1002/widm.1072). Cité p. 200.
- BOURGIN, F., *et al.* 2015. “Transferring global uncertainty estimates from gauged to ungauged catchments”. *Hydrology and Earth System Sciences* 19 (5) : 2535-2546. ISSN : 1027-5606. doi :[10.5194/hess-19-2535-2015](https://doi.org/10.5194/hess-19-2535-2015). Cité pp. 198, 200.
- BOURGIN, P., *et al.* 2010. *Description des caractéristiques morphologiques, climatiques et hydrologiques de 4436 bassins versants français. Guide d'utilisation de la base de données hydro-climatique*. Rapport Technique. Antony, France : IRSTEA. Cité pp. 68, 78, 80, 113, 204.
- BOYD, M. J., M. C. BUFILL et R. M. KNEE. 1993. “Pervious and impervious runoff in urban catchments”. *Hydrological Sciences Journal* 38 (6) : 463-478. ISSN : 0262-6667. doi :[10.1080/02626669309492699](https://doi.org/10.1080/02626669309492699). Cité pp. 30, 252.

- BRABEC, E., S. SCHULTE et P. L. RICHARDS. 2002. "Impervious surfaces and water quality : A review of current literature and its implications for watershed planning". *Journal of Planning Literature* 16 (4) : 499-514. ISSN : 0885-4122. doi : [10.1177/088541202400903563](https://doi.org/10.1177/088541202400903563). Cité pp. 22, 109, 160, 170, 260.
- BRANDES, D., G. J. CAVALLO et M. L. NILSON. 2005. "Base flow trends in urbanizing watersheds of the Delaware river basin". *JAWRA Journal of the American Water Resources Association* 41 (6) : 1377-1391. ISSN : 1752-1688. doi : [10.1111/j.1752-1688.2005.tb03806.x](https://doi.org/10.1111/j.1752-1688.2005.tb03806.x). Cité pp. 31, 32, 33, 36.
- BRANGER, E., *et al.* 2013. "Assessment of the influence of land use data on the water balance components of a peri-urban catchment using a distributed modelling approach". *Journal of Hydrology* 505 :312-325. ISSN : 0022-1694. doi : [10.1016/j.jhydro.2013.09.055](https://doi.org/10.1016/j.jhydro.2013.09.055). Cité pp. 39, 53.
- BRAUD, I., T. D. FLETCHER et H. ANDRIEU. 2013. "Hydrology of peri-urban catchments : Processes and modelling". *Journal of Hydrology*, Hydrology of peri-urban catchments : processes and modelling, 485 :1-4. ISSN : 0022-1694. doi : [10.1016/j.jhydro.2013.02.045](https://doi.org/10.1016/j.jhydro.2013.02.045). Cité pp. 43, 59, 61.
- BRAUD, I., *et al.* 2013. "Evidence of the impact of urbanization on the hydrological regime of a medium-sized periurban catchment in France". *Journal of Hydrology*, Hydrology of peri-urban catchments : processes and modelling, 485 :5-23. ISSN : 0022-1694. doi : [10.1016/j.jhydro.2012.04.049](https://doi.org/10.1016/j.jhydro.2012.04.049). Cité pp. 20, 33, 36, 186, 227, 251.
- BRAUD, I., *et al.* 2010. "The AVuPUR project (Assessing the Vulnerability of Peri-Urbans Rivers) : experimental set up, modelling strategy and first results". In *7th International Conference on sustainable techniques and strategies for urban water management*. Lyon, France. Cité p. 61.
- BREIMAN, L. 2001a. "Random Forests". *Machine Learning* 45 (1) : 5-32. ISSN : 1573-0565. doi : [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324). Cité pp. 150, 199, 209, 210.
- . 2001b. "Statistical modeling : The Two Cultures (with comments and a rejoinder by the author)". *Statistical Science* 16 (3) : 199-231. ISSN : 0883-4237, 2168-8745. doi : [10.1214/ss/1009213726](https://doi.org/10.1214/ss/1009213726). Cité p. 221.
- BREIMAN, L., et A. CUTLER. 2019. *Random Forests*. Visité le 06/05/2019. <https://www.stat.berkeley.edu/~breiman/RandomForests>. Cité p. 209.
- BRIGODE, P., L. OUDIN et C. PERRIN. 2013. "Hydrological model parameter instability : A source of additional uncertainty in estimating the hydrological impacts of climate change?" *Journal of Hydrology* 476 :410-425. ISSN : 0022-1694. doi : [10.1016/j.jhydro.2012.11.012](https://doi.org/10.1016/j.jhydro.2012.11.012). Cité p. 188.
- BRONSTERT, A., D. NIEHOFF et G. BÜRGER. 2002. "Effects of climate and land-use change on storm runoff generation : present knowledge and modelling capabilities". *Hydrological Processes* 16 (2) : 509-529. ISSN : 1099-1085. doi : [10.1002/hyp.326](https://doi.org/10.1002/hyp.326). Cité pp. 15, 16, 56, 228, 250.
- BRUN, S. E., et L. E. BAND. 2000. "Simulating runoff behavior in an urbanizing watershed". *Computers, Environment and Urban Systems* 24 (1) : 5-22. ISSN : 0198-9715. doi : [10.1016/S0198-9715\(99\)00040-X](https://doi.org/10.1016/S0198-9715(99)00040-X). Cité pp. 34, 35, 36, 53, 136, 158, 170, 186, 259.
- BRUNNER, M. I., J. SEIBERT et A.-C. FAVRE. 2018. "Representative sets of design hydrographs for ungauged catchments : A regional approach using probabilistic region memberships". *Advances in Water Resources* 112 :235-244. ISSN : 0309-1708. doi : [10.1016/j.advwatres.2017.12.018](https://doi.org/10.1016/j.advwatres.2017.12.018). Cité p. 199.
- BRUNNER, P., P. G. COOK et C. T. SIMMONS. 2009. "Hydrogeologic controls on disconnection between surface water and groundwater". *Water Resources Research* 45 (1). ISSN : 1944-7973. doi : [10.1029/2008WR006953](https://doi.org/10.1029/2008WR006953). Cité p. 15.
- BUCHANAN, B., *et al.* 2018. "Estimating dominant runoff modes across the conterminous United States". *Hydrological Processes* 32 (26) : 3881-3890. ISSN : 1099-1085. doi : [10.1002/hyp.13296](https://doi.org/10.1002/hyp.13296). Cité p. 199.
- BUDYKO, M. I. 1974. *Climate and Life*. Academic Press, Inc. T. 18. International geophysics series. New York : Academic Press. ISBN : 0-12-139450-6. doi : [10.1016/S0074-6142\(09\)60005-X](https://doi.org/10.1016/S0074-6142(09)60005-X). Cité p. 203.

- BUNN, S. E., et A. H. ARTHINGTON. 2002. "Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity". *Environmental Management* 30 (4) : 492-507. ISSN : 1432-1009. doi : [10.1007/s00267-002-2737-0](https://doi.org/10.1007/s00267-002-2737-0). Cité pp. 32, 33.
- BURNS, D., *et al.* 2005. "Effects of suburban development on runoff generation in the Croton River basin, New York, USA". *Journal of Hydrology* 311 (1) : 266-281. ISSN : 0022-1694. doi : [10.1016/j.jhydro.2005.01.022](https://doi.org/10.1016/j.jhydro.2005.01.022). Cité pp. 28, 35, 36, 186.
- BÜTTNER, G., *et al.* 2012. "Implementation and achievements of CLC2006". Cité p. 203.
- BUYANTUYEV, A., J. WU et C. GRIES. 2010. "Multiscale analysis of the urbanization pattern of the Phoenix metropolitan landscape of USA : Time, space and thematic resolution". *Landscape and Urban Planning* 94 (3) : 206-217. ISSN : 0169-2046. doi : [10.1016/j.landurbplan.2009.10.005](https://doi.org/10.1016/j.landurbplan.2009.10.005). Cité p. 32.
- CANTONE, J., et A. SCHMIDT. 2011. "Improved understanding and prediction of the hydrologic response of highly urbanized catchments through development of the Illinois Urban Hydrologic Model". *Water Resources Research* 47 (8). ISSN : 0043-1397. doi : [10.1029/2010WR009330](https://doi.org/10.1029/2010WR009330). Cité p. 49.
- CAPPUS, P. 1960. "Bassin expérimental d'Alrance. Étude des lois de l'écoulement. Application au calcul et la prévision des débits". *La Houille Blanche* : 493-520. ISSN : 0018-6368, 1958-5551. doi : [10.1051/lhb/1960007](https://doi.org/10.1051/lhb/1960007). Cité p. 145.
- CARBAJAL, J. P., et V. BELLOS. 2018. "An overview of the role of Machine Learning in hydraulic and hydrological modeling". *engrXiv*. doi : [10.31224/osf.io/wgm72](https://doi.org/10.31224/osf.io/wgm72). Cité p. 199.
- CARLISLE, D. M., *et al.* 2010. "Predicting the natural flow regime : models for assessing hydrological alteration in streams". *River Research and Applications* 26 (2) : 118-136. ISSN : 1535-1467. doi : [10.1002/rra.1247](https://doi.org/10.1002/rra.1247). Cité p. 199.
- CARRÉ, C., et J.-C. DEUTSCH. 2015. *L'eau dans la ville. Une amie qui nous fait la guerre*. Éditions de l'aube. ISBN : 978-2-8159-1090-3. Cité pp. 3, 5, 16, 42, 43.
- CASTIGLIONI, S., *et al.* 2010. "Calibration of rainfall-runoff models in ungauged basins : A regional maximum likelihood approach". *Advances in Water Resources*, Special Issue on Novel Insights in Hydrological Modelling, 33 (10) : 1235-1242. ISSN : 0309-1708. doi : [10.1016/j.advwatres.2010.04.009](https://doi.org/10.1016/j.advwatres.2010.04.009). Cité p. 198.
- CHANGNON, S. A., R. T. SHEALY et R. W. SCOTT. 1991. "Precipitation changes in fall, winter, and spring caused by St. Louis". *Journal of Applied Meteorology* 30 (1) : 126-134. ISSN : 0894-8763. doi : [10.1175/1520-0450\(1991\)030<0126:PCIFWA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1991)030<0126:PCIFWA>2.0.CO;2). Cité p. 17.
- CHEN, J., A. A. HILL et L. D. URBANO. 2009. "A GIS-based model for urban flood inundation". *Journal of Hydrology* 373 (1) : 184-192. ISSN : 0022-1694. doi : [10.1016/j.jhydro.2009.04.021](https://doi.org/10.1016/j.jhydro.2009.04.021). Cité p. 44.
- CHEN, R.-s., W.-N. CHUANG et S.-j. CHENG. 2014. "Effects of urbanization variables on model parameters for watershed divisions". *Hydrological Sciences Journal* 59 (6) : 1167-1183. ISSN : 0262-6667. doi : [10.1080/02626667.2014.910305](https://doi.org/10.1080/02626667.2014.910305). Cité pp. 31, 167, 200.
- CHENG, S.-j., C.-f. LEE et J.-h. LEE. 2010. "Effects of urbanization factors on model parameters". *Water Resources Management* 24 (4) : 775-794. ISSN : 1573-1650. doi : [10.1007/s11269-009-9471-9](https://doi.org/10.1007/s11269-009-9471-9). Cité p. 200.
- CHOCAT, B. 1997. "Aménagement urbain et hydrologie". *La Houille Blanche*, n° 7 : 12-19. ISSN : 0018-6368, 1958-5551. doi : [10.1051/lhb/1997057](https://doi.org/10.1051/lhb/1997057). Cité p. 16.
- CHOCAT, B., *et al.* 2001. "Urban drainage redefined : from stormwater removal to integrated management". *Water Science and Technology* 43 (5) : 61-68. ISSN : 0273-1223. doi : [10.2166/wst.2001.0251](https://doi.org/10.2166/wst.2001.0251). Cité p. 16.
- CHOI, J.-Y., *et al.* 2003. "GIS based long term hydrologic impact evaluation for watershed urbanization". *JAWRA Journal of the American Water Resources Association* 39 (3) : 623-635. ISSN : 1752-1688. doi : [10.1111/j.1752-1688.2003.tb03680.x](https://doi.org/10.1111/j.1752-1688.2003.tb03680.x). Cité pp. 35, 36.

- CHORMANSKI, J., *et al.* 2008. "Improving distributed runoff prediction in urbanized catchments with remote sensing based estimates of impervious surface cover". *Sensors* 8 (2) : 910-932. doi :[10.3390/s8020910](https://doi.org/10.3390/s8020910). Cité p. 47.
- CHUI, T. F. M., et D. H. TRINH. 2016. "Modelling infiltration enhancement in a tropical urban catchment for improved stormwater management". *Hydrological Processes* 30 (23) : 4405-4419. ISSN : 1099-1085. doi :[10.1002/hyp.10926](https://doi.org/10.1002/hyp.10926). Cité p. 21.
- CHUNG, E.-S., K. PARK et K. S. LEE. 2011. "The relative impacts of climate change and urbanization on the hydrological response of a Korean urban watershed". *Hydrological Processes* 25 (4) : 544-560. ISSN : 1099-1085. doi :[10.1002/hyp.7781](https://doi.org/10.1002/hyp.7781). Cité p. 264.
- CICCO, L. A. D., *et al.* 2018. *dataRetrieval : R packages for discovering and retrieving water data available from U.S. federal hydrologic web services*. Reston, VA : U.S. Geological Survey. doi :[10.5066/P9X4L3GE](https://doi.org/10.5066/P9X4L3GE). Cité pp. 68, 138, 170, 201, 230.
- CLAESSENS, L., *et al.* 2006. "Effect of historical changes in land use and climate on the water budget of an urbanizing watershed". *Water Resources Research* 42 (3). ISSN : 1944-7973. doi :[10.1029/2005WR004131](https://doi.org/10.1029/2005WR004131). Cité p. 37.
- CLARK, M. P., D. KAVETSKI et F. FENICIA. 2011. "Pursuing the method of multiple working hypotheses for hydrological modeling". *Water Resources Research* 47 (9). ISSN : 1944-7973. doi :[10.1029/2010WR009827](https://doi.org/10.1029/2010WR009827). Cité p. 248.
- CLOPPER, C. J., et E. S. PEARSON. 1934. "The use of confidence or fiducial limits illustrated in the case of the binomial". *Biometrika* 26 (4) : 404-413. ISSN : 0006-3444. doi :[10.1093/biomet/26.4.404](https://doi.org/10.1093/biomet/26.4.404). Cité p. 240.
- COLLISCHONN, W., et F. M. FAN. 2013. "Defining parameters for Eckhardt's digital baseflow filter". *Hydrological Processes* 27 (18) : 2614-2622. ISSN : 1099-1085. doi :[10.1002/hyp.9391](https://doi.org/10.1002/hyp.9391). Cité pp. 103, 145, 172, 232, 298.
- CONGEDO, L., *et al.* 2016. "Copernicus high-resolution layers for land cover classification in Italy". *Journal of Maps* 12 (5) : 1195-1205. doi :[10.1080/17445647.2016.1145151](https://doi.org/10.1080/17445647.2016.1145151). Cité pp. 171, 230.
- CORON, L., *et al.* 2017. "The suite of lumped GR hydrological models in an R package". *Environmental Modelling and Software* 94 :166-171. doi :[10.1016/j.envsoft.2017.05.002](https://doi.org/10.1016/j.envsoft.2017.05.002). Cité pp. 209, 238.
- CORTINOVIS, C., et D. GENELETTI. 2018. "Ecosystem services in urban plans : What is there, and what is still needed for better decisions". *Land Use Policy* 70 :298-312. ISSN : 0264-8377. doi :[10.1016/j.landusepol.2017.10.017](https://doi.org/10.1016/j.landusepol.2017.10.017). Cité p. 4.
- CRISS, R. E., et W. E. WINSTON. 2008. "Do Nash values have value? Discussion and alternate proposals". *Hydrological Processes* 22 (14) : 2723-2725. ISSN : 1099-1085. doi :[10.1002/hyp.7072](https://doi.org/10.1002/hyp.7072). Cité p. 240.
- CRISTIANO, E., M.-C. t. VELDHIJS et N. v. d. GIESEN. 2017. "Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas – a review". *Hydrology and Earth System Sciences* 21 (7) : 3859-3878. ISSN : 1027-5606. doi :[10.5194/hess-21-3859-2017](https://doi.org/10.5194/hess-21-3859-2017). Cité pp. 55, 228, 259.
- CRUTZEN, P. J. 2002. "The "anthropocene"". *Journal de Physique IV (Proceedings)* 12 (10) : 1-5. ISSN : 1155-4339. doi :[10.1051/jp4:20020447](https://doi.org/10.1051/jp4:20020447). Cité p. 3.
- CUO, L., *et al.* 2008. "Hydrologic prediction for urban watersheds with the Distributed Hydrology-Soil-Vegetation Model". *Hydrological Processes* 22 (21) : 4205-4213. ISSN : 1099-1085. doi :[10.1002/hyp.7023](https://doi.org/10.1002/hyp.7023). Cité pp. 46, 228.
- CUTLER, A., *et al.* 2018. *randomForest : Breiman and Cutler's random forests for classification and regression*. Visité le 21/05/2019. <https://CRAN.R-project.org/package=randomForest>. Cité p. 199.
- DE NIEL, J., *et al.* 2020. "Efficient approach for impact analysis of land cover changes on hydrological extremes by means of a lumped conceptual model". *Journal of Hydrology : Regional Studies* 28 :100666. ISSN : 2214-5818. doi :[10.1016/j.ejrh.2020.100666](https://doi.org/10.1016/j.ejrh.2020.100666). Cité p. 228.

- DEBBAGE, N., et J. M. SHEPHERD. 2018. "The influence of urban development patterns on streamflow characteristics in the Charlanta Megaregion". *Water Resources Research* 54 (5) : 3728-3747. ISSN : 1944-7973. doi :[10.1029/2017WR021594](https://doi.org/10.1029/2017WR021594). Cité p. 161.
- DEWALLE, D. R., *et al.* 2000. "Potential effects of climate change and urbanization on mean annual streamflow in the United States". *Water Resources Research* 36 (9) : 2655-2664. ISSN : 1944-7973. doi :[10.1029/2000WR900134](https://doi.org/10.1029/2000WR900134). Cité pp. 31, 33, 37, 168, 186.
- DÍAZ-URIARTE, R., et S. ALVAREZ DE ANDRÉS. 2006. "Gene selection and classification of microarray data using random forest". *BMC Bioinformatics* 7 (1) : 3. ISSN : 1471-2105. doi :[10.1186/1471-2105-7-3](https://doi.org/10.1186/1471-2105-7-3). Cité p. 209.
- DIEM, J. E., T. C. HILL et R. A. MILLIGAN. 2018. "Diverse multi-decadal changes in streamflow within a rapidly urbanizing region". *Journal of Hydrology* 556 :61-71. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2017.10.026](https://doi.org/10.1016/j.jhydrol.2017.10.026). Cité pp. 32, 34, 35, 36, 37, 186, 214, 227, 237.
- DIEZ-SIERRA, J., et M. del JESUS. 2019. "Subdaily rainfall estimation through daily rainfall downscaling using random forests in Spain". *Water* 11 (1) : 125. doi :[10.3390/w11010125](https://doi.org/10.3390/w11010125). Cité p. 199.
- DIXON, B., et J. EARLS. 2012. "Effects of urbanization on streamflow using SWAT with real and simulated meteorological data". *Applied Geography* 35 (1) : 174-190. ISSN : 0143-6228. doi :[10.1016/j.apgeog.2012.06.010](https://doi.org/10.1016/j.apgeog.2012.06.010). Cité p. 52.
- DIXON, P. G., et T. L. MOTE. 2003. "Patterns and causes of Atlanta's urban heat island–Initiated precipitation". *Journal of Applied Meteorology* 42 (9) : 1273-1284. ISSN : 0894-8763. doi :[10.1175/1520-0450\(2003\)042<1273:PACOU>2.0.CO;2](https://doi.org/10.1175/1520-0450(2003)042<1273:PACOU>2.0.CO;2). Cité p. 17.
- DOTTO, C. B. S., *et al.* 2011. "Performance and sensitivity analysis of stormwater models using a Bayesian approach and long-term high resolution data". *Environmental Modelling & Software* 26 (10) : 1225-1239. ISSN : 1364-8152. doi :[10.1016/j.envsoft.2011.03.013](https://doi.org/10.1016/j.envsoft.2011.03.013). Cité pp. 50, 51, 173, 250.
- DOU, J., *et al.* 2015. "Observed spatial characteristics of Beijing urban climate impacts on summer thunderstorms". *Journal of Applied Meteorology and Climatology* 54 (1) : 94-105. ISSN : 1558-8424. doi :[10.1175/JAMC-D-13-0355.1](https://doi.org/10.1175/JAMC-D-13-0355.1). Cité p. 17.
- DOW, C. L. 2007. "Assessing regional land-use/cover influences on New Jersey Pinelands streamflow through hydrograph analysis". *Hydrological Processes* 21 (2) : 185-197. ISSN : 1099-1085. doi :[10.1002/hyp.6232](https://doi.org/10.1002/hyp.6232). Cité pp. 34, 36.
- DROGUE, G., *et al.* 2002. "The applicability of a parsimonious model for local and regional prediction of runoff". *Hydrological Sciences Journal* 47 (6) : 905-920. ISSN : 0262-6667. doi :[10.1080/02626660209492999](https://doi.org/10.1080/02626660209492999). Cité p. 198.
- DUAN, Q., *et al.* 2006. "Model Parameter Estimation Experiment (MOPEX) : An overview of science strategy and major results from the second and third workshops". *Journal of Hydrology*, The model parameter estimation experiment, 320 (1) : 3-17. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2005.07.031](https://doi.org/10.1016/j.jhydrol.2005.07.031). Cité p. 198.
- DUCHARNE, A. 2009. "Reducing scale dependence in TOPMODEL using a dimensionless topographic index". *Hydrology and Earth System Sciences* 13 (12) : 2399-2412. ISSN : 1027-5606. doi :[10.5194/hess-13-2399-2009](https://doi.org/10.5194/hess-13-2399-2009). Cité p. 112.
- DUNNE, T., et R. D. BLACK. 1970. "An experimental investigation of runoff production in permeable soils". *Water Resources Research* 6 (2) : 478-490. ISSN : 1944-7973. doi :[10.1029/WR006i002p00478](https://doi.org/10.1029/WR006i002p00478). Cité p. 15.
- EASTON, Z. M., *et al.* 2007. "Hydrologic assessment of an urban variable source watershed in the northeast United States". *Water Resources Research* 43 (3). ISSN : 1944-7973. doi :[10.1029/2006WR005076](https://doi.org/10.1029/2006WR005076). Cité p. 51.
- EBEL, B. A., et K. LOAGUE. 2006. "Physics-based hydrologic-response simulation : Seeing through the fog of equifinality". *Hydrological Processes* 20 (13) : 2887-2900. ISSN : 1099-1085. doi :[10.1002/hyp.6388](https://doi.org/10.1002/hyp.6388). Cité pp. 56, 187.

- EBRAHIMIAN, A., B. N. WILSON et J. S. GULLIVER. 2016. "Improved methods to estimate the effective impervious area in urban catchments using rainfall-runoff data". *Journal of Hydrology* 536 :109-118. ISSN : 0022-1694. doi :10.1016/j.jhydro.2016.02.023. Cité pp. 32, 161, 220, 252.
- ECKART, K., Z. MCPHEE et T. BOLISSETTI. 2017. "Performance and implementation of low impact development – A review". *Science of The Total Environment* 607-608 :413-432. ISSN : 0048-9697. doi :10.1016/j.scitotenv.2017.06.254. Cité p. 21.
- ECKHARDT, K. 2005. "How to construct recursive digital filters for baseflow separation". *Hydrological Processes* 19 (2) : 507-515. ISSN : 1099-1085. doi :10.1002/hyp.5675. Cité pp. 83, 103, 145, 172, 232, 298.
- EDIJATNO *et al.* 1999. "GR3J : a daily watershed model with three free parameters". *Hydrological Sciences Journal* 44 (2) : 263-277. ISSN : 0262-6667. doi :10.1080/02626669909492221. Cité pp. 175, 209, 238, 329.
- ELLIOTT, A. H., et S. A. TROWSDALE. 2007. "A review of models for low impact urban stormwater drainage". *Environmental Modelling & Software*, Special section : Advanced Technology for Environmental Modelling, 22 (3) : 394-405. ISSN : 1364-8152. doi :10.1016/j.envsoft.2005.12.005. Cité p. 42.
- ELLIS, P. A., R. MACKAY et M. O. RIVETT. 2007. "Quantifying urban river-aquifer fluid exchange processes : A multi-scale problem". *Journal of Contaminant Hydrology*, Issues in urban hydrology : The emerging field of urban contaminant hydrology, 91 (1) : 58-80. ISSN : 0169-7722. doi :10.1016/j.jconhyd.2006.08.014. Cité p. 20.
- ESSE, W. R. van, *et al.* 2013. "The influence of conceptual model structure on model performance : a comparative study for 237 French catchments". *Hydrology and Earth System Sciences* 17 (10) : 4227-4239. ISSN : 1027-5606. doi :10.5194/hess-17-4227-2013. Cité pp. 168, 207, 233, 248.
- ETCHEVERS, P., *et al.* 2002. "Impact of a climate change on the Rhone river catchment hydrology". *Journal of Geophysical Research : Atmospheres* 107 (D16) : ACL 6-1-ACL 6-18. ISSN : 2156-2202. doi :10.1029/2001JD00490. Cité p. 16.
- EUSER, T., *et al.* 2015. "The effect of forcing and landscape distribution on performance and consistency of model structures". *Hydrological Processes* 29 (17) : 3727-3743. ISSN : 1099-1085. doi :10.1002/hyp.10445. Cité pp. 228, 263.
- FALCONE, J. A. 2011. *GAGES-II : Geospatial Attributes of Gages for Evaluating Streamflow*. USGS Unnumbered Series. Reston, VA : U.S. Geological Survey. Visité le 23/07/2019. <http://pubs.er.usgs.gov/publication/70046617>. Cité pp. 68, 77, 78, 79, 113, 138, 142, 160, 170, 201, 204, 230.
- FAO/IIASA/ISRIC/ISS-CAS/JRC. 2012. *Harmonized world soil database*. Rapp. tech. FAO, Rome, Italy et IIASA, Laxenburg, Austria. Cité pp. 118, 143, 206.
- FAULKNER, D. S., et S. BARBER. 2009. "Performance of the revitalised flood hydrograph method". *Journal of Flood Risk Management* 2 (4) : 254-261. ISSN : 1753-318X. doi :10.1111/j.1753-318X.2009.01042.x. Cité p. 53.
- FENICIA, E, J. J. McDONNELL et H. H. G. SAVENIJE. 2008. "Learning from model improvement : On the contribution of complementary data to process understanding". *Water Resources Research* 44 (6). ISSN : 1944-7973. doi :10.1029/2007WR006386. Cité p. 248.
- FERNÁNDEZ-BARRERA, A. H., *et al.* 2008. "Infiltration capacity assessment of urban pavements using the LCS permeameter and the CP infiltrometer". *Journal of Irrigation and Drainage Engineering* 134 (5) : 659-665. doi :10.1061/(ASCE)0733-9437(2008)134:5(659). Cité pp. 19, 30.
- FERNANDEZ, W., R. M. VOGEL et A. SANKARASUBRAMANIAN. 2000. "Regional calibration of a watershed model". *Hydrological Sciences Journal* 45 (5) : 689-707. ISSN : 0262-6667. doi :10.1080/02626660009492371. Cité p. 198.

- FICCHÌ, A. 2017. “An adaptive hydrological model for multiple time-steps : diagnostics and improvements based on fluxes consistency”. PHD Thesis, Université Pierre et Marie Curie-Paris VI. Cité pp. 41, 66, 74, 88, 102, 207, 208, 262, 308.
- FICCHÌ, A., C. PERRIN et V. ANDRÉASSIAN. 2019. “Hydrological modelling at multiple sub-daily time steps : Model improvement via flux-matching”. *Journal of Hydrology* 575 :1308-1327. ISSN : 0022-1694. doi:[10.1016/j.jhydrol.2019.05.084](https://doi.org/10.1016/j.jhydrol.2019.05.084). Cité pp. 9, 10, 146, 168, 173, 174, 229, 233, 234, 251, 304, 305.
- . 2016. “Impact of temporal resolution of inputs on hydrological model performance : An analysis based on 2400 flood events”. *Journal of Hydrology* 538 :454-470. ISSN : 0022-1694. doi:[10.1016/j.jhydrol.2016.04.016](https://doi.org/10.1016/j.jhydrol.2016.04.016). Cité p. 232.
- FIDAL, J., et T. KJELDSSEN. 2020. “Operational comparison of rainfall-runoff models through hypothesis testing”. *Journal of Hydrologic Engineering* 25 (4) : 04020005. ISSN : 1943-5584. doi:[10.1061/\(ASCE\)HE.1943-5584.0001892](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001892). Cité pp. 229, 240, 251.
- FITZHUGH, T. W., et R. M. VOGEL. 2011. “The impact of dams on flood flows in the United States”. *River Research and Applications* 27 (10) : 1192-1215. ISSN : 1535-1467. doi:[10.1002/rra.1417](https://doi.org/10.1002/rra.1417). Cité pp. 78, 169.
- FLETCHER, T. D., H. ANDRIEU et P. HAMEL. 2013. “Understanding, management and modelling of urban hydrology and its consequences for receiving waters : A state of the art”. *Advances in Water Resources*, 35th Year Anniversary Issue, 51 :261-279. ISSN : 0309-1708. doi:[10.1016/j.advwatres.2012.09.001](https://doi.org/10.1016/j.advwatres.2012.09.001). Cité pp. 5, 13, 16, 42, 43, 168, 206, 220, 227, 232, 259, 261.
- FLETCHER, T. D., *et al.* 2015. “SUDS, LID, BMPs, WSUD and more – The evolution and application of terminology surrounding urban drainage”. *Urban Water Journal* 12 (7) : 525-542. ISSN : 1573-062X. doi:[10.1080/1573062X.2014.916314](https://doi.org/10.1080/1573062X.2014.916314). Cité pp. 21, 43.
- FOLEY, J. A., *et al.* 2005. “Global consequences of land use”. *Science* 309 (5734) : 570-574. ISSN : 0036-8075, 1095-9203. doi:[10.1126/science.1111772](https://doi.org/10.1126/science.1111772). Cité p. 15.
- FOSTER, S. 2020. “Global policy overview of groundwater in urban development—A tale of 10 cities!” *Water* 12 (2) : 456. doi:[10.3390/w12020456](https://doi.org/10.3390/w12020456). Cité p. 20.
- FOSTER, S., et P. J. CHILTON. 2003. “Groundwater : the processes and global significance of aquifer degradation”. *Philosophical Transactions of the Royal Society of London. Series B : Biological Sciences* 358 (1440) : 1957-1972. doi:[10.1098/rstb.2003.1380](https://doi.org/10.1098/rstb.2003.1380). Cité p. 20.
- FRANCZYK, J., et H. CHANG. 2009. “The effects of climate change and urbanization on the runoff of the Rock Creek basin in the Portland metropolitan area, Oregon, USA”. *Hydrological Processes* 23 (6) : 805-815. ISSN : 1099-1085. doi:[10.1002/hyp.7176](https://doi.org/10.1002/hyp.7176). Cité p. 52.
- FUAMBA, M., *et al.* 2019. “Value of distributed water level and soil moisture data in the evaluation of a distributed hydrological model : Application to the PUMMA model in the Mercier catchment (6.6 km²) in France”. *Journal of Hydrology* 569 :753-770. ISSN : 0022-1694. doi:[10.1016/j.jhydrol.2018.12.035](https://doi.org/10.1016/j.jhydrol.2018.12.035). Cité p. 49.
- FURUSHO, C., K. CHANCIBAULT et H. ANDRIEU. 2013. “Adapting the coupled hydrological model ISBA-TOPMODEL to the long-term hydrological cycles of suburban rivers : Evaluation and sensitivity analysis”. *Journal of Hydrology*, Hydrology of peri-urban catchments : processes and modelling, 485 :139-147. ISSN : 0022-1694. doi:[10.1016/j.jhydrol.2012.06.059](https://doi.org/10.1016/j.jhydrol.2012.06.059). Cité pp. 47, 59, 250, 251, 262.
- GAO, H., *et al.* 2018. “Landscape heterogeneity and hydrological processes : a review of landscape-based hydrological models”. *Landscape Ecology* 33 (9) : 1461-1480. ISSN : 1572-9761. doi:[10.1007/s10980-018-0690-4](https://doi.org/10.1007/s10980-018-0690-4). Cité pp. 15, 161.
- GAUME, E., *et al.* 2009. “A compilation of data on European flash floods”. *Journal of Hydrology* 367 (1) : 70-78. ISSN : 0022-1694. doi:[10.1016/j.jhydrol.2008.12.028](https://doi.org/10.1016/j.jhydrol.2008.12.028). Cité p. 104.

- GENUER, R., *et al.* 2017. "Random forests for big data". *Big Data Research* 9 :28-46. ISSN : 2214-5796. doi :[10.1016/j.bdr.2017.07.003](https://doi.org/10.1016/j.bdr.2017.07.003). Cité p. 209.
- GHARARI, S., *et al.* 2014. "Using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration". *Hydrology and Earth System Sciences* 18 (12) : 4839-4859. ISSN : 1027-5606. doi :[10.5194/hess-18-4839-2014](https://doi.org/10.5194/hess-18-4839-2014). Cité pp. 228, 250, 263.
- GILBERT, J. K., et J. C. CLAUSEN. 2006. "Stormwater runoff quality and quantity from asphalt, paver, and crushed stone driveways in Connecticut". *Water Research* 40 (4) : 826-832. ISSN : 0043-1354. doi :[10.1016/j.watres.2005.12.006](https://doi.org/10.1016/j.watres.2005.12.006). Cité pp. 19, 30.
- GIRES, A., *et al.* 2017. "Fractal analysis of urban catchments and their representation in semi-distributed models : imperviousness and sewer system". *Hydrology and Earth System Sciences* 21 (5) : 2361-2375. ISSN : 1027-5606. doi :[10.5194/hess-21-2361-2017](https://doi.org/10.5194/hess-21-2361-2017). Cité pp. 32, 55.
- GIRONÁS, J., *et al.* 2009. "A morpho-climatic instantaneous unit hydrograph model for urban catchments based on the kinematic wave approximation". *Journal of Hydrology* 377 (3) : 317-334. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2009.08.030](https://doi.org/10.1016/j.jhydrol.2009.08.030). Cité p. 45.
- . 2010. "Evaluation of methods for representing urban terrain in storm-water modeling". *Journal of Hydrologic Engineering* 15 (1) : 1-14. ISSN : 1084-0699. doi :[10.1061/\(ASCE\)HE.1943-5584.0000142](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000142). Cité p. 112.
- GISLASON, P. O., J. A. BENEDIKTSSON et J. R. SVEINSSON. 2006. "Random Forests for land cover classification". *Pattern Recognition Letters, Pattern Recognition in Remote Sensing (PRRS 2004)*, 27 (4) : 294-300. ISSN : 0167-8655. doi :[10.1016/j.patrec.2005.08.011](https://doi.org/10.1016/j.patrec.2005.08.011). Cité p. 199.
- GLEESON, T., *et al.* 2014. "A glimpse beneath earth's surface : GLObal HYdrogeology MaPS (GLHYMPS) of permeability and porosity". *Geophysical Research Letters* 41 (11) : 3891-3898. ISSN : 0094-8276. doi :[10.1002/2014GL059856](https://doi.org/10.1002/2014GL059856). Cité pp. 118, 144, 205, 206.
- GÖBEL, P., *et al.* 2004. "Near-natural stormwater management and its effects on the water budget and groundwater surface in urban areas taking account of the hydrogeological conditions". *Journal of Hydrology, Urban Hydrology*, 299 (3) : 267-283. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2004.08.013](https://doi.org/10.1016/j.jhydrol.2004.08.013). Cité pp. 20, 45, 46, 186.
- GRAEFF, T., *et al.* 2012. "Predicting event response in a nested catchment with generalized linear models and a distributed watershed model". *Hydrological Processes* 26 (24) : 3749-3769. ISSN : 1099-1085. doi :[10.1002/hyp.8463](https://doi.org/10.1002/hyp.8463). Cité p. 136.
- GRAYSON, R. B., *et al.* 1997. "Preferred states in spatial soil moisture patterns : Local and nonlocal controls". *Water Resources Research* 33 (12) : 2897-2908. ISSN : 1944-7973. doi :[10.1029/97WR02174](https://doi.org/10.1029/97WR02174). Cité p. 15.
- GREGORY, J. H., *et al.* 2006. "Effect of urban soil compaction on infiltration rate". *Journal of Soil and Water Conservation* 61 (3) : 117-124. ISSN : 0022-4561, 1941-3300. Cité p. 18.
- GRIMMOND, C., et T. OKE. 1991. "An evapotranspiration-interception model for urban areas". *Water Resources Research* 27 (7) : 1739-1755. ISSN : 0043-1397. doi :[10.1029/91WR00557](https://doi.org/10.1029/91WR00557). Cité pp. 17, 50.
- GROVE, M., J. HARBOR et B. ENGEL. 1998. "Composite vs. distributed Curve Numbers : Effects on estimates of storm runoff depths". *JAWRA Journal of the American Water Resources Association* 34 (5) : 1015-1023. ISSN : 1752-1688. doi :[10.1111/j.1752-1688.1998.tb04150.x](https://doi.org/10.1111/j.1752-1688.1998.tb04150.x). Cité p. 30.
- GUAN, M., N. SILLANPÄÄ et H. KOIVUSALO. 2015. "Modelling and assessment of hydrological changes in a developing urban catchment". *Hydrological Processes* 29 (13) : 2880-2894. ISSN : 1099-1085. doi :[10.1002/hyp.10410](https://doi.org/10.1002/hyp.10410). Cité p. 39.
- . 2016. "Storm runoff response to rainfall pattern, magnitude and urbanization in a developing urban catchment". *Hydrological Processes* 30 (4) : 543-557. ISSN : 1099-1085. doi :[10.1002/hyp.10624](https://doi.org/10.1002/hyp.10624). Cité p. 31.

- GUIMBERTEAU, M., *et al.* 2017. "Impacts of future deforestation and climate change on the hydrology of the Amazon Basin : a multi-model analysis with a new set of land-cover change scenarios". *Hydrology and Earth System Sciences* 21 (3) : 1455-1475. ISSN : 1027-5606. doi :[10.5194/hess-21-1455-2017](https://doi.org/10.5194/hess-21-1455-2017). Cité p. 15.
- GUPTA, H. V., *et al.* 2014. "Large-sample hydrology : a need to balance depth with breadth". *Hydrology and Earth System Sciences* 18 (2) : 463-477. ISSN : 1027-5606. doi :[10.5194/hess-18-463-2014](https://doi.org/10.5194/hess-18-463-2014). Cité pp. 57, 65, 168, 228, 229, 250, 258, 262.
- GUPTA, H. V., T. WAGENER et Y. LIU. 2008. "Reconciling theory with observations : elements of a diagnostic approach to model evaluation". *Hydrological Processes* 22 (18) : 3802-3813. ISSN : 1099-1085. doi :[10.1002/hyp.6989](https://doi.org/10.1002/hyp.6989). Cité p. 248.
- GUPTA, H. V., *et al.* 2009. "Decomposition of the mean squared error and NSE performance criteria : Implications for improving hydrological modelling". *Journal of Hydrology* 377 (1) : 80-91. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2009.08.003](https://doi.org/10.1016/j.jhydrol.2009.08.003). Cité pp. 175, 209, 238, 329.
- HAASE, D. 2009. "Effects of urbanisation on the water balance – A long-term trajectory". *Environmental Impact Assessment Review* 29 (4) : 211-219. ISSN : 0195-9255. doi :[10.1016/j.eiar.2009.01.002](https://doi.org/10.1016/j.eiar.2009.01.002). Cité pp. 38, 228.
- HAMDI, R., P. TERMONIA et P. BAGUIS. 2011. "Effects of urbanization and climate change on surface runoff of the Brussels Capital Region : a case study using an urban soil-vegetation-atmosphere-transfer model". *International Journal of Climatology* 31 (13) : 1959-1974. ISSN : 1097-0088. doi :[10.1002/joc.2207](https://doi.org/10.1002/joc.2207). Cité p. 45.
- HAMEL, P., E. DALY et T. D. FLETCHER. 2013. "Source-control stormwater management for mitigating the impacts of urbanisation on baseflow : A review". *Journal of Hydrology, Hydrology of peri-urban catchments : processes and modelling*, 485 :201-211. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2013.01.001](https://doi.org/10.1016/j.jhydrol.2013.01.001). Cité pp. 20, 21, 33.
- HAMEL, P., et T. D. FLETCHER. 2014. "Modelling the impact of stormwater source control infiltration techniques on catchment baseflow". *Hydrological Processes* 28 (24) : 5817-5831. ISSN : 1099-1085. doi :[10.1002/hyp.10069](https://doi.org/10.1002/hyp.10069). Cité pp. 50, 60, 229, 251, 262.
- HANCOCK, P. J. 2002. "Human impacts on the stream-groundwater exchange zone". *Environmental Management* 29 (6) : 763-781. ISSN : 1432-1009. doi :[10.1007/s00267-001-0064-5](https://doi.org/10.1007/s00267-001-0064-5). Cité p. 20.
- HARDEGREE, S. P., *et al.* 2008. "Evaluation of NEXRAD radar precipitation products for natural resource applications". *Rangeland Ecology & Management* 61 (3) : 346-353. ISSN : 1550-7424. doi :[10.2111/07-036.1](https://doi.org/10.2111/07-036.1). Cité p. 204.
- HARRISON, J., et J. Y. KIM. 2020. *RSelenium : R Bindings for 'Selenium WebDriver'*. Visité le 13/07/2020. <https://CRAN.R-project.org/package=RSelenium>. Cité p. 68.
- HAWLEY, R. J., et B. P. BLEDSOE. 2011. "How do flow peaks and durations change in suburbanizing semi-arid watersheds? A southern California case study". *Journal of Hydrology* 405 (1) : 69-82. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2011.05.011](https://doi.org/10.1016/j.jhydrol.2011.05.011). Cité p. 35.
- HE, X., *et al.* 2016. "Spatial downscaling of precipitation using adaptable random forests". *Water Resources Research* 52 (10) : 8217-8237. ISSN : 1944-7973. doi :[10.1002/2016WR019034](https://doi.org/10.1002/2016WR019034). Cité p. 199.
- HEGGEN, R. J. 2001. "Normalized Antecedent Precipitation Index". *Journal of Hydrologic Engineering* 6 (5) : 377-381. ISSN : 1084-0699. doi :[10.1061/\(ASCE\)1084-0699\(2001\)6:5\(377\)](https://doi.org/10.1061/(ASCE)1084-0699(2001)6:5(377)). Cité p. 145.
- HEJAZI, M. I., et G. E. MOGLEN. 2007. "Regression-based approach to low flow prediction in the Maryland Piedmont region under joint climate and land use change". *Hydrological Processes* 21 (14) : 1793-1801. ISSN : 1099-1085. doi :[10.1002/hyp.6374](https://doi.org/10.1002/hyp.6374). Cité p. 36.

- HENGL, T., *et al.* 2018. "Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables". *PeerJ* 6 :e5518. ISSN : 2167-8359. doi :[10.7717/peerj.5518](https://doi.org/10.7717/peerj.5518). Cité p. 220.
- HEWLETT, J. D., J. C. FORTSON et G. B. CUNNINGHAM. 1977. "The effect of rainfall intensity on storm flow and peak discharge from forest land". *Water Resources Research* 13 (2) : 259-266. ISSN : 1944-7973. doi :[10.1029/WR013i002p00259](https://doi.org/10.1029/WR013i002p00259). Cité pp. 136, 161.
- HILLIER, B., et A. PENN. 1992. "Dense civilisations : the shape of cities in the 21st century". *Applied Energy, Rational Use of Energy in Urban Regeneration*, 43 (1) : 41-66. ISSN : 0306-2619. doi :[10.1016/0306-2619\(92\)90073-K](https://doi.org/10.1016/0306-2619(92)90073-K). Cité p. 3.
- HOLKO, L., *et al.* 2011. "Flashiness of mountain streams in Slovakia and Austria". *Journal of Hydrology* 405 (3) : 392-401. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2011.05.038](https://doi.org/10.1016/j.jhydrol.2011.05.038). Cité pp. 141, 142, 173, 204.
- HOLLIS, G. E. 1977. "Water yield changes after the urbanization of the Canon's Brook catchment, Harlow, England". *Hydrological Sciences Bulletin* 22 (1) : 61-75. ISSN : 0303-6936. doi :[10.1080/02626667709491694](https://doi.org/10.1080/02626667709491694). Cité pp. 33, 35, 36, 37, 186.
- HOLLIS, G. E., et J. C. OVENDEN. 1988. "The quantity of stormwater runoff from ten stretches of road, a car park and eight roofs in Hertfordshire, England during 1983". *Hydrological Processes* 2 (3) : 227-243. ISSN : 1099-1085. doi :[10.1002/hyp.3360020304](https://doi.org/10.1002/hyp.3360020304). Cité pp. 4, 18.
- HOMER, C., *et al.* 2007. "Completion of the 2001 national land cover database for the conterminous United States". *Photogrammetric Engineering and Remote Sensing* 73 (4) : 337-341. Cité pp. 31, 90, 201.
- HOMER, C., *et al.* 2015. "Completion of the 2011 National Land Cover Database for the Conterminous United States – Representing a decade of land cover change information". *Photogrammetric Engineering & Remote Sensing* 81 (5) : 345-354. ISSN : 0099-1112. doi :[10.1016/S0099-1112\(15\)30100-2](https://doi.org/10.1016/S0099-1112(15)30100-2). Cité pp. 138, 171, 201.
- HOMER, C., *et al.* 2004. "Development of a 2001 National Land-Cover Database for the United States". *Photogrammetric Engineering & Remote Sensing* 70 (7) : 829-840. Cité p. 201.
- HORVAT, D. J., *et al.* 2011. "The refreshed WSR-88 level II data collection and distribution network". WSR-88D Radar Operations Center, Norman, Oklahoma. Cité p. 204.
- HRACHOWITZ, M., *et al.* 2013. "A decade of Predictions in Ungauged Basins (PUB)—a review". *Hydrological Sciences Journal* 58 (6) : 1198-1255. ISSN : 0262-6667. doi :[10.1080/02626667.2013.803183](https://doi.org/10.1080/02626667.2013.803183). Cité p. 198.
- HRACHOWITZ, M., *et al.* 2014. "Process consistency in models : The importance of system signatures, expert knowledge, and process complexity". *Water Resources Research* 50 (9) : 7445-7469. ISSN : 1944-7973. doi :[10.1002/2014WR015484](https://doi.org/10.1002/2014WR015484). Cité p. 228.
- HRACHOWITZ, M., et M. P. CLARK. 2017. "HESS Opinions : The complementary merits of competing modelling philosophies in hydrology". *Hydrology and Earth System Sciences* 21 (8) : 3953-3973. ISSN : 1027-5606. doi :[10.5194/hess-21-3953-2017](https://doi.org/10.5194/hess-21-3953-2017). Cité pp. 7, 41, 228.
- HUANG, H.-j., *et al.* 2008a. "Effect of growing watershed imperviousness on hydrograph parameters and peak discharge". *Hydrological Processes* 22 (13) : 2075-2085. ISSN : 1099-1085. doi :[10.1002/hyp.6807](https://doi.org/10.1002/hyp.6807). Cité pp. 35, 228, 248.
- HUANG, S.-y., *et al.* 2008b. "Identifying peak-imperviousness-recurrence relationships on a growing-impervious watershed, Taiwan". *Journal of Hydrology* 362 (3) : 320-336. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2008.09.002](https://doi.org/10.1016/j.jhydrol.2008.09.002). Cité pp. 19, 54.
- HUNDECHA, Y., T. B. M. J. OUARDA et A. BÁRDOSSY. 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 Resources Research* 44 (1). ISSN : 1944-7973. doi :[10.1029/2006WR005439](https://doi.org/10.1029/2006WR005439). Cité p. 198.

- HUSAIN, S. Z., S. BÉLAIR et S. LEROYER. 2014. "Influence of soil moisture on urban microclimate and surface-layer meteorology in Oklahoma City". *Journal of Applied Meteorology and Climatology* 53 (1) : 83-98. ISSN : 1558-8424. doi:[10.1175/JAMC-D-13-0156.1](https://doi.org/10.1175/JAMC-D-13-0156.1). Cité p. 17.
- HUTCHINS, M. G., *et al.* 2017. "Integrated modeling in urban hydrology : reviewing the role of monitoring technology in overcoming the issue of 'big data' requirements". *WIREs Water* 4 (1) : e1177. ISSN : 2049-1948. doi:[10.1002/wat2.1177](https://doi.org/10.1002/wat2.1177). Cité p. 43.
- IM, S., *et al.* 2008. "Assessing the impacts of land use changes on watershed hydrology using MIKE SHE". *Environmental Geology* 57 (1) : 231. ISSN : 1866-6299. doi:[10.1007/s00254-008-1303-3](https://doi.org/10.1007/s00254-008-1303-3). Cité p. 48.
- IORGULESCU, I., et K. J. BEVEN. 2004. "Nonparametric direct mapping of rainfall-runoff relationships : An alternative approach to data analysis and modeling?" *Water Resources Research* 40 (8). ISSN : 1944-7973. doi:[10.1029/2004WR003094](https://doi.org/10.1029/2004WR003094). Cité p. 221.
- ISIK, S., *et al.* 2013. "Modeling effects of changing land use/cover on daily streamflow : An Artificial Neural Network and curve number based hybrid approach". *Journal of Hydrology, Hydrology of peri-urban catchments : processes and modelling*, 485:103-112. ISSN : 0022-1694. doi:[10.1016/j.jhydro1.2012.08.032](https://doi.org/10.1016/j.jhydro1.2012.08.032). Cité pp. 145, 248.
- JACOBSON, C. R. 2011. "Identification and quantification of the hydrological impacts of imperviousness in urban catchments : A review". *Journal of Environmental Management* 92 (6) : 1438-1448. ISSN : 0301-4797. doi:[10.1016/j.jenvman.2011.01.018](https://doi.org/10.1016/j.jenvman.2011.01.018). Cité pp. 4, 5, 16, 259.
- JAKEMAN, A. J., et G. M. HORNBERGER. 1993. "How much complexity is warranted in a rainfall-runoff model?" *Water Resources Research* 29 (8) : 2637-2649. ISSN : 1944-7973. doi:[10.1029/93WR00877](https://doi.org/10.1029/93WR00877). Cité p. 55.
- JAKEMAN, A. J., I. G. LITTLEWOOD et P. G. WHITEHEAD. 1990. "Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments". *Journal of Hydrology* 117 (1) : 275-300. ISSN : 0022-1694. doi:[10.1016/0022-1694\(90\)90097-H](https://doi.org/10.1016/0022-1694(90)90097-H). Cité pp. 9, 173, 313.
- JANKOWFSKY, S., *et al.* 2014. "Assessing anthropogenic influence on the hydrology of small peri-urban catchments : Development of the object-oriented PUMMA model by integrating urban and rural hydrological models". *Journal of Hydrology* 517 :1056-1071. ISSN : 0022-1694. doi:[10.1016/j.jhydro1.2014.06.034](https://doi.org/10.1016/j.jhydro1.2014.06.034). Cité pp. 49, 228.
- JANKOWFSKY, S., *et al.* 2013. "Comparison of catchment and network delineation approaches in complex suburban environments : application to the Chaudanne catchment, France". *Hydrological Processes* 27 (25) : 3747-3761. ISSN : 1099-1085. doi:[10.1002/hyp.9506](https://doi.org/10.1002/hyp.9506). Cité p. 61.
- JAVELLE, P., *et al.* 2010. "Flash flood warning at ungauged locations using radar rainfall and antecedent soil moisture estimations". *Journal of Hydrology, Flash Floods : Observations and Analysis of Hydrometeorological Controls*, 394 (1) : 267-274. ISSN : 0022-1694. doi:[10.1016/j.jhydro1.2010.03.032](https://doi.org/10.1016/j.jhydro1.2010.03.032). Cité p. 146.
- JHA, A. K., R. BLOCH et J. LAMOND. 2012. *Cities and flooding*. The World Bank. ISBN : 978-0-8213-8866-2. doi:[10.1596/978-0-8213-8866-2](https://doi.org/10.1596/978-0-8213-8866-2). Cité p. 32.
- JIA, Y., *et al.* 2001. "Development of WEP model and its application to an urban watershed". *Hydrological Processes* 15 (11) : 2175-2194. ISSN : 1099-1085. doi:[10.1002/hyp.275](https://doi.org/10.1002/hyp.275). Cité pp. 46, 228.
- JIM, C. Y. 1998. "Physical and chemical properties of a Hong Kong roadside soil in relation to urban tree growth". *Urban Ecosystems* 2 (2) : 171-181. ISSN : 1573-1642. doi:[10.1023/A:1009585700191](https://doi.org/10.1023/A:1009585700191). Cité p. 18.
- JIN, M., J. M. SHEPHERD et M. D. KING. 2005. "Urban aerosols and their variations with clouds and rainfall : A case study for New York and Houston". *Journal of Geophysical Research : Atmospheres* 110 (D10). ISSN : 2156-2202. doi:[10.1029/2004JD005081](https://doi.org/10.1029/2004JD005081). Cité p. 17.
- JOTHITYANGKOON, C., M. SIVAPALAN et D. L. FARMER. 2001. "Process controls of water balance variability in a large semi-arid catchment : downward approach to hydrological model development". *Journal of Hydrology* 254 (1) : 174-198. ISSN : 0022-1694. doi:[10.1016/S0022-1694\(01\)00496-6](https://doi.org/10.1016/S0022-1694(01)00496-6). Cité p. 248.

- KALANTARI, Z., *et al.* 2017. "Urbanization development under climate change : Hydrological responses in a peri-urban Mediterranean catchment". *Land Degradation & Development* 28 (7) : 2207-2221. ISSN : 1099-145X. doi : [10.1002/ldr.2747](https://doi.org/10.1002/ldr.2747). Cité p. 38.
- KARLSSON, I. B., *et al.* 2016. "Combined effects of climate models, hydrological model structures and land use scenarios on hydrological impacts of climate change". *Journal of Hydrology* 535 :301-317. ISSN : 0022-1694. doi : [10.1016/j.jhydrol.2016.01.069](https://doi.org/10.1016/j.jhydrol.2016.01.069). Cité p. 264.
- KAUFFMAN, G. J., *et al.* 2009. "Link between impervious cover and base flow in the White Clay Creek Wild and Scenic Watershed in Delaware". *Journal of Hydrologic Engineering* 14 (4) : 324-334. doi : [10.1061/\(ASCE\)1084-0699\(2009\)14:4\(324\)](https://doi.org/10.1061/(ASCE)1084-0699(2009)14:4(324)). Cité p. 36.
- KIM, U., et J. J. KALUARACHCHI. 2008. "Application of parameter estimation and regionalization methodologies to ungauged basins of the Upper Blue Nile River Basin, Ethiopia". *Journal of Hydrology* 362 (1) : 39-56. ISSN : 0022-1694. doi : [10.1016/j.jhydrol.2008.08.016](https://doi.org/10.1016/j.jhydrol.2008.08.016). Cité p. 198.
- KIM, Y.-Y., K.-K. LEE et I. SUNG. 2001. "Urbanization and the groundwater budget, metropolitan Seoul area, Korea". *Hydrogeology Journal* 9 (4) : 401-412. ISSN : 1435-0157. doi : [10.1007/s100400100139](https://doi.org/10.1007/s100400100139). Cité p. 20.
- KIRCHNER, J. W. 2009. "Catchments as simple dynamical systems : Catchment characterization, rainfall-runoff modeling, and doing hydrology backward". *Water Resources Research* 45 (2). ISSN : 1944-7973. doi : [10.1029/2008WR006912](https://doi.org/10.1029/2008WR006912). Cité p. 145.
- . 2006. "Getting the right answers for the right reasons : Linking measurements, analyses, and models to advance the science of hydrology". *Water Resources Research* 42 (3). ISSN : 1944-7973. doi : [10.1029/2005WR004362](https://doi.org/10.1029/2005WR004362). Cité pp. 65, 89, 228.
- KJELDSEN, T. R. 2010. "Modelling the impact of urbanization on flood frequency relationships in the UK". *Hydrology Research* 41 (5) : 391-405. ISSN : 0029-1277. doi : [10.2166/nh.2010.056](https://doi.org/10.2166/nh.2010.056). Cité pp. 31, 53.
- KJELDSEN, T. R., J. D. MILLER et J. C. PACKMAN. 2013. "Modelling design flood hydrographs in catchments with mixed urban and rural land cover". *Hydrology Research* 44 (6) : 1040-1057. ISSN : 0029-1277. doi : [10.2166/nh.2013.158](https://doi.org/10.2166/nh.2013.158). Cité pp. 7, 53, 59, 200, 229, 250, 262.
- KLEIDORFER, M., *et al.* 2009. "Impact of input data uncertainties on urban stormwater model parameters". *Water Science and Technology* 60 (6) : 1545-1554. ISSN : 0273-1223. doi : [10.2166/wst.2009.493](https://doi.org/10.2166/wst.2009.493). Cité p. 51.
- KLEIN, R. D. 1979. "Urbanization and stream quality impairment". *JAWRA Journal of the American Water Resources Association* 15 (4) : 948-963. ISSN : 1752-1688. doi : [10.1111/j.1752-1688.1979.tb01074.x](https://doi.org/10.1111/j.1752-1688.1979.tb01074.x). Cité p. 36.
- KLEMEŠ, V. 1986. "Operational testing of hydrological simulation models". *Hydrological Sciences Journal* 31 (1) : 13-24. ISSN : 0262-6667. doi : [10.1080/02626668609491024](https://doi.org/10.1080/02626668609491024). Cité pp. 175, 198, 237, 250.
- KNOBEN, W. J. M., *et al.* 2019. "Modular Assessment of Rainfall-Runoff Models Toolbox (MARRMoT) v1.2 : an open-source, extendable framework providing implementations of 46 conceptual hydrologic models as continuous state-space formulations". *Geoscientific Model Development* 12 (6) : 2463-2480. ISSN : 1991-959X. doi : [10.5194/gmd-12-2463-2019](https://doi.org/10.5194/gmd-12-2463-2019). Cité p. 248.
- KOHLER, M. A., et R. K. LINSLEY. 1951. *Predicting the runoff from storm rainfall*. Weather Bureau Research Papers. U.S. Department of Commerce, Weather Bureau. Cité p. 145.
- KONRAD, C. P., et D. B. BOOTH. 2005. "Hydrologic changes in urban streams and their ecological significance". In *American Fisheries Society Symposium*, 47 :157-177. Cité pp. 35, 36.
- KONRAD, C. P., D. B. BOOTH et S. J. BURGESS. 2005. "Effects of urban development in the Puget Lowland, Washington, on interannual streamflow patterns : Consequences for channel form and streambed disturbance". *Water Resources Research* 41 (7). ISSN : 1944-7973. doi : [10.1029/2005WR004097](https://doi.org/10.1029/2005WR004097). Cité p. 19, 33.

- KRAUSE, P., D. P. BOYLE et F. BÄSE. 2005. "Comparison of different efficiency criteria for hydrological model assessment". In *Advances in Geosciences*, 5 :89-97. Copernicus GmbH. doi :10.5194/adgeo-5-89-2005. Cité p. 250.
- KURSA, M. B., et W. R. RUDNICKI. 2020. *Boruta : Wrapper algorithm for all relevant feature selection*. Visité le 06/10/2020. <https://CRAN.R-project.org/package=Boruta>. Cité p. 150.
- . 2010. "Feature selection with the Boruta package". *Journal of Statistical Software* 36 (1) : 1-13. ISSN : 1548-7660. doi :10.18637/jss.v036.i11. Cité p. 150.
- LANCIA, M., et al. 2019. "Analysis of groundwater resources in densely populated urban watersheds with a complex tectonic setting : Shenzhen, southern China". *Hydrogeology Journal* 27 (1) : 183-194. ISSN : 1435-0157. doi :10.1007/s10040-018-1867-2. Cité p. 186.
- LANGANKE, T., et al. 2016. *Copernicus Land Monitoring Service – High Resolution Layer Imperviousness : product specifications document*. Rapport technique. European Environment Agency. <https://land.copernicus.eu/user-corner/technical-library/hrl-imperviousness-technical-document-product-2015>. Cité p. 91.
- LARRARTE, F. 2006. "Velocity fields within sewers : An experimental study". *Flow Measurement and Instrumentation* 17 (5) : 282-290. ISSN : 0955-5986. doi :10.1016/j.flowmeasinst.2006.08.001. Cité p. 19.
- LE MOINE, N. 2008. "Le bassin versant de surface vu par le souterrain : une voie d'amélioration des performances et du réalisme des modèles pluie-débit?" Thèse de doctorat, Université Pierre et Marie Curie-Paris VI. Cité pp. 15, 28, 66, 88, 112, 207, 262, 308.
- LE MOINE, N., et al. 2007. "How can rainfall-runoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments". *Water Resources Research* 43 (6). ISSN : 1944-7973. doi :10.1029/2006WR005608. Cité pp. 233, 251.
- LEBAUER, D. 2019. *An R package for using the Harmonized World Soil Database (HWSD) : dlebauer/rhwsd*. Visité le 29/05/2019. <https://github.com/dlebauer/rhwsd>. Cité p. 206.
- LEBECHEREL, L., V. ANDRÉASSIAN et C. PERRIN. 2013. "On regionalizing the Turc-Mezentsev water balance formula". *Water Resources Research* 49 (11) : 7508-7517. ISSN : 1944-7973. doi :10.1002/2013WR013575. Cité pp. 170, 211.
- LEE, J.-Y., et al. 2005. "Evaluation of hydrologic data obtained from a local groundwater monitoring network in a metropolitan city, Korea". *Hydrological Processes* 19 (13) : 2525-2537. ISSN : 1099-1085. doi :10.1002/hyp.5689. Cité p. 19.
- LEE, J. G., et J. P. HEANEY. 2003. "Estimation of urban imperviousness and its impacts on storm water systems". *Journal of Water Resources Planning and Management* 129 (5) : 419-426. doi :10.1061/(ASCE)0733-9496(2003)129:5(419). Cité p. 30.
- LEGATES, D. R., et G. J. MCCABE. 1999. "Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation". *Water Resources Research* 35 (1) : 233-241. ISSN : 1944-7973. doi :10.1029/1998WR900018. Cité p. 250.
- LELEU, I., et al. 2014. "La refonte du système d'information national pour la gestion et la mise à disposition des données hydrométriques". *La Houille Blanche*, n° 1 : 25-32. ISSN : 0018-6368, 1958-5551. doi :10.1051/lhb/2014004. Cité pp. 68, 77, 171, 201, 230.
- LEMONSU, A., V. MASSON et E. BERTHIER. 2007. "Improvement of the hydrological component of an urban soil-vegetation-atmosphere-transfer model". *Hydrological Processes* 21 (16) : 2100-2111. ISSN : 1099-1085. doi :10.1002/hyp.6373. Cité p. 45.
- LEOPOLD, L. B. 1968. "Hydrology for urban land planning : A guidebook on the hydrologic effects of urban land use". *Geological Survey Circular* 554. Cité p. 186.

- LEPEŠKA, T., *et al.* 2020. "Urbanization—Its hidden impact on water losses : Prądnik river basin, Lesser Poland". *Water* 12 (7) : 1958. doi : [10.3390/w12071958](https://doi.org/10.3390/w12071958). Cité p. 36.
- LERAT, J., *et al.* 2012. "Do internal flow measurements improve the calibration of rainfall-runoff models?" *Water Resources Research* 48 (2). ISSN : 1944-7973. doi : [10.1029/2010WR010179](https://doi.org/10.1029/2010WR010179). Cité p. 239.
- LERNER, D. N. 1990. "Groundwater recharge in urban areas". *Atmospheric Environment. Part B. Urban Atmosphere* 24 (1) : 29-33. ISSN : 0957-1272. doi : [10.1016/0957-1272\(90\)90006-G](https://doi.org/10.1016/0957-1272(90)90006-G). Cité pp. 19, 20.
- . 2002. "Identifying and quantifying urban recharge : a review". *Hydrogeology Journal* 10 (1) : 143-152. ISSN : 1435-0157. doi : [10.1007/s10040-001-0177-1](https://doi.org/10.1007/s10040-001-0177-1). Cité p. 20.
- LHOMME, J., C. BOUVIER et J.-L. PERRIN. 2004. "Applying a GIS-based geomorphological routing model in urban catchments". *Journal of Hydrology, Urban Hydrology*, 299 (3) : 203-216. ISSN : 0022-1694. doi : [10.1016/j.jhydrol.2004.08.006](https://doi.org/10.1016/j.jhydrol.2004.08.006). Cité pp. 47, 50, 250.
- LI, C., *et al.* 2020. "Impacts of urbanization on watershed water balances across the conterminous United States". *Water Resources Research* 56 (7) : e2019WR026574. ISSN : 1944-7973. doi : [10.1029/2019WR026574](https://doi.org/10.1029/2019WR026574). Cité p. 16.
- LIAKOS, K. G., *et al.* 2018. "Machine learning in agriculture : A review". *Sensors* 18 (8) : 2674. doi : [10.3390/s18082674](https://doi.org/10.3390/s18082674). Cité p. 199.
- LIAN, X., *et al.* 2018. "Partitioning global land evapotranspiration using CMIP5 models constrained by observations". *Nature Climate Change* 8 (7) : 640-646. ISSN : 1758-6798. doi : [10.1038/s41558-018-0207-9](https://doi.org/10.1038/s41558-018-0207-9). Cité p. 15.
- LIAW, A., et M. WIENER. 2002. "Classification and regression by randomForest". *R News* 2 (3) : 18-22. ISSN : 1609-3631. Cité pp. 150, 209.
- LIN, K., *et al.* 2014. "Xinjiang model combined with Curve Number to simulate the effect of land use change on environmental flow". *Journal of Hydrology* 519 : 3142-3152. ISSN : 0022-1694. doi : [10.1016/j.jhydrol.2014.10.049](https://doi.org/10.1016/j.jhydrol.2014.10.049). Cité p. 228.
- LIN, Y., et K. E. MITCHELL. 2005. "The NCEP stage II/IV hourly precipitation analyses : Development and applications". In *Proceedings of the 19th conference on Hydrology*, 9-13. Citeseer. Cité pp. 70, 77, 138, 171.
- LIU, J., et D. NIYOGLI. 2019. "Meta-analysis of urbanization impact on rainfall modification". *Scientific Reports* 9 (1) : 1-14. ISSN : 2045-2322. doi : [10.1038/s41598-019-42494-2](https://doi.org/10.1038/s41598-019-42494-2). Cité p. 17.
- LIVNEH, B., *et al.* 2015. "A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and Southern Canada 1950–2013". *Scientific Data* 2 (1) : 150042. ISSN : 2052-4463. doi : [10.1038/sdata.2015.42](https://doi.org/10.1038/sdata.2015.42). Cité p. 69.
- LOBLIGEIS, F. 2014. "Mieux connaître la distribution spatiale des pluies améliore-t-il la modélisation des crues? Diagnostic sur 181 bassins versants français". Thèse de doctorat, AgroParisTech (Paris). Cité pp. 71, 74, 83, 88.
- MANSELL, M., et F. ROLLET. 2006. "Water balance and the behaviour of different paving surfaces". *Water and Environment Journal* 20 (1) : 7-10. ISSN : 1747-6593. doi : [10.1111/j.1747-6593.2005.00015.x](https://doi.org/10.1111/j.1747-6593.2005.00015.x). Cité p. 18.
- MARTIN, E. H., C. KELLEHER et T. WAGENER. 2012. "Has urbanization changed ecological streamflow characteristics in Maine (USA)?" *Hydrological Sciences Journal* 57 (7) : 1337-1354. ISSN : 0262-6667. doi : [10.1080/02626667.2012.707318](https://doi.org/10.1080/02626667.2012.707318). Cité pp. 167, 169.
- MASSMANN, C. 2020. "Identification of factors influencing hydrologic model performance using a top-down approach in a large number of U.S. catchments". *Hydrological Processes* 34 (1) : 4-20. ISSN : 1099-1085. doi : [10.1002/hyp.13566](https://doi.org/10.1002/hyp.13566). Cité p. 248.

- MATHEVET, T. 2005. "Quels modèles pluie-débit globaux pour le pas de temps horaire? Développement empirique et comparaison de modèles sur un large échantillon de bassins versants". Thèse de doctorat, ENGREF (Paris). Cité pp. 174, 207, 262, 308, 313.
- MCDONNELL, J. J., *et al.* 2007. "Moving beyond heterogeneity and process complexity : A new vision for watershed hydrology". *Water Resources Research* 43 (7). ISSN : 1944-7973. doi :[10.1029/2006WR005467](https://doi.org/10.1029/2006WR005467). Cité pp. 15, 250.
- MCGRANE, S. J. 2016. "Impacts of urbanisation on hydrological and water quality dynamics, and urban water management : a review". *Hydrological Sciences Journal* 61 (13) : 2295-2311. ISSN : 0262-6667. doi :[10.1080/02626667.2015.1128084](https://doi.org/10.1080/02626667.2015.1128084). Cité pp. 5, 22, 31, 34, 158, 227, 259.
- MCINTYRE, N., et M. MARSHALL. 2010. "Identification of rural land management signals in runoff response". *Hydrological Processes* 24 (24) : 3521-3534. ISSN : 1099-1085. doi :[10.1002/hyp.7774](https://doi.org/10.1002/hyp.7774). Cité p. 187.
- MCINTYRE, N., *et al.* 2014. "Modelling the hydrological impacts of rural land use change". *Hydrology Research* 45 (6) : 737-754. ISSN : 0029-1277. doi :[10.2166/nh.2013.145](https://doi.org/10.2166/nh.2013.145). Cité pp. 39, 228.
- MCPMAHON, G., *et al.* 2003. "Use of stage data to characterize hydrologic conditions in an urbanizing environment". *JAWRA Journal of the American Water Resources Association* 39 (6) : 1529-1546. ISSN : 1752-1688. doi :[10.1111/j.1752-1688.2003.tb04437.x](https://doi.org/10.1111/j.1752-1688.2003.tb04437.x). Cité pp. 30, 31, 37.
- MCMILLAN, H. K., *et al.* 2011. "Hydrological field data from a modeller's perspective : Part 1. Diagnostic tests for model structure". *Hydrological Processes* 25 (4) : 511-522. ISSN : 1099-1085. doi :[10.1002/hyp.7841](https://doi.org/10.1002/hyp.7841). Cité p. 248.
- MEI, Y., et E. N. ANAGNOSTOU. 2015. "A hydrograph separation method based on information from rainfall and runoff records". *Journal of Hydrology* 523 :636-649. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2015.01.083](https://doi.org/10.1016/j.jhydrol.2015.01.083). Cité pp. 103, 145, 172, 232, 298.
- MEJÍA, A. I., et G. E. MOGLEN. 2010. "Impact of the spatial distribution of imperviousness on the hydrologic response of an urbanizing basin". *Hydrological Processes* 24 (23) : 3359-3373. ISSN : 1099-1085. doi :[10.1002/hyp.7755](https://doi.org/10.1002/hyp.7755). Cité pp. 30, 48, 161, 220.
- MEJÍA, A. I., *et al.* 2015. "Anthropogenic controls from urban growth on flow regimes". *Advances in Water Resources* 84 :125-135. ISSN : 0309-1708. doi :[10.1016/j.advwatres.2015.08.010](https://doi.org/10.1016/j.advwatres.2015.08.010). Cité pp. 35, 36, 187.
- MERZ, R., G. BLÖSCHL et J. PARAJKA. 2006a. "Spatio-temporal variability of event runoff coefficients". *Journal of Hydrology* 331 (3) : 591-604. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2006.06.008](https://doi.org/10.1016/j.jhydrol.2006.06.008). Cité p. 146.
- MERZ, R., et G. BLÖSCHL. 2009. "A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria". *Water Resources Research* 45 (1). ISSN : 1944-7973. doi :[10.1029/2008WR007163](https://doi.org/10.1029/2008WR007163). Cité pp. 15, 136, 145, 160, 161, 298.
- . 2004. "Regionalisation of catchment model parameters". *Journal of Hydrology* 287 (1) : 95-123. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2003.09.028](https://doi.org/10.1016/j.jhydrol.2003.09.028). Cité pp. 198, 214, 251.
- MERZ, R., J. PARAJKA et G. BLÖSCHL. 2009. "Scale effects in conceptual hydrological modeling". *Water Resources Research* 45 (9). ISSN : 1944-7973. doi :[10.1029/2009WR007872](https://doi.org/10.1029/2009WR007872). Cité p. 79.
- MERZ, R., G. BLÖSCHL et J. D. PARAJKA. 2006b. "Regionalization methods in rainfall-runoff modelling using large catchment samples". *IAHS Publication* 307 :117-125. Cité p. 199.
- MERZ, R., J. PARAJKA et G. BLÖSCHL. 2011. "Time stability of catchment model parameters : Implications for climate impact analyses". *Water Resources Research* 47 (2). ISSN : 1944-7973. doi :[10.1029/2010WR009505](https://doi.org/10.1029/2010WR009505). Cité p. 187.
- MICHEL, C., V. ANDRÉASSIAN et C. PERRIN. 2005. "Soil Conservation Service Curve Number method : How to mend a wrong soil moisture accounting procedure?" *Water Resources Research* 41 (2). ISSN : 1944-7973. doi :[10.1029/2004WR003191](https://doi.org/10.1029/2004WR003191). Cité p. 248.

- MILLER, J. D., et S. GREBBY. 2014. "Mapping long-term temporal change in imperviousness using topographic maps". *International Journal of Applied Earth Observation and Geoinformation* 30 :9-20. ISSN : 0303-2434. doi :[10.1016/j.jag.2014.01.002](https://doi.org/10.1016/j.jag.2014.01.002). Cité p. 31.
- MILLER, J. D., et T. HESS. 2017. "Urbanisation impacts on storm runoff along a rural-urban gradient". *Journal of Hydrology* 552 :474-489. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2017.06.025](https://doi.org/10.1016/j.jhydrol.2017.06.025). Cité pp. 32, 33, 35, 37, 136, 137, 158, 160, 186, 227, 234, 237, 259.
- MILLER, J. D., *et al.* 2014. "Assessing the impact of urbanization on storm runoff in a peri-urban catchment using historical change in impervious cover". *Journal of Hydrology* 515 :59-70. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2014.04.011](https://doi.org/10.1016/j.jhydrol.2014.04.011). Cité pp. 35, 37, 39, 49.
- MITCHELL, V. G., et C. DIAPER. 2005. "UVQ : A tool for assessing the water and contaminant balance impacts of urban development scenarios". *Water Science and Technology* 52 (12) : 91-98. ISSN : 0273-1223. doi :[10.2166/wst.2005.0435](https://doi.org/10.2166/wst.2005.0435). Cité p. 52.
- MITCHELL, V. G., R. G. MEIN et T. A. MCMAHON. 2001. "Modelling the urban water cycle". *Environmental Modelling & Software* 16 (7) : 615-629. ISSN : 1364-8152. doi :[10.1016/S1364-8152\(01\)00029-9](https://doi.org/10.1016/S1364-8152(01)00029-9). Cité p. 52.
- MITCHELL, V. G., *et al.* 2008. "Linking urban water balance and energy balance models to analyse urban design options". *Hydrological Processes* 22 (16) : 2891-2900. ISSN : 1099-1085. doi :[10.1002/hyp.6868](https://doi.org/10.1002/hyp.6868). Cité p. 50.
- MONTANARI, A., *et al.* 2013. "'Panta Rhei—Everything Flows" : Change in hydrology and society—The IAHS Scientific Decade 2013–2022". *Hydrological Sciences Journal* 58 (6) : 1256-1275. ISSN : 0262-6667. doi :[10.1080/02626667.2013.809088](https://doi.org/10.1080/02626667.2013.809088). Cité p. 13.
- MORRIS, B., *et al.* 2007. "Using linked process models to improve urban groundwater management : an example from Doncaster England". *Water and Environment Journal* 21 (4) : 229-240. ISSN : 1747-6593. doi :[10.1111/j.1747-6593.2006.00067.x](https://doi.org/10.1111/j.1747-6593.2006.00067.x). Cité p. 52.
- MUÑOZ, P., *et al.* 2018. "Flash-flood forecasting in an Andean Mountain catchment—Development of a step-wise methodology based on the random forest algorithm". *Water* 10 (11) : 1519. doi :[10.3390/w10111519](https://doi.org/10.3390/w10111519). Cité p. 199.
- 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. ISSN : 0022-1694. doi :[10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6). Cité pp. 209, 238, 330.
- NAVRATIL, O., *et al.* 2013. "Hydrogeomorphic adjustments of stream channels disturbed by urban runoff (Yzeron River basin, France)". *Journal of Hydrology, Hydrology of peri-urban catchments : processes and modelling*, 485 :24-36. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2012.01.036](https://doi.org/10.1016/j.jhydrol.2012.01.036). Cité p. 19.
- NIEHOFF, D., U. FRITSCH et A. BRONSTERT. 2002. "Land-use impacts on storm-runoff generation : scenarios of land-use change and simulation of hydrological response in a meso-scale catchment in SW-Germany". *Journal of Hydrology, Advances in Flood Research*, 267 (1) : 80-93. ISSN : 0022-1694. doi :[10.1016/S0022-1694\(02\)00142-7](https://doi.org/10.1016/S0022-1694(02)00142-7). Cité pp. 39, 56, 228.
- NIYOGI, D., *et al.* 2011. "Urban modification of thunderstorms : An observational storm climatology and model case study for the Indianapolis Urban Region". *Journal of Applied Meteorology and Climatology* 50 (5) : 1129-1144. ISSN : 1558-8424. doi :[10.1175/2010JAMC1836.1](https://doi.org/10.1175/2010JAMC1836.1). Cité p. 17.
- NORBIATO, D., *et al.* 2009. "Controls on event runoff coefficients in the eastern Italian Alps". *Journal of Hydrology* 375 (3) : 312-325. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2009.06.044](https://doi.org/10.1016/j.jhydrol.2009.06.044). Cité pp. 145, 298.
- O'DRISCOLL, M., *et al.* 2010. "Urbanization effects on watershed hydrology and in-stream processes in the Southern United States". *Water* 2 (3) : 605-648. doi :[10.3390/w2030605](https://doi.org/10.3390/w2030605). Cité pp. 34, 80.
- OBLED, C., I. ZIN et B. HINGRAY. 2009. "Choix des pas de temps et d'espace pour des modélisations parcimonieuses en hydrologie des crues". *La Houille Blanche*, n° 5 : 81-87. ISSN : 0018-6368, 1958-5551. doi :[10.1051/lhb/2009059](https://doi.org/10.1051/lhb/2009059). Cité p. 130.

- OGDEN, F. L., *et al.* 2017. "Comment on "Beyond the SCS-CN method : A theoretical framework for spatially lumped rainfall-runoff response" by M. S. Bartlett et al." *Water Resources Research* 53 (7) : 6345-6350. ISSN : 1944-7973. doi :[10.1002/2016WR020176](https://doi.org/10.1002/2016WR020176). Cité pp. 55, 248.
- OGDEN, F. L., *et al.* 2011. "Relative importance of impervious area, drainage density, width function, and subsurface storm drainage on flood runoff from an urbanized catchment". *Water Resources Research* 47 (12). ISSN : 1944-7973. doi :[10.1029/2011WR010550](https://doi.org/10.1029/2011WR010550). Cité pp. 46, 228.
- OKE, T. R. 1982. "The energetic basis of the urban heat island". *Quarterly Journal of the Royal Meteorological Society* 108 (455) : 1-24. ISSN : 1477-870X. doi :[10.1002/qj.49710845502](https://doi.org/10.1002/qj.49710845502). Cité pp. 17, 18.
- OLDEN, J. D., et N. L. POFF. 2003. "Redundancy and the choice of hydrologic indices for characterizing streamflow regimes". *River Research and Applications* 19 (2) : 101-121. ISSN : 1535-1467. doi :[10.1002/rra.700](https://doi.org/10.1002/rra.700). Cité p. 198.
- OTT, B., et S. UHLENBROOK. 2004. "Quantifying the impact of land-use changes at the event and seasonal time scale using a process-oriented catchment model". *Hydrology and Earth System Sciences* 8 (1) : 62-78. ISSN : 1027-5606. doi :[10.5194/hess-8-62-2004](https://doi.org/10.5194/hess-8-62-2004). Cité p. 48.
- LOUDIN, L., *et al.* 2010. "Are seemingly physically similar catchments truly hydrologically similar?" *Water Resources Research* 46 (11). ISSN : 1944-7973. doi :[10.1029/2009WR008887](https://doi.org/10.1029/2009WR008887). Cité pp. 198, 200.
- LOUDIN, L., *et al.* 2006a. "Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations". *Water Resources Research* 42 (7). ISSN : 1944-7973. doi :[10.1029/2005WR004636](https://doi.org/10.1029/2005WR004636). Cité pp. 175, 238, 239.
- LOUDIN, L., *et al.* 2006b. "How informative is land-cover for the regionalization of the GR4J rainfall-runoff model? Lessons of a downward approach". *IAHS Publication* 307 :246-255. Cité pp. 198, 200.
- LOUDIN, L., *et al.* 2018. "Hydrological impacts of urbanization at the catchment scale". *Journal of Hydrology* 559 :774-786. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2018.02.064](https://doi.org/10.1016/j.jhydrol.2018.02.064). Cité pp. 16, 30, 31, 32, 35, 36, 97, 161, 168, 186, 200, 201, 220, 251, 258, 264.
- LOUDIN, L., *et al.* 2004. "Locating the sources of low-pass behavior within rainfall-runoff models". *Water Resources Research* 40 (11). ISSN : 1944-7973. doi :[10.1029/2004WR003291](https://doi.org/10.1029/2004WR003291). Cité p. 75.
- LOUDIN, L., *et al.* 2008. "Spatial proximity, physical similarity, regression and ungauged catchments : A comparison of regionalization approaches based on 913 French catchments". *Water Resources Research* 44 (3). ISSN : 1944-7973. doi :[10.1029/2007WR006240](https://doi.org/10.1029/2007WR006240). Cité pp. 198, 200, 211.
- LOUDIN, L., *et al.* 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) : 290-306. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2004.08.026](https://doi.org/10.1016/j.jhydrol.2004.08.026). Cité pp. 72, 74, 141, 172, 204, 232.
- PARAJKA, J., R. MERZ et G. BLÖSCHL. 2005. "A comparison of regionalisation methods for catchment model parameters". *Hydrology and Earth System Sciences* 9 (3) : 157-171. ISSN : 1027-5606. doi :[10.5194/hess-9-157-2005](https://doi.org/10.5194/hess-9-157-2005). Cité p. 198.
- PARAJKA, J., G. BLÖSCHL et R. MERZ. 2007. "Regional calibration of catchment models : Potential for ungauged catchments". *Water Resources Research* 43 (6). ISSN : 1944-7973. doi :[10.1029/2006WR005271](https://doi.org/10.1029/2006WR005271). Cité p. 198.
- PATHIRAJA, S., *et al.* 2018. "Time-varying parameter models for catchments with land use change : the importance of model structure". *Hydrology and Earth System Sciences* 22 (5) : 2903-2919. ISSN : 1027-5606. doi :[10.5194/hess-22-2903-2018](https://doi.org/10.5194/hess-22-2903-2018). Cité pp. 39, 187, 228.
- PAUL, M. J., et J. L. MEYER. 2001. "Streams in the urban landscape". *Annual Review of Ecology and Systematics* 32 (1) : 333-365. doi :[10.1146/annurev.ecolsys.32.081501.114040](https://doi.org/10.1146/annurev.ecolsys.32.081501.114040). Cité pp. 16, 19.

- PAYAN, J.-L. 2007. "Prise en compte des barrages-réservoirs dans un modèle global pluie-débit". Thèse de doctorat, ENGREF (Paris). Cité p. 78.
- PEREDO, D. 2019. *Approche probabiliste de la prévision immédiate des risques d'inondation sur des bassins versants soumis aux crues éclair*. Rapport de comité de thèse. Antony, France : IRSTEA. Cité pp. 326, 331.
- PERRIN, C. 2000. "Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative". Thèse de doctorat, Grenoble INPG. Cité pp. 174, 313.
- PERRIN, C., C. MICHEL et V. ANDRÉASSIAN. 2003. "Improvement of a parsimonious model for streamflow simulation". *Journal of Hydrology* 279 (1) : 275-289. ISSN : 0022-1694. doi : [10.1016/S0022-1694\(03\)00225-7](https://doi.org/10.1016/S0022-1694(03)00225-7). Cité pp. 200, 207, 208, 233, 234, 237, 305, 308, 309.
- PERRIN, C., *et al.* 2007. "Impact of limited streamflow data on the efficiency and the parameters of rainfall—runoff models". *Hydrological Sciences Journal* 52 (1) : 131-151. ISSN : 0262-6667. doi : [10.1623/hysj.52.1.131](https://doi.org/10.1623/hysj.52.1.131). Cité pp. 79, 170, 230.
- PETCHPRAYOON, P., *et al.* 2010. "Hydrological impacts of land use/land cover change in a large river basin in central-northern Thailand". *International Journal of Climatology* 30 (13) : 1917-1930. ISSN : 1097-0088. doi : [10.1002/joc.2131](https://doi.org/10.1002/joc.2131). Cité pp. 35, 37.
- PETRUCCI, G., et C. BONHOMME. 2014. "The dilemma of spatial representation for urban hydrology semi-distributed modelling : Trade-offs among complexity, calibration and geographical data". *Journal of Hydrology* 517 :997-1007. ISSN : 0022-1694. doi : [10.1016/j.jhydrol.2014.06.019](https://doi.org/10.1016/j.jhydrol.2014.06.019). Cité pp. 50, 56, 58, 228.
- POFF, N. L., B. P. BLEDSOE et C. O. CUHACIYAN. 2006. "Hydrologic variation with land use across the contiguous United States : Geomorphic and ecological consequences for stream ecosystems". *Geomorphology*, 37th Binghamton Geomorphology Symposium, 79 (3) : 264-285. ISSN : 0169-555X. doi : [10.1016/j.geomorph.2006.06.032](https://doi.org/10.1016/j.geomorph.2006.06.032). Cité pp. 16, 34.
- PONCELET, C. 2016. "Du bassin au paramètre : jusqu'où peut-on régionaliser un modèle hydrologique conceptuel?". Thèse de doctorat, Université Pierre et Marie Curie-Paris VI. Cité pp. 15, 102, 203.
- PRASKIEVICZ, S., et H. CHANG. 2009. "A review of hydrological modelling of basin-scale climate change and urban development impacts". *Progress in Physical Geography : Earth and Environment* 33 (5) : 650-671. ISSN : 0309-1333. doi : [10.1177/0309133309348098](https://doi.org/10.1177/0309133309348098). Cité p. 16.
- PRICE, K. 2011. "Effects of watershed topography, soils, land use, and climate on baseflow hydrology in humid regions : A review". *Progress in Physical Geography : Earth and Environment* 35 (4) : 465-492. ISSN : 0309-1333. doi : [10.1177/0309133311402714](https://doi.org/10.1177/0309133311402714). Cité p. 136.
- PRIETO, C., *et al.* 2019. "Flow prediction in ungauged catchments using probabilistic Random Forests regionalization and new statistical adequacy tests". *Water Resources Research*. ISSN : 1944-7973. doi : [10.1029/2018WR023254](https://doi.org/10.1029/2018WR023254). Cité p. 199.
- PROSDOCIMI, I., T. R. KJELDSSEN et J. D. MILLER. 2015. "Detection and attribution of urbanization effect on flood extremes using nonstationary flood-frequency models". *Water Resources Research* 51 (6) : 4244-4262. ISSN : 1944-7973. doi : [10.1002/2015WR017065](https://doi.org/10.1002/2015WR017065). Cité pp. 35, 38, 228.
- PUTRO, B., *et al.* 2016. "An empirical investigation of climate and land-use effects on water quantity and quality in two urbanising catchments in the southern United Kingdom". *Science of The Total Environment* 548-549 :164-172. ISSN : 0048-9697. doi : [10.1016/j.scitotenv.2015.12.132](https://doi.org/10.1016/j.scitotenv.2015.12.132). Cité p. 37.
- R CORE TEAM. 2019. *R : A language and environment for statistical computing*. Vienna, Austria : R Foundation for Statistical Computing. Cité pp. 68, 201, 209.
- RAGAB, R., *et al.* 2003a. "Experimental study of water fluxes in a residential area : 1. Rainfall, roof runoff and evaporation : the effect of slope and aspect". *Hydrological Processes* 17 (12) : 2409-2422. ISSN : 1099-1085. doi : [10.1002/hyp.1250](https://doi.org/10.1002/hyp.1250). Cité pp. 4, 18.

- RAGAB, R., *et al.* 2003b. "Experimental study of water fluxes in a residential area : 2. Road infiltration, runoff and evaporation". *Hydrological Processes* 17 (12) : 2423-2437. ISSN : 1099-1085. doi :[10.1002/hyp.1251](https://doi.org/10.1002/hyp.1251). Cité pp. 4, 18, 161.
- RAMIER, D., E. BERTHIER et H. ANDRIEU. 2004. "An urban lysimeter to assess runoff losses on asphalt concrete plates". *Physics and Chemistry of the Earth, Parts A/B/C, Anthropogenic impacts on catchment processes*, 29 (11) : 839-847. ISSN : 1474-7065. doi :[10.1016/j.pce.2004.05.011](https://doi.org/10.1016/j.pce.2004.05.011). Cité pp. 4, 18.
- . 2011. "The hydrological behaviour of urban streets : long-term observations and modelling of runoff losses and rainfall–runoff transformation". *Hydrological Processes* 25 (14) : 2161-2178. ISSN : 1099-1085. doi :[10.1002/hyp.7968](https://doi.org/10.1002/hyp.7968). Cité pp. 18, 19, 161.
- RAZAVI, T., et P. COULIBALY. 2013. "Streamflow prediction in ungauged basins : Review of regionalization methods". *Journal of Hydrologic Engineering* 18 (8) : 958-975. ISSN : 1943-5584. doi :[10.1061/\(ASCE\)HE.1943-5584.0000690](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000690). Cité pp. 220, 261.
- READ, J. S., *et al.* 2015. "geoknife : Reproducible web-processing of large gridded datasets". *Ecography*. doi :[10.1111/ecog.01880](https://doi.org/10.1111/ecog.01880). Cité pp. 70, 138, 141, 171, 204, 230.
- REDFERN, T. W., *et al.* 2016. "Current understanding of hydrological processes on common urban surfaces". *Progress in Physical Geography : Earth and Environment* 40 (5) : 699-713. ISSN : 0309-1333. doi :[10.1177/0309133316652819](https://doi.org/10.1177/0309133316652819). Cité pp. 5, 19, 30, 39, 161, 258, 259.
- ROCHE, P.-A., J. MIQUEL et E. GAUME. 2012. "Processus physiques à l'échelle des bassins versants". In *Hydrologie quantitative : Processus, modèles et aide à la décision*, sous la dir. de P.-A. ROCHE, J. MIQUEL et E. GAUME, 41-94. Ingénierie et développement durable. Paris : Springer. ISBN : 978-2-8178-0106-3. doi :[10.1007/978-2-8178-0106-3_2](https://doi.org/10.1007/978-2-8178-0106-3_2). Cité p. 15.
- RODRIGUEZ-ITURBE, I., et J. B. VALDÉS. 1979. "The geomorphologic structure of hydrologic response". *Water Resources Research* 15 (6) : 1409-1420. ISSN : 1944-7973. doi :[10.1029/WR015i006p01409](https://doi.org/10.1029/WR015i006p01409). Cité p. 15.
- RODRIGUEZ, E., H. ANDRIEU et Y. ZECH. 2000. "Evaluation of a distributed model for urban catchments using a 7-year continuous data series". *Hydrological Processes* 14 (5) : 899-914. ISSN : 1099-1085. doi :[10.1002/\(SICI\)1099-1085\(20000415\)14:5<899::AID-HYP977>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1099-1085(20000415)14:5<899::AID-HYP977>3.0.CO;2-R). Cité p. 44.
- RODRIGUEZ, E., H. ANDRIEU et J.-D. CREUTIN. 2003. "Surface runoff in urban catchments : morphological identification of unit hydrographs from urban databanks". *Journal of Hydrology* 283 (1) : 146-168. ISSN : 0022-1694. doi :[10.1016/S0022-1694\(03\)00246-4](https://doi.org/10.1016/S0022-1694(03)00246-4). Cité pp. 19, 49, 56, 228, 232.
- RODRIGUEZ, E., H. ANDRIEU et E. MORENA. 2008. "A distributed hydrological model for urbanized areas – Model development and application to case studies". *Journal of Hydrology* 351 (3) : 268-287. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2007.12.007](https://doi.org/10.1016/j.jhydrol.2007.12.007). Cité pp. 49, 173.
- RODRIGUEZ, E., E. BOCHER et K. CHANCIBAULT. 2013. "Terrain representation impact on periurban catchment morphological properties". *Journal of Hydrology*, Hydrology of peri-urban catchments : processes and modelling, 485 :54-67. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2012.11.023](https://doi.org/10.1016/j.jhydrol.2012.11.023). Cité pp. 61, 112.
- RODRIGUEZ, E., *et al.* 2020. "Groundwater contribution to sewer network baseflow in an urban catchment-case study of Pin Sec catchment, Nantes, France". *Water* 12 (3) : 689. doi :[10.3390/w12030689](https://doi.org/10.3390/w12030689). Cité p. 20.
- ROSE, S., et N. E. PETERS. 2001. "Effects of urbanization on streamflow in the Atlanta area (Georgia, USA) : a comparative hydrological approach". *Hydrological Processes* 15 (8) : 1441-1457. ISSN : 1099-1085. doi :[10.1002/hyp.218](https://doi.org/10.1002/hyp.218). Cité pp. 33, 35, 36, 37.
- ROUGÉ, C., et X. CAI. 2014. "Crossing-scale hydrological impacts of urbanization and climate variability in the Greater Chicago area". *Journal of Hydrology* 517 :13-27. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2014.05.005](https://doi.org/10.1016/j.jhydrol.2014.05.005). Cité pp. 35, 36, 37.

- ROY, A. H., et W. D. SHUSTER. 2009. "Assessing impervious surface connectivity and applications for watershed management". *JAWRA Journal of the American Water Resources Association* 45 (1) : 198-209. ISSN : 1752-1688. doi : [10.1111/j.1752-1688.2008.00271.x](https://doi.org/10.1111/j.1752-1688.2008.00271.x). Cité p. 30.
- SAADI, M., L. OUDIN et P. RIBSTEIN. 2020. "Crossing the rural-urban boundary in hydrological modelling : How do conceptual rainfall-runoff models handle the specificities of urbanized catchments?" *Hydrological Processes* 34 (15) : 3331-3346. ISSN : 1099-1085. doi : [10.1002/hyp.13808](https://doi.org/10.1002/hyp.13808). Cité pp. 228, 229, 233, 237.
- . 2019. "Random forest ability in regionalizing hourly hydrological model parameters". *Water* 11 (8) : 1540. doi : [10.3390/w11081540](https://doi.org/10.3390/w11081540). Cité pp. 137, 189, 230, 251.
- SALAVATI, B. 2015. "Impact de l'urbanisation sur la réponse hydrologique des bassins versants urbains". Thèse de doctorat, Université Pierre et Marie Curie-Paris VI. Cité pp. 5, 14, 32, 39, 66, 258.
- SALAVATI, B., *et al.* 2016. "Modeling approaches to detect land-use changes : Urbanization analyzed on a set of 43 US catchments". *Journal of Hydrology* 538 :138-151. ISSN : 0022-1694. doi : [10.1016/j.jhydrol.2016.04.010](https://doi.org/10.1016/j.jhydrol.2016.04.010). Cité pp. 35, 36, 39, 200, 220, 227.
- SALVADOR, R., C. BAUTISTA-CAPETILLO et E. PLAYÁN. 2011. "Irrigation performance in private urban landscapes : A study case in Zaragoza (Spain)". *Landscape and Urban Planning* 100 (3) : 302-311. ISSN : 0169-2046. doi : [10.1016/j.landurbplan.2010.12.018](https://doi.org/10.1016/j.landurbplan.2010.12.018). Cité p. 18.
- SALVADORE, E., J. BRONDERS et O. BATELAAN. 2015. "Hydrological modelling of urbanized catchments : A review and future directions". *Journal of Hydrology* 529 :62-81. ISSN : 0022-1694. doi : [10.1016/j.jhydrol.2015.06.028](https://doi.org/10.1016/j.jhydrol.2015.06.028). Cité pp. 6, 7, 13, 18, 19, 28, 30, 42, 44, 109, 160, 168, 169, 170, 173, 186, 206, 220, 228, 230, 232, 259, 260, 261.
- SAMANIEGO, L., R. KUMAR et S. ATTINGER. 2010. "Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale". *Water Resources Research* 46 (5). ISSN : 1944-7973. doi : [10.1029/2008WR007327](https://doi.org/10.1029/2008WR007327). Cité p. 198.
- SANTAMOURIS, M. 2015. "Analyzing the heat island magnitude and characteristics in one hundred Asian and Australian cities and regions". *Science of The Total Environment* 512-513 :582-598. ISSN : 0048-9697. doi : [10.1016/j.scitotenv.2015.01.060](https://doi.org/10.1016/j.scitotenv.2015.01.060). Cité p. 18.
- SANTOS, L. 2018. "Que peut-on attendre des Super Modèles en hydrologie? Évaluation d'une approche de combinaison dynamique de modèles pluie-débit". Thèse de doctorat, AgroParisTech (Paris). Cité p. 322.
- SANTOS, L., G. THIREL et C. PERRIN. 2018. "Technical note : Pitfalls in using log-transformed flows within the KGE criterion". *Hydrology and Earth System Sciences* 22 (8) : 4583-4591. ISSN : 1027-5606. doi : [10.5194/hess-22-4583-2018](https://doi.org/10.5194/hess-22-4583-2018). Cité pp. 175, 238.
- SANZANA, P., *et al.* 2019. "Impact of urban growth and high residential irrigation on streamflow and groundwater levels in a peri-urban semiarid catchment". *JAWRA Journal of the American Water Resources Association* 55 (3) : 720-739. ISSN : 1752-1688. doi : [10.1111/1752-1688.12743](https://doi.org/10.1111/1752-1688.12743). Cité pp. 33, 36, 53, 186, 228.
- SCHAEFLI, B., et H. V. GUPTA. 2007. "Do Nash values have value?" *Hydrological Processes* 21 (15) : 2075-2080. ISSN : 1099-1085. doi : [10.1002/hyp.6825](https://doi.org/10.1002/hyp.6825). Cité p. 250.
- SCHNEIDER, A., *et al.* 2017. "Global-scale river network extraction based on high-resolution topography and constrained by lithology, climate, slope, and observed drainage density". *Geophysical Research Letters* 44 (6) : 2773-2781. ISSN : 1944-8007. doi : [10.1002/2016GL071844](https://doi.org/10.1002/2016GL071844). Cité p. 112.
- SCHWARTZ, S. S., et B. SMITH. 2014. "Slowflow fingerprints of urban hydrology". *Journal of Hydrology* 515 :116-128. ISSN : 0022-1694. doi : [10.1016/j.jhydrol.2014.04.019](https://doi.org/10.1016/j.jhydrol.2014.04.019). Cité p. 36.
- SEFTON, C. E. M., et S. M. HOWARTH. 1998. "Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales". *Journal of Hydrology* 211 (1) : 1-16. ISSN : 0022-1694. doi : [10.1016/S0022-1694\(98\)00163-2](https://doi.org/10.1016/S0022-1694(98)00163-2). Cité p. 198.

- SEIBERT, J. 1999. "Regionalisation of parameters for a conceptual rainfall-runoff model". *Agricultural and Forest Meteorology* 98-99 :279-293. ISSN : 0168-1923. doi :[10.1016/S0168-1923\(99\)00105-7](https://doi.org/10.1016/S0168-1923(99)00105-7). Cité p. 198.
- SEIBERT, J., et J. J. MCDONNELL. 2010. "Land-cover impacts on streamflow : a change-detection modelling approach that incorporates parameter uncertainty". *Hydrological Sciences Journal* 55 (3) : 316-332. ISSN : 0262-6667. doi :[10.1080/02626661003683264](https://doi.org/10.1080/02626661003683264). Cité p. 187.
- SEIBERT, J., et H. J. v. MEERVELD. 2016. "Hydrological change modeling : Challenges and opportunities". *Hydrological Processes* 30 (26) : 4966-4971. ISSN : 1099-1085. doi :[10.1002/hyp.10999](https://doi.org/10.1002/hyp.10999). Cité p. 39.
- SELVALINGAM, S., S. Y. LIONG et P. C. MANOHARAN. 1987. "Use of RORB and SWMM models to an urban catchment in Singapore". *Advances in Water Resources* 10 (2) : 78-86. ISSN : 0309-1708. doi :[10.1016/0309-1708\(87\)90012-1](https://doi.org/10.1016/0309-1708(87)90012-1). Cité p. 50.
- SHEEDER, S. A., J. D. ROSS et T. N. CARLSON. 2002. "Dual urban and rural hydrograph signals in three small watersheds". *JAWRA Journal of the American Water Resources Association* 38 (4) : 1027-1040. ISSN : 1752-1688. doi :[10.1111/j.1752-1688.2002.tb05543.x](https://doi.org/10.1111/j.1752-1688.2002.tb05543.x). Cité p. 30.
- SHEN, C. 2018. "A transdisciplinary review of deep learning research and its relevance for water resources scientists". *Water Resources Research* 54 (11) : 8558-8593. ISSN : 1944-7973. doi :[10.1029/2018WR022643](https://doi.org/10.1029/2018WR022643). Cité pp. 198, 199.
- SHEPHERD, J. M. 2005. "A review of current investigations of urban-induced rainfall and recommendations for the future". *Earth Interactions* 9 (12) : 1-27. ISSN : 1087-3562. doi :[10.1175/EI156.1](https://doi.org/10.1175/EI156.1). Cité p. 17.
- SHEPHERD, J. M., H. PIERCE et A. J. NEGRI. 2002. "Rainfall modification by major urban areas : Observations from spaceborne rain radar on the TRMM satellite". *Journal of Applied Meteorology* 41 (7) : 689-701. ISSN : 0894-8763. doi :[10.1175/1520-0450\(2002\)041<0689:RMBMUA>2.0.CO;2](https://doi.org/10.1175/1520-0450(2002)041<0689:RMBMUA>2.0.CO;2). Cité p. 17.
- SHUSTER, W. D., R. GEHRING et J. GERKEN. 2007. "Prospects for enhanced groundwater recharge via infiltration of urban storm water runoff : A case study". *Journal of Soil and Water Conservation* 62 (3) : 129-137. ISSN : 0022-4561, 1941-3300. Cité p. 20.
- SHUSTER, W. D., *et al.* 2005. "Impacts of impervious surface on watershed hydrology : A review". *Urban Water Journal* 2 (4) : 263-275. ISSN : 1573-062X. doi :[10.1080/15730620500386529](https://doi.org/10.1080/15730620500386529). Cité p. 30.
- SILLANPÄÄ, N., et H. KOIVUSALO. 2015. "Impacts of urban development on runoff event characteristics and unit hydrographs across warm and cold seasons in high latitudes". *Journal of Hydrology* 521 :328-340. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2014.12.008](https://doi.org/10.1016/j.jhydrol.2014.12.008). Cité pp. 35, 37, 39, 78, 158, 162, 186, 239, 259.
- SIMMONS, D. L., et R. J. REYNOLDS. 1982. "Effects of urbanization on base flow of selected south-shore streams, Long Island, New York". *JAWRA Journal of the American Water Resources Association* 18 (5) : 797-805. ISSN : 1752-1688. doi :[10.1111/j.1752-1688.1982.tb00075.x](https://doi.org/10.1111/j.1752-1688.1982.tb00075.x). Cité pp. 33, 36.
- SINGH, R., B. MAHESHWARI et H. M. MALANO. 2009. "Developing a conceptual model for water accounting in peri-urban catchments". In *18th World IMACS Congress and MODSIM09*, 13-17. Cairns, Australia. Cité pp. 7, 52, 59, 262.
- SINGH, V. P., et D. A. WOOLHISER. 2002. "Mathematical modeling of watershed hydrology". *Journal of Hydrologic Engineering* 7 (4) : 270-292. ISSN : 1084-0699. doi :[10.1061/\(ASCE\)1084-0699\(2002\)7:4\(270\)](https://doi.org/10.1061/(ASCE)1084-0699(2002)7:4(270)). Cité pp. 15, 41, 220.
- SIVAPALAN, M., *et al.* 2003. "IAHS Decade on Predictions in Ungauged Basins (PUB), 2003-2012 : Shaping an exciting future for the hydrological sciences". *Hydrological Sciences Journal* 48 (6) : 857-880. ISSN : 0262-6667. doi :[10.1623/hysj.48.6.857.51421](https://doi.org/10.1623/hysj.48.6.857.51421). Cité p. 198.
- SMITH, B. K., *et al.* 2013. "Spectrum of storm event hydrologic response in urban watersheds". *Water Resources Research* 49 (5) : 2649-2663. ISSN : 1944-7973. doi :[10.1002/wrcr.20223](https://doi.org/10.1002/wrcr.20223). Cité pp. 35, 136, 158, 259.

- SMITH, J. A., *et al.* 2002. "The regional hydrology of extreme floods in an urbanizing drainage basin". *Journal of Hydrometeorology* 3 (3) : 267-282. ISSN : 1525-755X. doi :[10.1175/1525-7541\(2002\)003<0267:TRHOEF>2.0.CO;2](https://doi.org/10.1175/1525-7541(2002)003<0267:TRHOEF>2.0.CO;2). Cité pp. 35, 38.
- SOPHOCLEOUS, M. 2002. "Interactions between groundwater and surface water : the state of the science". *Hydrogeology Journal* 10 (1) : 52-67. ISSN : 1435-0157. doi :[10.1007/s10040-001-0170-8](https://doi.org/10.1007/s10040-001-0170-8). Cité p. 15.
- SOULIS, K. X., *et al.* 2009. "Investigation of the direct runoff generation mechanism for the analysis of the SCS-CN method applicability to a partial area experimental watershed". *Hydrology and Earth System Sciences* 13 (5) : 605-615. ISSN : 1027-5606. doi :[10.5194/hess-13-605-2009](https://doi.org/10.5194/hess-13-605-2009). Cité p. 145.
- SOULSBY, C., C. BIRKEL et D. TETZLAFF. 2014. "Assessing urbanization impacts on catchment transit times". *Geophysical Research Letters* 41 (2) : 442-448. ISSN : 1944-8007. doi :[10.1002/2013GL058716](https://doi.org/10.1002/2013GL058716). Cité p. 37.
- STAVROPULOS-LAFFAILLE, X., *et al.* 2018. "Improvements to the hydrological processes of the Town Energy Balance model (TEB-Veg, SURFEX v7.3) for urban modelling and impact assessment". *Geoscientific Model Development* 11 (10) : 4175-4194. ISSN : 1991-959X. doi :[10.5194/gmd-11-4175-2018](https://doi.org/10.5194/gmd-11-4175-2018). Cité pp. 49, 228.
- SULTANA, Z., *et al.* 2018. "Assessment of business interruption of flood-affected companies using random forests". *Water* 10 (8) : 1049. doi :[10.3390/w10081049](https://doi.org/10.3390/w10081049). Cité p. 199.
- SUN, S., *et al.* 2017. "Urban hydrologic trend analysis based on rainfall and runoff data analysis and conceptual model calibration". *Hydrological Processes* 31 (6) : 1349-1359. ISSN : 1099-1085. doi :[10.1002/hyp.11109](https://doi.org/10.1002/hyp.11109). Cité pp. 35, 37, 51.
- SURVEY, U. G. 1994. *Nationwide summary of US Geological Survey regional regression equations for estimating magnitude and frequency of floods for ungaged sites, 1993*. Water-Resources Investigations Report 94-4002. Reston, VA. Visité le 16/10/2020. doi :[10.3133/wri944002](https://doi.org/10.3133/wri944002). <https://pubs.er.usgs.gov/publication/wri944002>. Cité p. 136.
- TABARY, P., *et al.* 2013. "A 10-year (1997–2006) reanalysis of Quantitative Precipitation Estimation over France : methodology and first results". *IAHS Publication* 351 :255-260. Cité pp. 71, 77, 171, 204, 230.
- TAHA, H. 1997. "Urban climates and heat islands : albedo, evapotranspiration, and anthropogenic heat". *Energy and Buildings* 25 (2) : 99-103. ISSN : 0378-7788. doi :[10.1016/S0378-7788\(96\)00999-1](https://doi.org/10.1016/S0378-7788(96)00999-1). Cité p. 17.
- TARASOVA, L., *et al.* 2018a. "Exploring controls on rainfall-runoff events : 1. Time series-based event separation and temporal dynamics of event runoff response in Germany". *Water Resources Research* 54 (10) : 7711-7732. ISSN : 1944-7973. doi :[10.1029/2018WR022587](https://doi.org/10.1029/2018WR022587). Cité pp. 84, 136, 145, 161, 298.
- TARASOVA, L., *et al.* 2018b. "Exploring controls on rainfall-runoff events : 2. Regional patterns and spatial controls of event characteristics in Germany". *Water Resources Research* 54 (10) : 7688-7710. ISSN : 1944-7973. doi :[10.1029/2018WR022588](https://doi.org/10.1029/2018WR022588). Cité pp. 136, 160.
- TE LINDE, A. H., J. C. J. H. AERTS et J. C. J. KWADIJK. 2010. "Effectiveness of flood management measures on peak discharges in the Rhine basin under climate change". *Journal of Flood Risk Management* 3 (4) : 248-269. ISSN : 1753-318X. doi :[10.1111/j.1753-318X.2010.01076.x](https://doi.org/10.1111/j.1753-318X.2010.01076.x). Cité p. 52.
- TETZLAFF, D., M. GROTTKER et C. LEIBUNDGUT. 2005. "Hydrological criteria to assess changes of flow dynamic in urban impacted catchments". *Physics and Chemistry of the Earth, Parts A/B/C, Integrated Water Resource Assessment*, 30 (6) : 426-431. ISSN : 1474-7065. doi :[10.1016/j.pce.2005.06.008](https://doi.org/10.1016/j.pce.2005.06.008). Cité pp. 35, 37.
- THIELEN, J., et J. -.-D. CREUTIN. 1997. "An urban hydrological model with high spatial resolution rainfall from a meteorological model". *Journal of Hydrology* 200 (1) : 58-83. ISSN : 0022-1694. doi :[10.1016/S0022-1694\(97\)00002-4](https://doi.org/10.1016/S0022-1694(97)00002-4). Cité p. 51.
- THORNTON, P. E., *et al.* 2016. "Daymet : Daily surface weather data on a 1-km grid for North America, version 3". *ORNL DAAC*. doi :[10.3334/ORNLDAAC/1328](https://doi.org/10.3334/ORNLDAAC/1328). Cité pp. 70, 72, 77, 138, 141, 171, 204, 230.

- TONG, S. T. Y. 1990. "The hydrologic effects of urban land use : A case study of the little Miami River Basin". *Landscape and Urban Planning* 19 (1) : 99-105. ISSN : 0169-2046. doi :[10.1016/0169-2046\(90\)90037-3](https://doi.org/10.1016/0169-2046(90)90037-3). Cité p. 35.
- TROCH, P. A., *et al.* 2015. "Catchment coevolution : A useful framework for improving predictions of hydrological change?" *Water Resources Research* 51 (7) : 4903-4922. ISSN : 1944-7973. doi :[10.1002/2015WR017032](https://doi.org/10.1002/2015WR017032). Cité p. 16.
- TYRALIS, H., G. PAPACHARALAMPOUS et A. LANGOUSIS. 2019. "A brief review of random forests for water scientists and practitioners and their recent history in water resources". *Water* 11 (5) : 910. doi :[10.3390/w11050910](https://doi.org/10.3390/w11050910). Cité pp. 198, 199, 200, 209, 220, 221.
- UNITED NATIONS. 2019. *World Urbanization Prospects : The 2018 Revision, Highlights*. Department of Economic and Social Affairs, Population Division, United Nations. New York. ISBN : 978-92-1-148318-5. Cité pp. 3, 4.
- VALEO, C., et S. M. A. MOIN. 2000. "Grid-resolution effects on a model for integrating urban and rural areas". *Hydrological Processes* 14 (14) : 2505-2525. ISSN : 1099-1085. doi :[10.1002/1099-1085\(20001015\)14:14<2505::AID-HYP111>3.0.CO;2-3](https://doi.org/10.1002/1099-1085(20001015)14:14<2505::AID-HYP111>3.0.CO;2-3). Cité pp. 45, 59.
- VALÉRY, A. 2010. "Modélisation précipitations - débit sous influence nivale. Élaboration d'un module neige et évaluation sur 380 bassins versants". Thèse de doctorat, AgroParisTech (Paris). Cité p. 78.
- VALÉRY, A., V. ANDRÉASSIAN et C. PERRIN. 2014. "'As simple as possible but not simpler' : What is useful in a temperature-based snow-accounting routine? Part 2 – Sensitivity analysis of the Cemaneige snow accounting routine on 380 catchments". *Journal of Hydrology* 517 :1176-1187. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2014.04.058](https://doi.org/10.1016/j.jhydrol.2014.04.058). Cité p. 78.
- VAN LOON, A. F., *et al.* 2019. "Using paired catchments to quantify the human influence on hydrological droughts". *Hydrology and Earth System Sciences* 23 (3) : 1725-1739. ISSN : 1027-5606. doi :[10.5194/hess-23-1725-2019](https://doi.org/10.5194/hess-23-1725-2019). Cité p. 169.
- VANDEWIELE, G. L., et A. ELIAS. 1995. "Monthly water balance of ungauged catchments obtained by geographical regionalization". *Journal of Hydrology* 170 (1) : 277-291. ISSN : 0022-1694. doi :[10.1016/0022-1694\(95\)02681-E](https://doi.org/10.1016/0022-1694(95)02681-E). Cité p. 198.
- VERBEIREN, B., *et al.* 2013. "Assessing urbanisation effects on rainfall-runoff using a remote sensing supported modelling strategy". *International Journal of Applied Earth Observation and Geoinformation* 21 :92-102. ISSN : 0303-2434. doi :[10.1016/j.jag.2012.08.011](https://doi.org/10.1016/j.jag.2012.08.011). Cité pp. 32, 47.
- VERDIN, K. L. 2017. *Hydrologic Derivatives for Modeling and Analysis—A new global high-resolution database*. USGS Numbered Series 1053. Reston, VA : U.S. Geological Survey. Visité le 25/10/2018. <http://pubs.er.usgs.gov/publication/ds1053>. Cité pp. 111, 112, 143, 205.
- VERSINI, P.-A., *et al.* 2015. "Assessment of the hydrological impacts of green roof : From building scale to basin scale". *Journal of Hydrology* 524 :562-575. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2015.03.020](https://doi.org/10.1016/j.jhydrol.2015.03.020). Cité p. 21.
- VIDAL, J.-P., *et al.* 2010. "A 50-year high-resolution atmospheric reanalysis over France with the Safran system". *International Journal of Climatology* 30 (11) : 1627-1644. ISSN : 1097-0088. doi :[10.1002/joc.2003](https://doi.org/10.1002/joc.2003). Cité pp. 72, 77, 171, 204, 230.
- VIETZ, G. J., *et al.* 2014. "Ecologically relevant geomorphic attributes of streams are impaired by even low levels of watershed effective imperviousness". *Geomorphology* 206 :67-78. ISSN : 0169-555X. doi :[10.1016/j.geomorph.2013.09.019](https://doi.org/10.1016/j.geomorph.2013.09.019). Cité p. 19.
- VÖRÖSMARTY, C. J., *et al.* 2010. "Global threats to human water security and river biodiversity". *Nature* 467 (7315) : 555-561. ISSN : 1476-4687. doi :[10.1038/nature09440](https://doi.org/10.1038/nature09440). Cité p. 136.

- VOS, N. J. de, T. H. M. RIENTJES et H. V. GUPTA. 2010. "Diagnostic evaluation of conceptual rainfall-runoff models using temporal clustering". *Hydrological Processes* 24 (20) : 2840-2850. ISSN : 1099-1085. doi :[10.1002/hyp.7698](https://doi.org/10.1002/hyp.7698). Cité p. 265.
- VREBOS, D., *et al.* 2014. "Water displacement by sewer infrastructure in the Grote Nete catchment, Belgium, and its hydrological regime effects". *Hydrology and Earth System Sciences* 18 (3) : 1119-1136. ISSN : 1027-5606. doi :[10.5194/hess-18-1119-2014](https://doi.org/10.5194/hess-18-1119-2014). Cité p. 48.
- VUORINEN, H. S., P. S. JUUTI et T. S. KATKO. 2007. "History of water and health from ancient civilizations to modern times". *Water Supply* 7 (1) : 49-57. ISSN : 1606-9749. doi :[10.2166/ws.2007.006](https://doi.org/10.2166/ws.2007.006). Cité p. 3.
- WAKODE, H. B., *et al.* 2018. "Impact of urbanization on groundwater recharge and urban water balance for the city of Hyderabad, India". *International Soil and Water Conservation Research* 6 (1) : 51-62. ISSN : 2095-6339. doi :[10.1016/j.iswcr.2017.10.003](https://doi.org/10.1016/j.iswcr.2017.10.003). Cité p. 20.
- WALSH, C. J., *et al.* 2005. "The urban stream syndrome : current knowledge and the search for a cure". *Journal of the North American Benthological Society* 24 (3) : 706-723. ISSN : 0887-3593. doi :[10.1899/04-028.1](https://doi.org/10.1899/04-028.1). Cité pp. 16, 19.
- WANG, D., et M. HEJAZI. 2011. "Quantifying the relative contribution of the climate and direct human impacts on mean annual streamflow in the contiguous United States". *Water Resources Research* 47 (10). ISSN : 1944-7973. doi :[10.1029/2010WR010283](https://doi.org/10.1029/2010WR010283). Cité p. 37.
- WANG, Z., *et al.* 2015. "Flood hazard risk assessment model based on random forest". *Journal of Hydrology* 527 :1130-1141. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2015.06.008](https://doi.org/10.1016/j.jhydrol.2015.06.008). Cité p. 199.
- WEILER, M., et K. BEVEN. 2015. "Do we need a Community Hydrological Model?" *Water Resources Research* 51 (9) : 7777-7784. ISSN : 1944-7973. doi :[10.1002/2014WR016731](https://doi.org/10.1002/2014WR016731). Cité p. 41.
- WENG, Q., et D. LU. 2008. "A sub-pixel analysis of urbanization effect on land surface temperature and its interplay with impervious surface and vegetation coverage in Indianapolis, United States". *International Journal of Applied Earth Observation and Geoinformation* 10 (1) : 68-83. ISSN : 0303-2434. doi :[10.1016/j.jag.2007.05.002](https://doi.org/10.1016/j.jag.2007.05.002). Cité p. 17.
- WHEATER, H., et E. EVANS. 2009. "Land use, water management and future flood risk". *Land Use Policy, Land Use Futures*, 26 :S251-S264. ISSN : 0264-8377. doi :[10.1016/j.landusepol.2009.08.019](https://doi.org/10.1016/j.landusepol.2009.08.019). Cité p. 32.
- WICKHAM, J., *et al.* 2014. "The Multi-Resolution Land Characteristics (MRLC) Consortium — 20 years of development and integration of USA National Land Cover Data". *Remote Sensing* 6 (8) : 7424-7441. doi :[10.3390/rs6087424](https://doi.org/10.3390/rs6087424). Cité p. 230.
- WIJESEKARA, G. N., *et al.* 2012. "Assessing the impact of future land-use changes on hydrological processes in the Elbow River watershed in southern Alberta, Canada". *Journal of Hydrology, Hydrology Conference 2010*, 412-413 :220-232. ISSN : 0022-1694. doi :[10.1016/j.jhydrol.2011.04.018](https://doi.org/10.1016/j.jhydrol.2011.04.018). Cité p. 56.
- WILES, T. J., et J. M. SHARP. 2008. "The secondary permeability of impervious cover". *Environmental and Engineering Geoscience* 14 (4) : 251-265. ISSN : 1078-7275. doi :[10.2113/gsegeosci.14.4.251](https://doi.org/10.2113/gsegeosci.14.4.251). Cité p. 19.
- WINTER, T. C., *et al.* 1998. *Ground water and surface water : A single resource*. Report 1139. doi :[10.3133/cir1139](https://doi.org/10.3133/cir1139). <http://pubs.er.usgs.gov/publication/cir1139>. Cité p. 15.
- XIAO, Q., *et al.* 2007. "Hydrologic processes at the urban residential scale". *Hydrological Processes* 21 (16) : 2174-2188. ISSN : 1099-1085. doi :[10.1002/hyp.6482](https://doi.org/10.1002/hyp.6482). Cité p. 48.
- XU, Q., *et al.* 2014. "Review on water leakage control in distribution networks and the associated environmental benefits". *Journal of Environmental Sciences* 26 (5) : 955-961. ISSN : 1001-0742. doi :[10.1016/S1001-0742\(13\)60569-0](https://doi.org/10.1016/S1001-0742(13)60569-0). Cité pp. 18, 20.

- YADAV, M., T. WAGENER et H. GUPTA. 2007. "Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins". *Advances in Water Resources* 30 (8) : 1756-1774. ISSN : 0309-1708. doi : [10.1016/j.advwatres.2007.01.005](https://doi.org/10.1016/j.advwatres.2007.01.005). Cité p. 198.
- YANG, G., *et al.* 2011. "The impact of urban development on hydrologic regime from catchment to basin scales". *Landscape and Urban Planning* 103 (2) : 237-247. ISSN : 0169-2046. doi : [10.1016/j.landurbplan.2011.08.003](https://doi.org/10.1016/j.landurbplan.2011.08.003). Cité p. 158.
- YANG, J., *et al.* 2014. "Hydrologic response of a tropical watershed to urbanization". *Journal of Hydrology* 517 : 538-546. ISSN : 0022-1694. doi : [10.1016/j.jhydrol.2014.05.053](https://doi.org/10.1016/j.jhydrol.2014.05.053). Cité p. 35.
- YANG, L., *et al.* 2013. "Urbanization and climate change : An examination of nonstationarities in urban flooding". *Journal of Hydrometeorology* 14 (6) : 1791-1809. ISSN : 1525-755X. doi : [10.1175/JHM-D-12-095.1](https://doi.org/10.1175/JHM-D-12-095.1). Cité p. 35.
- YANG, Y., *et al.* 1999. "Quantification of groundwater recharge in the city of Nottingham, UK". *Environmental Geology* 38 (3) : 183-198. ISSN : 1432-0495. doi : [10.1007/s002540050414](https://doi.org/10.1007/s002540050414). Cité p. 20.
- YOUNG, P. 2003. "Top-down and data-based mechanistic modelling of rainfall-flow dynamics at the catchment scale". *Hydrological Processes* 17 (11) : 2195-2217. ISSN : 1099-1085. doi : [10.1002/hyp.1328](https://doi.org/10.1002/hyp.1328). Cité p. 248.
- ZECH, Y., *et al.* 1994. "Rainfall-runoff modelling of partly urbanized watersheds : Comparison between a distributed model using GIS and other models sensitivity analysis". *Water Science and Technology* 29 (1-2) : 163-170. ISSN : 0273-1223. doi : [10.2166/wst.1994.0662](https://doi.org/10.2166/wst.1994.0662). Cité p. 44.
- ZÉGRE, N., *et al.* 2010. "In lieu of the paired catchment approach : Hydrologic model change detection at the catchment scale". *Water Resources Research* 46 (11). ISSN : 1944-7973. doi : [10.1029/2009WR008601](https://doi.org/10.1029/2009WR008601). Cité pp. 40, 169.
- ZEKTSER, I. S., et L. G. EVERETT. 2004. *Groundwater resources of the world and their use*. T. IHP-VI, Series on Groundwater No. 6. Paris, France : UNESCO. ISBN : 978-92-9220-007-7. Cité p. 186.
- ZEKTSER, S., H. A. LOÁICIGA et J. T. WOLF. 2005. "Environmental impacts of groundwater overdraft : selected case studies in the southwestern United States". *Environmental Geology* 47 (3) : 396-404. ISSN : 1432-0495. doi : [10.1007/s00254-004-1164-3](https://doi.org/10.1007/s00254-004-1164-3). Cité p. 20.
- ZHANG, Y., et W. SHUSTER. 2014. "Impacts of spatial distribution of impervious areas on runoff response of hillslope catchments : Simulation study". *Journal of Hydrologic Engineering* 19 (6) : 1089-1100. ISSN : 1943-5584. doi : [10.1061/\(ASCE\)HE.1943-5584.0000905](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000905). Cité pp. 30, 39, 136, 158, 259.
- ZHANG, Y., *et al.* 2018. "Predicting runoff signatures using regression and hydrological modeling approaches". *Water Resources Research* 54 (10) : 7859-7878. ISSN : 1944-7973. doi : [10.1029/2018WR023325](https://doi.org/10.1029/2018WR023325). Cité pp. 150, 199, 209.
- ZHANG, Z., *et al.* 2008. "Reducing uncertainty in predictions in ungauged basins by combining hydrologic indices regionalization and multiobjective optimization". *Water Resources Research* 44 (12). ISSN : 1944-7973. doi : [10.1029/2008WR006833](https://doi.org/10.1029/2008WR006833). Cité p. 198.
- ZHOU, F., *et al.* 2013. "Hydrological response to urbanization at different spatio-temporal scales simulated by coupling of CLUE-S and the SWAT model in the Yangtze River Delta region". *Journal of Hydrology, Hydrology of peri-urban catchments : processes and modelling*, 485 : 113-125. ISSN : 0022-1694. doi : [10.1016/j.jhydrol.2012.12.040](https://doi.org/10.1016/j.jhydrol.2012.12.040). Cité p. 39.
- ZHOU, Y., *et al.* 2010. "Modeling watershed rainfall-runoff relations using impervious surface-area data with high spatial resolution". *Hydrogeology Journal* 18 (6) : 1413-1423. ISSN : 1435-0157. doi : [10.1007/s10040-010-0618-9](https://doi.org/10.1007/s10040-010-0618-9). Cité p. 47.
- ZHOU, Z., *et al.* 2017. "The complexities of urban flood response : Flood frequency analyses for the Charlotte metropolitan region". *Water Resources Research* 53 (8) : 7401-7425. ISSN : 1944-7973. doi : [10.1002/2016WR019997](https://doi.org/10.1002/2016WR019997). Cité pp. 35, 136, 158, 186, 234, 239, 259.

- ZIEGLER, A., et I. R. KÖNIG. 2014. "Mining data with random forests : current options for real-world applications". *Wiley Interdisciplinary Reviews : Data Mining and Knowledge Discovery* 4 (1) : 55-63. ISSN : 1942-4795. doi :[10.1002/widm.1114](https://doi.org/10.1002/widm.1114). Cité p. 209.
- ZOPPOU, C. 2001. "Review of urban storm water models". *Environmental Modelling & Software* 16 (3) : 195-231. ISSN : 1364-8152. doi :[10.1016/S1364-8152\(00\)00084-0](https://doi.org/10.1016/S1364-8152(00)00084-0). Cité p. 42.

ANNEXE A

MÉTHODE DE SÉPARATION DES ÉVÉNEMENTS

Many event-separation algorithms have been reported in the literature (Mei and Anagnostou, 2015; Merz and Blöschl, 2009; Norbiato *et al.*, 2009; Tarasova *et al.*, 2018a). The method adopted in this work is empirical, and consists of the following steps.

Step 1: Baseflow separation

Using numerical filtering-based methods is more practical when working on a large sample of catchments (Tarasova *et al.*, 2018a). We used a combination of the constant- k method (Blume, Zehe, and Bronstert, 2007) and a recursive digital filtering algorithm (Collischonn and Fan, 2013; Mei and Anagnostou, 2015; Eckhardt, 2005). We chose this method for its physical soundness: It hypothesizes that the baseflow is associated with the low-frequency component of the total streamflow. Tarasova *et al.* (2018a) compared five methods of baseflow extraction, and the method used herein has behaved acceptably.

Step 2: Estimation of characteristic response times

The direct runoff Q_d , that is considered to be the catchment response to rainfall events, was computed as the difference between the total flow and the baseflow. To constrain the event selection, we determined a set of catchment characteristic response times based on the cross-correlation of the time series of rainfall and direct flow, besides the direct flow auto-correlation.

Three characteristic response times were determined (Figure A.1):

- The lag time Lag (in hours), which is the positive shift Δt that maximizes the correlation between $Q_d(t + \Delta t)$ and $P(t)$.
- Sometimes, the correlogram of $Q_d(t + \Delta t)$ and $P(t)$ does not show a sharp maximum and can attain significantly high values for $\Delta t > Lag$ or $\Delta t < 0$. A second metric of response time Lag_{max} (in hours) was calculated to account for this issue. The correlogram is scaled by its maximum. Three limits were defined: (1) $Limit_1$ corresponded to the peak of the scaled correlogram for $\Delta t < 0$; (2) $Limit_2$ was taken as the peak of the scaled correlogram for $\Delta t > dec$, where dec corresponded to the moment when the correlogram reincreased for the first time; and (3) an arbitrary limit of $Limit_3 = 0.2$ (i.e., 20% of the maximum value of the correlogram). The maximum of these three limits, noted $LimitPQd$, was then explored in the scaled correlogram, and Lag_{max} was the shift corresponding to this limit.
- The auto-correlogram of Q_d was also investigated to determine the memory Mem (in

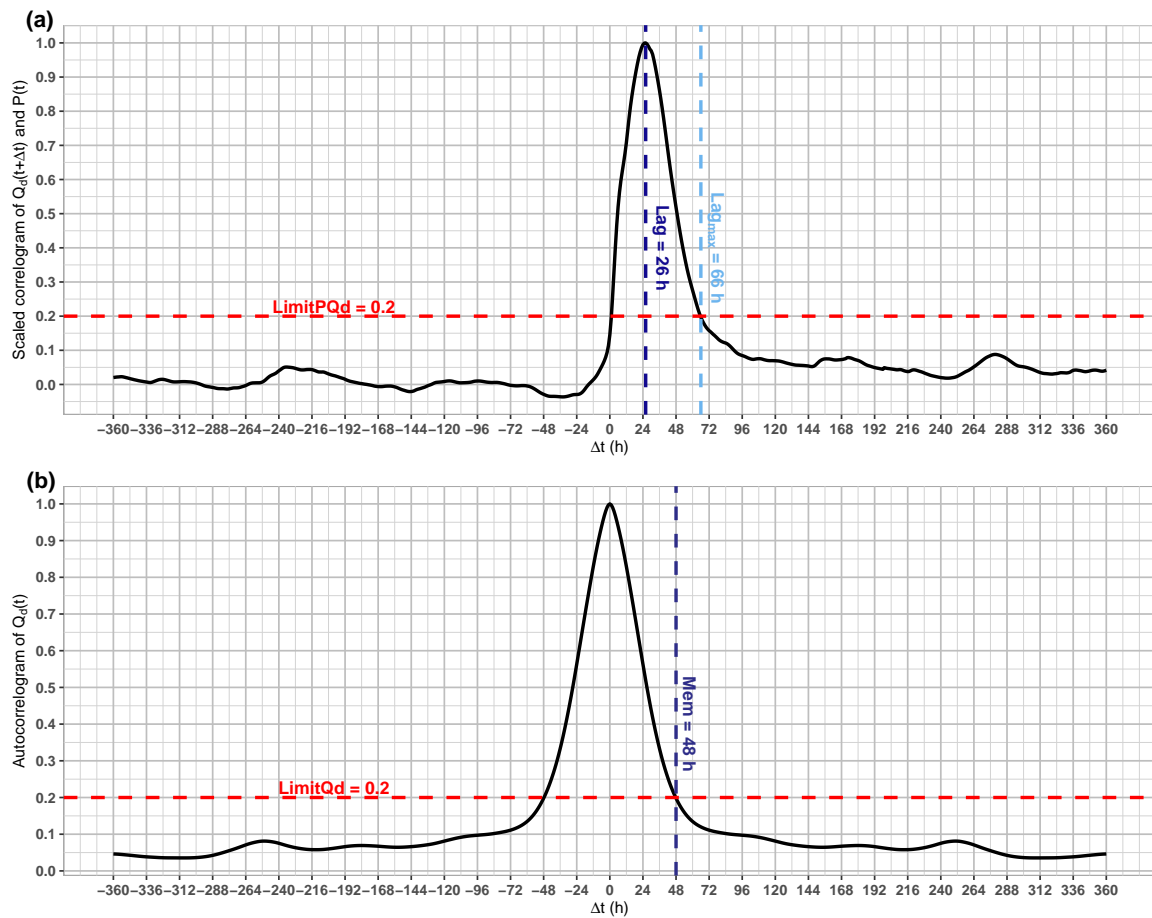


Figure A.1. Examples of correlograms computed for the catchment drained by the USGS station 01658000 Mattawoman Creek near Pomonkey, Maryland, which has an area of 143.6 km^2 . (a) Scaled correlogram of hourly direct discharge Q_d and hourly precipitation P , and (b) autocorrelogram of Q_d .

hours) of the catchment. Normally, the auto-correlogram should decrease for any $\Delta t > 0$. Mem corresponded to the peak of the auto-correlogram for $\Delta t > dec_{Q_d}$, where dec_{Q_d} is the moment at which the auto-correlogram reincreased for the first time. When no reincrease was registered, Mem corresponded to the shift at which an autocorrelation of $LimitQd = 0.2$ was attained for the first time.

Step 3: Estimation of the start of the event

The time series of $Q_d(t)$ were screened to look for a peak flow Q_d^p . Once found, a limit Q_d^{start} function of Q_d^p was determined. Over a time window TW_{start} , the first estimation of the start of the event corresponds to the moment when $Q_d(t) = Q_d^{start}$. The rainfall depth over a time window TWP_{start} before the start should be less than 10% of the rainfall depth between the start and the peak moments.

Step 4: Determination of the end of the event

As in step 3, a limit Q_d^{end} function of Q_d^p is determined. A time window TW_{end} after the peak is screened to identify the moment when $Q_d(t) = Q_d^{end}$. The rainfall depth between the peak and the end should be less than $p\%$ of the rainfall depth between the start and the peak.

Step 5: Additional conditions

Once the start and the end of an event are determined, the event must satisfy the following conditions: (1) At the peak moment, the ratio of the baseflow to the total flow should be less than the whole-period baseflow index of the catchment BFI , estimated from the baseflow separation step; (2) The rainfall and direct runoff time series should not contain missing values between the start and the end of the event; (3) The sum of the event rainfall depth should be higher than 2 mm; (4) The total volume of direct runoff should be less than the total volume of event rainfall (i.e., runoff ratio less than 1); and (5) The total duration of the event should be greater than 20% of Mem and less than 120% of Mem .

Finally, event RR is taken as the ratio of the direct runoff depth to the rainfall depth. Table A.1 summarizes the steps and the corresponding thresholds.

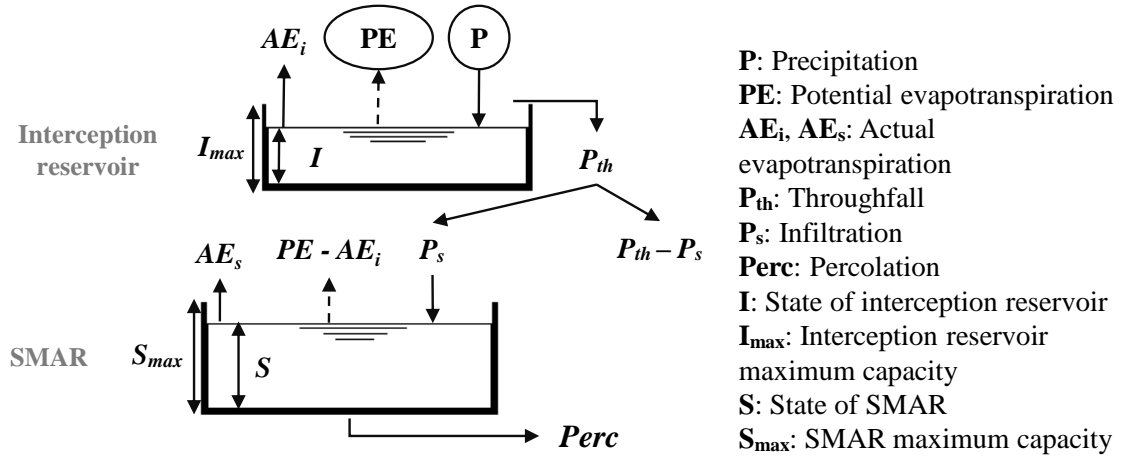
Table A.1. Summary of the steps of the event separation method.

Steps	Settings
Start of the event	<ul style="list-style-type: none"> • $Q_d^{start} = \max\left(\frac{Q_d^p}{10}, 0.05 \cdot (Q_d^p - Q_d^0)\right)$, where Q_d^0 is the minimum of $Q_d(t)$ for $t \in [t_{peak} - Lag - Lag_{max}, t_{peak}]$, where t_{peak} is the peak moment • $TW_{start} = [t_{peak} - Lag - Lag_{max}, t_{peak}]$ • $TWP_{start} = [t_{deb} - Lag_{max}, t_{deb}]$, where t_{deb} is the start moment
End of the event	<ul style="list-style-type: none"> • $Q_d^{end} = \max\left(\frac{Q_d^p}{10}, 0.05 \cdot (Q_d^p - Q_d^1)\right)$, where Q_d^1 is the minimum of $Q_d(t)$ for $t \in [t_{peak}, t_{peak} + Mem]$, where t_{peak} is the peak moment • $TW_{end} = [t_{peak}, t_{peak} + Mem]$ • $p = 50$
Additional conditions	<ul style="list-style-type: none"> • The ratio of baseflow to total flow at t_{peak} should be less than BFI • No missing values are allowed • Direct runoff depth should be less than the rainfall depth • The rainfall depth should be higher than 2 mm • Very short or very long events are eliminated (the event duration should be between $0.2 \cdot Mem$ and $1.2 \cdot Mem$)

ANNEXE **B**

ÉQUATIONS DU RÉSERVOIR DE L'HUMIDITÉ DU SOL

As a proxy for antecedent soil moisture (*ASM*), we used the soil moisture accounting reservoir (*SMAR*) of the GR4H hourly model (Ficchi, Perrin, and Andréassian, 2019). Detailed equations are given in Figure B.1. First, the hourly throughfall P_{th} is estimated using the equations under Step 1 (Figure B.1), where P is the hourly precipitation, PE the hourly potential evapotranspiration, and I the interception reservoir state. Second, when PE is not satisfied solely by the interception reservoir content, a depth AE_s is subtracted from the *SMAR* (Step 2). Third, we estimated the proportion of P_{th} that feeds the *SMAR*, noted P_s , which represents the amount of water that is expected to satisfy the initial losses through soil storage. Fourth, a depth representing the percolation is removed from the *SMAR*. P_s and AE_s are both a function of the state of the reservoir S , and its maximum capacity S_{max} . Finally, the pre-event catchment *ASM* is taken as the ratio S/S_{max} at the start of the event. This ratio varies between 0 and 1, where a value of 0 indicates a completely dry condition, and 1 indicates a completely wet condition.



Step 1: Estimation of throughfall P_{th}

- 1.1 $AE_i = \min(PE; I + P)$
- 1.2 $I := \max(0; I + P - PE)$
- 1.3 $P_{th} = 0$ if $I \leq I_{max}$ / $P_{th} = I - I_{max}$ if $I > I_{max}$
- 1.4 $I := I - P_{th}$

Step 2: Subtraction of PE-induced losses

$$AE_s = S * \frac{(2 - \frac{S}{S_{max}}) \tanh(\frac{PE - AE_i}{S_{max}})}{1 + (1 - \frac{S}{S_{max}}) \tanh(\frac{PE - AE_i}{S_{max}})} ; \quad S := S - AE_s$$

Step 3: Feeding the SMAR

$$P_s = S_{max} * \frac{(1 - (\frac{S}{S_{max}})^2) \tanh(\frac{P_{th}}{S_{max}})}{1 + \frac{S}{S_{max}} \tanh(\frac{P_{th}}{S_{max}})} ; \quad S := S + P_s$$

Step 4: Subtraction of percolation

$$Perc = S * \left(1 - \left(1 + \left(\frac{4}{21} \frac{S}{S_{max}} \right)^4 \right)^{-0.25} \right) ; \quad S := S - Perc$$

Figure B.1. Description of the soil moisture accounting reservoir (SMAR) equations. All variables are in millimeters. The interception reservoir maximum capacity I_{max} was estimated by matching the flux of throughfall at the hourly and daily time steps (Ficchi, Perrin, and Andréassian, 2019). Daily throughfall was estimated by comparing daily P with PE (Perrin, Michel, and Andréassian, 2003). For the 852 catchments used in this study, I_{max} was between 2.00 mm and 4.00 mm with a median value of 2.75 mm. I is initiated at 0 mm whereas S is initiated at $0.5 \cdot S_{max}$.

ANNEXE C

COMPLÉMENTS AU CHAPITRE 4

C.1 Équations du modèle GR4H

La structure du modèle GR4H (acronyme de **G**énie **R**ural à **4** paramètres au pas de temps **H**oraire) présentée dans Figure 4.3a est héritée du modèle journalier GR4J dont les équations peuvent être retrouvées dans Perrin, Michel et Andréassian (2003). Cette structure a été testée au pas de temps horaire par Mathevet (2005) et Le Moine (2008). Ficchi (2017) a modifié la structure du modèle en insérant un réservoir d'interception pour améliorer la cohérence des flux internes sur différents pas de temps.

C.1.1 Fonction de production

La fonction de production nous permet de déduire les précipitations nettes à partir des précipitations brutes en soustrayant les pertes par évapotranspiration réelle et infiltration dans le sol.

D'abord, on calcule l'évapotranspiration réelle à partir du réservoir d'interception E_i (mm) par l'équation :

$$E_i = \min(E, I + P) \quad (\text{C.1.1})$$

où E est la lame d'évapotranspiration potentielle (mm), P est la lame de précipitation (mm), et I (mm) le contenu du réservoir d'interception au début du pas de temps.

Quand $E_i = E$, l'évapotranspiration potentielle est complètement satisfaite par la somme du contenu du réservoir d'interception et la précipitation P . Dans le cas contraire, $I + P$ s'évapore complètement et le réservoir devient vide. Ainsi, son contenu est mis à jour par :

$$I := \max(0, I + P - E) \quad (\text{C.1.2})$$

La précipitation non interceptée P_{th} (*throughfall* en anglais, en mm) est estimée par :

$$P_{th} = \begin{cases} 0 & \text{si } I \leq I_{max} \\ I - I_{max} & \text{si } I > I_{max} \end{cases} \quad (\text{C.1.3})$$

I_{max} (mm) est un paramètre estimé de manière à minimiser la différence entre les précipitations non interceptées (1) au pas de temps horaire (Équation C.1.3) et (2) au pas de temps journalier. Pour estimer les précipitations non interceptées au pas de temps journalier, on fait une simple comparaison entre l'évapotranspiration et les précipitations journalières (voir Perrin, Michel et Andréassian, 2003). La valeur de I_{max} est choisie parmi l'ensemble des valeurs {0.1, 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 2.25, 2.5, 2.75, 3.0, 4.0, 5.0, 7.5, 10.0, 15.0}.

Le niveau du réservoir est de nouveau mis à jour par :

$$I := I - P_{th} \quad (C.1.4)$$

Arrivant au sol, si E n'est pas complètement satisfaite, une lame d'eau E_s est retranchée du réservoir de production suivant :

$$E_s = \frac{Prod_{GR} \cdot \left(2 - \frac{Prod_{GR}}{X_{1,GR}}\right) \cdot \tanh\left(\frac{E-E_i}{X_{1,GR}}\right)}{1 + \left(1 - \frac{Prod_{GR}}{X_{1,GR}}\right) \cdot \tanh\left(\frac{E-E_i}{X_{1,GR}}\right)} \quad (C.1.5)$$

où $Prod_{GR}$ (mm) est l'état dynamique du réservoir de production et $X_{1,GR}$ (mm) sa capacité.

L'évapotranspiration réelle AE_{GR} (mm) est donc :

$$AE_{GR} = E_i + E_s \quad (C.1.6)$$

Le niveau du réservoir de production est mis à jour par :

$$Prod_{GR} := \max(0, Prod_{GR} - E_s) \quad (C.1.7)$$

La lame d'eau P_s (mm) qui alimente le réservoir de production est estimée par :

$$P_s = \frac{X_{1,GR} \cdot \left(1 - \left(\frac{Prod_{GR}}{X_{1,GR}}\right)^2\right) \cdot \tanh\left(\frac{P_{th}}{X_{1,GR}}\right)}{1 + \frac{Prod_{GR}}{X_1} \cdot \tanh\left(\frac{P_{th}}{X_{1,GR}}\right)} \quad (C.1.8)$$

On peut alors calculer la percolation du réservoir de production $Perc$ (mm) suivant :

$$Perc = Prod_{GR} \cdot \left(1 - \left(1 + \left(\frac{4}{21} \cdot \frac{Prod_{GR}}{X_{1,GR}}\right)^4\right)^{-\frac{1}{4}}\right) \quad (C.1.9)$$

Finalement, le réservoir de production est mis à jour par :

$$Prod_{GR} := Prod_{GR} - Perc \quad (C.1.10)$$

La percolation $Perc$ et $P_{th} - P_s$ constituent les précipitations nettes PR_{GR} :

$$PR_{GR} = Perc + (P_{th} - P_s) \quad (C.1.11)$$

qui est l'entrée de la fonction de transfert.

C.1.2 Fonction de transfert

La fonction de transfert contient deux branches : la branche du débit rapide et la branche du débit lent. $0.1 \cdot PR_{GR}$ passe par la branche du débit rapide alors que $0.9 \cdot PR_{GR}$ passe par la branche du débit lent.

Dans les deux branches, on se sert des hydrogrammes unitaires $UH1$ pour la branche lente et $UH2$ pour la branche rapide. Pour estimer leurs ordonnées, on détermine les courbes cumulées

des temps de transfert $SH1$ et $SH2$, en fonction du paramètre $X_{4,GR}$ (h) selon :

$$SH_1(t) = \begin{cases} 0 & \text{pour } t \leq 0 \\ \left(\frac{t}{X_{4,GR}}\right)^{1.25} & \text{pour } 0 < t < X_{4,GR} \\ 1 & \text{pour } t \geq X_{4,GR} \end{cases} \quad (C.1.12)$$

et

$$SH_2(t) = \begin{cases} 0 & \text{pour } t \leq 0 \\ \frac{1}{2} \left(\frac{t}{X_{4,GR}}\right)^{1.25} & \text{pour } 0 < t \leq X_{4,GR} \\ 1 - \frac{1}{2} \left(2 - \frac{t}{X_{4,GR}}\right)^{1.25} & \text{pour } X_{4,GR} < t < 2 \cdot X_{4,GR} \\ 1 & \text{pour } t \geq 2 \cdot X_{4,GR} \end{cases} \quad (C.1.13)$$

Pour chaque heure h , les ordonnées des hydrogrammes unitaires $UH1$ et $UH2$ sont estimées par :

$$\begin{cases} UH1(h) = SH_1(h) - SH_1(h-1) \\ UH2(h) = SH_2(h) - SH_2(h-1) \end{cases} \quad (C.1.14)$$

Branche lente. Une convolution entre $UH1$ et $0.9 \cdot PR_{GR}$ donne le débit $Q_{9,GR}$. Ce débit alimente le réservoir de routage de capacité $X_{3,GR}$ (mm) et de niveau dynamique $Rout_{GR}$ (mm). Avant d'alimenter le réservoir de routage, nous estimons la lame d'eau potentiellement échangeable avec le réservoir souterrain, notée F (mm), par :

$$F = X_{2,GR} \left(\frac{Rout_{GR}}{X_{3,GR}}\right)^{3.5} \cdot \Delta t \quad (C.1.15)$$

avec $\Delta t = 1$ h, et $X_{2,GR}$ (mm/h) le taux d'échange, positif en cas d'apports d'eau du réservoir souterrain et négatif en cas d'exports d'eau vers le réservoir souterrain.

Le niveau du réservoir de routage est mis à jour par :

$$Rout_{GR} := \max(0, Rout_{GR} + Q_{9,GR} + F) \quad (C.1.16)$$

Ensuite, le débit lent $Q_{r,GR}$ (mm) est calculé suivant :

$$Q_{r,GR} = Rout_{GR} \left(1 - \left(1 + \left(\frac{Rout_{GR}}{X_{3,GR}} \right)^4 \right)^{-\frac{1}{4}} \right) \quad (C.1.17)$$

Le niveau du réservoir de routage est mis à jour par :

$$Rout_{GR} := Rout_{GR} - Q_{r,GR} \quad (C.1.18)$$

Branche rapide. Une convolution entre $UH2$ et $0.1 \cdot PR_{GR}$ donne le débit $Q_{1,GR}$. Ce débit est ensuite comparé avec F pour estimer le débit rapide $Q_{d,GR}$:

$$Q_{d,GR} = \max(0, Q_{1,GR} + F) \quad (C.1.19)$$

Finalement, le débit total $Q_{sim,GR}$ (mm) est la somme de $Q_{r,GR}$ et $Q_{d,GR}$:

$$Q_{sim,GR} = Q_{d,GR} + Q_{r,GR} \quad (C.1.20)$$

La lame d'eau échangée avec le réservoir souterrain $AExch$ (mm) est selon les cas :

$$AExch = \begin{cases} -Rout_{GR} - Q_{9,GR} + F & \text{si } Rout_{GR} + Q_{9,GR} + F \leq 0 \text{ et } Q_{1,GR} + F \geq 0 \\ -Q_{1,GR} + F & \text{si } Rout_{GR} + Q_{9,GR} + F \geq 0 \text{ et } Q_{1,GR} + F \leq 0 \\ -Rout_{GR} - Q_{9,GR} - Q_{1,GR} & \text{si } Rout_{GR} + Q_{9,GR} + F \leq 0 \text{ et } Q_{1,GR} + F \leq 0 \\ 2F & \text{sinon.} \end{cases} \quad (C.1.21)$$

Les pertes totales $LOSS_{GR}$ sont exprimées par :

$$LOSS_{GR} = AE_{GR} + AExch \quad (C.1.22)$$

où AE_{GR} est l'évapotranspiration réelle (Équation C.1.6).

C.2 Équations du modèle IHACRES

La version de la structure du modèle IHACRES (acronyme de *Identification of unit Hydrographs And Component flows from Rainfall, Evaporation and Streamflow data*; Jakeman, Littlewood et Whitehead, 1990) utilisée dans ces travaux est reprise de la réécriture au pas de temps journalier de Perrin (2000, Annexe 1, pp. 376-382) et testée au pas de temps horaire par (Mathevet, 2005, Annexe 1, p. 365).

C.2.1 Fonction de production

D'abord, on détermine un indice d'évapotranspiration ES en fonction de l'évapotranspiration potentielle E et du paramètre produit $X_{1,IHAC} \cdot X_{6,IHAC}$ (mm) :

$$ES = \min\left(a, \frac{E}{X_{1,IHAC} \cdot X_{6,IHAC}}\right) \quad (C.2.1)$$

où $a = 30.0$ est utilisé comme limitation quand $X_{1,IHAC} \cdot X_{6,IHAC}$ est trop faible.

L'indice d'humidité S_{IHAC} (-) au début du pas de temps est mémorisé dans S_0 , puis il est augmenté linéairement grâce aux précipitations et diminué exponentiellement par l'action de l'évapotranspiration selon :

$$\begin{cases} S_0 & := S_{IHAC} \\ S_{IHAC} & := \left(S_0 + \frac{P}{X_{1,IHAC}}\right) \exp(-ES) \end{cases} \quad (C.2.2)$$

L'évapotranspiration réelle AE_{IHAC} (mm) est considérée proportionnelle à la différence entre $S_0 + \frac{P}{X_{1,IHAC}}$ et $\left(S_0 + \frac{P}{X_{1,IHAC}}\right) \exp(-ES)$:

$$AE_{IHAC} \propto \left(S_0 + \frac{P}{X_{1,IHAC}}\right) (1 - \exp(-ES)) \quad (C.2.3)$$

Après la mise à jour de l'indice d'humidité S_{IHAC} (Équation C.2.2), les précipitations nettes PR_{IHAC} (mm) sont déterminées par :

$$PR_{IHAC} = 0.5 \cdot (S_0 + S_{IHAC}) \cdot P \quad (\text{C.2.4})$$

$0.5 \cdot (S_0 + S_{IHAC})$ peut dépasser 1 et cela donne la possibilité au modèle de fermer le bilan en rajoutant une entrée supplémentaire aux précipitations.

C.2.2 Fonction de transfert

PR_{IHAC} est divisée en $X_{2,IHAC} \cdot PR_{IHAC}$ destinée à passer par la branche rapide et $(1 - X_{2,IHAC}) \cdot PR_{IHAC}$ par la branche lente, avec $0 \leq X_{2,IHAC} \leq 1$.

Branche rapide. Le niveau du réservoir linéaire T_{IHAC} (mm) est mis à jour par :

$$T_{IHAC} := T_{IHAC} + X_{2,IHAC} \cdot PR_{IHAC} \quad (\text{C.2.5})$$

Ensuite, le débit rapide $Q_{T,IHAC}$ (mm) est estimé en fonction de T_{IHAC} par :

$$Q_{T,IHAC} = \frac{T_{IHAC}}{X_{3,IHAC}} \quad (\text{C.2.6})$$

avec $X_{3,IHAC} \geq 1$.

Finalement, le niveau du réservoir du routage rapide est de nouveau mis à jour :

$$T_{IHAC} := T_{IHAC} - Q_{T,IHAC} \quad (\text{C.2.7})$$

Branche lente. De manière similaire, le réservoir linéaire du routage lent R_{IHAC} (mm) est mis à jour par :

$$R_{IHAC} := R_{IHAC} + (1 - X_{2,IHAC}) \cdot PR_{IHAC} \quad (\text{C.2.8})$$

Ensuite, le débit lent $Q_{R,IHAC}$ (mm) est estimé en fonction de R_{IHAC} par :

$$Q_{R,IHAC} = \frac{R_{IHAC}}{X_{3,IHAC} \cdot X_{4,IHAC}} \quad (\text{C.2.9})$$

avec $X_{3,IHAC} \cdot X_{4,IHAC} \geq 1$.

Finalement, le niveau du réservoir du routage rapide est de nouveau mis à jour :

$$R_{IHAC} := R_{IHAC} - Q_{R,IHAC} \quad (\text{C.2.10})$$

La somme des débits rapide et lent $Q_{T,IHAC} + Q_{R,IHAC}$ est décalée de $X_{5,IHAC}$ heures pour donner le débit simulé $Q_{sim,IHAC}$ (en mm).

ANNEXE D

UNE MÉTHODE D'ESTIMATION DU TAUX MOYEN DE L'IMPERMÉABILITÉ EFFECTIVE

Geophysical Research Abstracts
Vol. 21, EGU2019-8306, 2019
EGU General Assembly 2019
© Author(s) 2019. CC Attribution 4.0 license.



The "health" of a catchment: How to assess the Effective Impervious Area?

Ludovic Oudin, Mohamed Saadi, and Pierre Ribstein

Sorbonne Université, CNRS, EPHE, UMR METIS, F-75005, Paris, France (ludovic.oudin@upmc.fr)

From an environmental standpoint, a hydrological catchment's health is in inverse proportion with its urbanization stage. As mentioned through different studies, urbanization has generally led to higher peak discharge, higher flood volumes flowing during reduced response time, but also to a poorer water quality. To predict (resp. assess) the impact of a future (resp. previous) urbanization of a catchment, one must diagnose the urban state of the catchment. To this end, the Effective Impervious Area (EIA) or Directly Connected Impervious Area (DCIA) measures were conceived as alternative urban metrics to what hydrologists define as the runoff-rainfall ratio (yield) of the catchment. At the catchment scale, it seems to be a very useful measure, since there may be a spatial combination of pervious and impervious areas that can lead to impact mitigation. Yet, it is still difficult to quantify the EIA of a catchment. An "easy" way consists in using empirical relationships to derive EIA from Total Imperviousness Area (TIA), where TIA is estimated from land cover databases. More rigorous approaches require more datasets, such as hydrologic time series and urban databanks, which are not generally available. Using a large sample of 365 French and American urbanized catchments, for which rainfall and runoff hourly time series were prepared, we propose a pragmatic approach to estimate the EIA using land cover databases, hydrographic network and population density. This approach is validated in an event-based framework, by comparing the resulting EIA values with the slope of runoff-rainfall linearly derived relationships for each catchment.

The "health" of a catchment: how to assess the Effective Impervious Area?

Ludovic Oudin, Mohamed Saadi and Pierre Ribstein

Sorbonne Université, CNRS, EPHE, UMR METIS, F-75005, Paris, France, mohamed.saadi@upmc.fr

Objectives

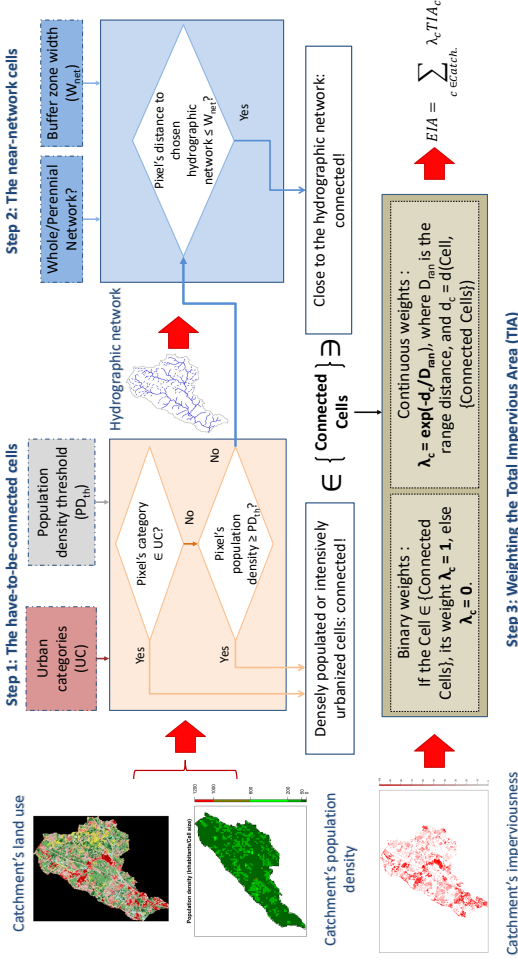
From an environmental standpoint, a hydrological catchment's health is in inverse proportion with its urbanization stage. In order to predict (resp. assess) the impact of a future (resp. previous) urbanization of a catchment, the Effective Impervious Area can be useful, as it represents the part of the catchment that is (1) **impervious** and (2) **connected to the drainage system**. Estimation of EIA can be achieved using several empirical relationships involving Total Impervious Area (TIA), which in turn is derived from spatial datasets. We would like to avoid using such empirical relationships, as they may work for only few catchments, but rather estimate EIA having the basic knowledge about the catchment's land use and hydrography.

A new pragmatic approach

We propose a method of estimating EIA at the catchment scale knowing

- (1) The catchment's land use and imperviousness
- (2) Population density of the catchment and
- (3) The natural hydrographic network.

The first step determines the pixels that are intensively urbanized or with a high population density. The second step selects the cells that are in the vicinity of the hydrographic network. The third step weights catchment's cells (Walsh and Kunapo, 2009; Mroczek et al., 2018) and allows for a catchment-scale estimation of EIA.



Step 3: Weighting the total Impervious Area (TIA)

Inputs

The methodology was applied on a 236-sample of American and French catchments for which TIA ∈ [5%;70%] and the drained areas range from 1.1 km² to 5247 km². For this, the available databases include:

- GHSL Population Density available for 2015 (resolution: 250 m)
- 2011's National Land Cover Data product for the American catchments (resolution: 30 m)
- 2012's CORINE Land Cover product for the French catchments (resolution: 100 m)
- National Hydrography Dataset and BD Carthage.

How efficient is the method?

The computed values of EIA are hydrologically validated against estimated **mean runoff coefficients RC**. First, the events separation was conducted on hourly precipitations-discharge time series through an algorithm proposed by Lobligeois et al. (2014). Then, the RCs were determined using the method proposed by Boyd et al. (1993) and the Ebrahiman et al. (2016) criterion, since it allows **avoiding the combined events in the determination of the RC**, thus focusing only on the EIA-produced runoff.

RCs are also compared to two imperviousness benchmarks:

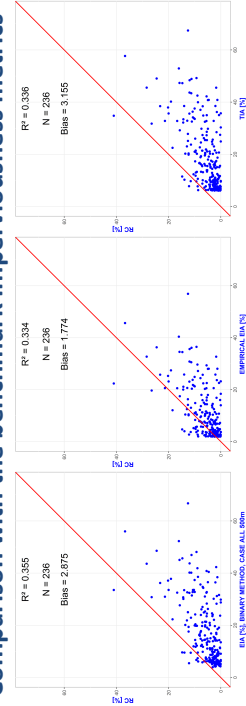
- TIA values
 - An empirical method of estimating EIA (Alley and Veenhuis, 1983): $EIA = 0.15 * TIA^{1.41}$
- The coefficient of determination R² and the Bias (ratio of means) are the considered evaluation criteria.

Optimal setting of the approach

- (1) The urban categories were limited to **developed, medium intensity**
- (2) PD_{th} = 1500 inhabitants/km²
- (3) **Hydrographic network: Whole network**
- (4) For the binary case W_{net} ∈ {50m; 500m; 1500m} and for the continuous case W_{net} = 0m, D_{min} ∈ {10²m; 10³m; 10⁴m}.

	R ² (EIA, RC)	Bias (EIA _{net} /RC _{net})
TIA	0.34	3.16
EMPIRICAL EIA	0.33	1.77
BINARY	W _{net} = 50m	2.37
	W _{net} = 500m	2.88
	W _{net} = 1500m	3.01
CONTINUOUS	D _{min} = 100m	2.78
	D _{min} = 1000m	2.99
	D _{min} = 10000m	3.02

Comparison with the benchmark imperviousness metrics



The best in terms of R²
In terms of R²:
✔ **The best:** R² = 0.36 achieved using the Binary method with W_{net} = 500 m
✘ **The worst:** R² = 0.33 obtained by the Empirical EIA
 In terms of Bias:
✔ **The best:** Bias = 1.77 achieved using the Empirical EIA
✘ **The worst:** Bias = 3.16 obtained by TIA

Conclusions

- The proposed method improves both R² and Bias compared to TIA, though not satisfactorily.
- The biases are larger as the range distance or the buffer zone width gets bigger, i.e. as we converge to TIA.
- The tested empirical EIA improves significantly the Bias but slightly deteriorates R².

Perspectives

The results in terms of bias remain unsatisfying. Thus, we are currently testing cases where:

- The distances (W_{net}, D_{min}) are smaller
- The urban categories are restricted to highly developed areas.

A more refined approach that investigates each pixel's hydraulic course is in development.

References

Walsh and Kunapo (2009). The importance of urban flow paths in determining... Mroczek et al. (2018). Testing the Accuracy of Three Empirical Equations for Determining the... Lobligeois et al. (2014). When does higher spatial resolution rainfall information improve streamflow... Alley and Veenhuis (1983). Effective imperviousness in urban catchments using remote sensing data. J. Hydrol. Soc. 1, 38, 463-478.

ANNEXE **E**

**COMPLÉMENTS AU CHAPITRE 6 :
CLASSEMENT DES 43 MODÈLES TESTÉS
SUR LES 273 BASSINS VERSANTS
URBANISÉS**

Cette annexe détaille 43 modèles testés sur les 273 bassins versants urbanisés, dont 6 structures ont été montrées dans le Chapitre 6. D'abord, nous décrivons les 43 structures de modèle testées. Ensuite, nous montrons le protocole de classification des structures selon leurs performances en calage et en test. Enfin, nous montrons un tableau récapitulatif des rangs de chaque modèle selon les critères de performance adoptés.

E.1 Modèles testés

E.1.1 6 structures conceptuelles non dédiées aux bassins urbanisés

Il s'agit de modèles testés principalement pour les bassins versants non urbanisés, incluant :

1. Le modèle GR4H (**Modèle 01**), que nous avons considéré comme modèle de départ dans les travaux de cette thèse. GR4H constitue la base de la plupart des tests montrés dans la présente annexe. Sa structure est montrée dans les Figures 4.3, 5.3 et 6.3. Elle contient 4 paramètres calés par optimisation numérique. Les équations sont détaillées dans l'Annexe C.
2. Le modèle IHACRES (**Modèle 02**), dont la structure est présentée dans la Figure 4.3 et les équations détaillées dans l'Annexe C. Cette structure contient 6 paramètres optimisés.
3. Le modèle HYMOD (**Modèle 03**), qui contient 5 paramètres optimisés et dont les équations peuvent être retrouvées dans Santos (2018, Annexe B).
4. Une structure hybride dont la fonction de production est celle de IHACRES et la fonction de routage est celle de GR4H (**Modèle 04**). Cette structure contient 5 paramètres optimisés (voir Annexe C) : les paramètres $X_{1,IHAC}$ et $X_{6,IHAC}$ pour la production, et les paramètres $X_{2,GR}$, $X_{3,GR}$ et $X_{4,GR}$ pour le routage.
5. Une structure hybride dont la fonction de production est celle de IHACRES et la fonction de routage est celle de GR4H (**Modèle 05**), avec optimisation du paramètre de partage entre écoulement lent et écoulement rapide X_6 (voir Figure 6.3), fixé initialement à 0.1 (c'est-à-dire 10% des précipitations nettes PR passent par la branche rapide). Rappelons que ce paramètre varie entre 0 et 1. Cette structure contient 6 paramètres optimisés (voir Annexe C) : les paramètres $X_{1,IHAC}$ et $X_{6,IHAC}$ pour la production, et les paramètres $X_{2,GR}$, $X_{3,GR}$, $X_{4,GR}$ et X_6 pour le routage.
6. Une structure hybride dont la fonction de production est celle de GR4H et la fonction de routage est celle de IHACRES (**Modèle 06**). Cette structure contient 5 paramètres

optimisés (voir Annexe C) : le paramètre $X_{1,GR}$ pour la production, et les paramètres $X_{2,IHAC}$, $X_{3,IHAC}$, $X_{4,IHAC}$ et $X_{5,IHAC}$ pour le routage.

E.1.2 6 structures issues de l'optimisation de paramètres fixes et de la suppression de l'échange entre le réservoir souterrain et les branches de routage

La structure de départ GR4H contient certains paramètres fixes qui gèrent interviennent dans la détermination de flux internes du modèle, comme la percolation et le partage entre écoulement lent et écoulement rapide. Nous avons testé le calage numérique de ces paramètres sur les bassins versants urbanisés. En outre, nous avons testé la suppression de l'échange entre le réservoir souterrain et la branche lente ou la branche rapide¹⁷. Ces tentatives incluent :

1. Optimisation du paramètre de partage entre écoulement rapide et écoulement lent X_6 (**Modèle 07, MR5H dans Chapitre 6**). Cette structure contient 5 paramètres à optimiser.
2. Optimisation du paramètre de partage entre écoulement rapide et écoulement lent X_6 , avec suppression de l'échange entre le réservoir souterrain et la branche rapide (**Modèle 08**). Cette structure contient 5 paramètres à optimiser.
3. Optimisation du paramètre de partage entre écoulement rapide et écoulement lent X_6 , avec suppression de l'échange entre le réservoir souterrain et la branche lente (**Modèle 09**). Cette structure contient 5 paramètres à optimiser.
4. Optimisation du paramètre de percolation X_7 (**Modèle 10**), figurant dans l'équation de la percolation suivante (voir Équation C.1.9) :

$$Perc = Prod_{GR} \cdot \left(1 - \left(1 + \left(\frac{4}{21} \cdot \frac{Prod_{GR}}{X_{1,GR}} \right)^{X_7} \right)^{-\frac{1}{X_7}} \right)$$

Ce paramètre est fixé à $X_7 = 4$ dans la structure initiale. En l'optimisant, le nombre de paramètres à caler numériquement devient 5.

5. Optimisation du paramètre de percolation X_7 et du paramètre de partage entre écoulement lent et écoulement rapide X_6 (**Modèle 11**). Cette structure contient 6 paramètres à optimiser.
6. Optimisation du paramètre de percolation X_7 et du paramètre de partage entre écoulement lent et écoulement rapide X_6 , avec suppression d'échange entre la branche rapide et le réservoir souterrain (**Modèle 12**). Cette structure contient 6 paramètres à optimiser.

¹⁷Notons que la suppression de l'échange sur les deux branches en même temps dégrade nettement les performances du modèle.

E.1.3 12 structures avec prise en compte du taux moyen d'imperméabilité totale TIA

L'ensemble de ces modifications tente de prendre en compte dans la structure du modèle le taux moyen d'imperméabilité totale TIA , variant au pas de temps annuel, pour (i) estimer le ruissellement sur la partie imperméable du bassin, ou (ii) prendre en compte la présence de surfaces imperméables ayant des vitesses d'écoulement importantes dans la fonction de routage.

Rappel de notation : X_5 (variant entre 0 et 1) est un paramètre de court-circuitage qui permet d'établir une liaison directe entre deux flux du modèle. Par exemple, dans la Figure 6.3, la liaison est établie entre le flux de précipitations non interceptées P_{th} et les précipitations nettes PR , ce qui est équivalent à court-circuiter le réservoir de production en considérant $X_5 \cdot P_{th}$ comme partie des précipitations nettes PR . X_6 (variant de 0 à 1) est le paramètre de partage entre écoulement rapide et écoulement lent, dont la valeur est fixée à 0.1 dans la structure de départ.

Les 12 tentatives de prise en compte de TIA dans le modèle incluent :

1. Partir de la structure de départ GR4H et faire dépendre le temps de base de l'hydrogramme unitaire $UH2$ sur la branche rapide $2X_{4,GR}$ (Figure 4.3) à la moyenne des valeurs annuelles de TIA sur la période de calage ou de test (**Modèle 13**). Au lieu de $2X_{4,GR}$ heures, le temps de base de l'hydrogramme unitaire $UH2$ devient $2X_{4,GR,rap}$ heures, où :

$$X_{4,GR,rap} = X_{4,GR} \cdot \left(1 - \left(\frac{TIA_{m,period}}{TIA_{max}} \right)^2 \right)$$

avec $TIA_{m,period}$ la moyenne des valeurs de TIA sur la période de calage ou de test, et $TIA_{max} = 0.6$ (valeur maximale de TIA de l'échantillon d'étude). Plus $TIA_{m,period}$ est important, moindre est le temps de base de l'hydrogramme unitaire $UH2$, ce qui permet d'accélérer le routage sur la branche rapide en fonction du niveau d'urbanisation du bassin versant. Cette structure contient 4 paramètres à optimiser. En plus de cette modification, nous avons testé une autre structure en rajoutant une optimisation du paramètre de partage X_6 et une suppression de l'échange entre la branche rapide et le réservoir souterrain (**Modèle 14**). Celle-ci comprend 5 paramètres à optimiser.

2. Fixer $X_5 = TIA$ au pas de temps annuel, en établissant un court-circuitage de P à PR (**Modèle 15**), et de P_{th} à PR (**Modèle 16, MU4H dans Chapitre 6**). Le paramètre X_6 étant fixé à 0.1, ces deux structures contiennent chacune 4 paramètres à optimiser.
3. Fixer $X_5 = TIA$ au pas de temps annuel, en établissant un court-circuitage de P à PR (**Modèle 17**), et de P_{th} à PR (**Modèle 18**). TIA est également prise en compte dans le

paramètre de partage entre écoulement lent et écoulement rapide X_6 , avec $X_6 = X_8 \cdot TIA$, et $X_8 \in [0, 1]$. X_8 est ainsi optimisé au lieu de X_6 . X_6 peut être désormais interprété comme étant le taux d'imperméabilité effective du bassin versant. Ces deux structures contiennent chacune 5 paramètres à optimiser.

4. Fixer $X_5 = TIA$ et optimiser X_6 , où X_6 varie entre 0 et 1. Six schémas de court-circuitage ont été testés : de P à PR (**Modèle 19**) ; de P_{th} à PR (**Modèle 20, MU5H dans Chapitre 6**) ; de P à $X_6 \cdot PR$, c'est-à-dire la branche de routage rapide (**Modèle 21**) ; de P_{th} à $X_6 \cdot PR$ (**Modèle 22**) ; de P à $X_6 \cdot PR$, c'est-à-dire la branche de routage rapide mais cette fois sans échange avec le réservoir souterrain sur cette branche (**Modèle 23**) ; et finalement de P_{th} à $X_6 \cdot PR$, avec suppression de l'échange entre le réservoir souterrain et la branche rapide (**Modèle 24**). Chacune de ces structures contient 5 paramètres à optimiser.

E.1.4 15 structures avec optimisation de X_5 et X_6

Ce paquet de modifications combinent l'optimisation de X_5 avec ou sans optimisation de X_6 , conjuguées à différents schémas de court-circuitage de différentes parties du modèle. Ces 15 structures incluent :

1. Optimiser X_5 (entre 0 et 1) et fixer $X_6 = 0.1$. Six schémas de court-circuitage ont été testés : de P à PR (**Modèle 25**) ; de P_{th} à PR (**Modèle 26, MUOpt5H dans Chapitre 6**) ; de P à $X_6 \cdot PR$, c'est-à-dire la branche de routage rapide (**Modèle 27**) ; de P_{th} à $X_6 \cdot PR$ (**Modèle 28**) ; de P à $X_6 \cdot PR$, c'est-à-dire la branche de routage rapide mais sans échange avec le réservoir souterrain sur cette branche (**Modèle 29**) ; et finalement de P_{th} à $X_6 \cdot PR$, avec suppression de l'échange entre le réservoir souterrain et la branche rapide (**Modèle 30**). Chacune de ces structures contient 5 paramètres à optimiser.
2. Optimiser X_5 (entre 0 et 1) et X_6 (entre 0 et 1). Neuf schémas de court-circuitage ont été testés : de P à PR (**Modèle 31**) ; de P_{th} à PR (**Modèle 32, MU6H dans Chapitre 6**) ; de P à $Q_{sim,GR}$ (**Modèle 33**), ce qui équivaut à convertir immédiatement $X_5 \cdot P$ en débit à l'exutoire, court-circuitant ainsi tout le modèle ; de P_{th} à $Q_{sim,GR}$ (**Modèle 34**) ; de PR à $Q_{sim,GR}$ (**Modèle 35**), ce qui équivaut à court-circuiter toute la fonction de routage ; de P à $X_6 \cdot PR$, c'est-à-dire la branche de routage rapide (**Modèle 36**) ; de P_{th} à $X_6 \cdot PR$ (**Modèle 37**) ; de P à $X_6 \cdot PR$, c'est-à-dire la branche de routage rapide mais sans échange avec le réservoir souterrain sur cette branche (**Modèle 38**) ; et finalement de P_{th} à $X_6 \cdot PR$, avec suppression de l'échange entre le réservoir souterrain et la branche rapide (**Modèle 39**). Chacune de ces structures contient 6 paramètres à optimiser.

E.1.5 4 structures comprenant des formulations alternatives du réservoir de production pour prendre en compte l'intensité de précipitation au pas de temps horaire

Les surfaces imperméables peuvent être conçues comme des surfaces avec un taux d'infiltration très faible par rapport aux surfaces naturelles. À l'échelle du bassin versant, elles peuvent être perçues comme des surfaces où prédomine un ruissellement qui est fonction de l'intensité de précipitation quelque soit la saison. Des tentatives de prise en compte plus explicite de l'intensité de précipitation dans la fonction de production du modèle GR4H ont été donc testées.

La formulation du réservoir de production du modèle GR4H (Figure 3.3) stipule que la lame d'eau instantanée qui l'alimente p_s est fonction du niveau du réservoir $\frac{Prod_{GR}}{X_{1,GR}}$, selon l'équation :

$$\frac{p_s}{p_{th}} = 1 - \left(\frac{Prod_{GR}}{X_{1,GR}} \right)^2 \quad (E.1.1)$$

Peredo (2019, pp. 19-20) a proposé une formulation prenant en compte l'intensité de précipitation, selon :

$$\frac{p_s}{p_{th}} = 1 - \frac{\left(\frac{Prod_{GR}}{X_{1,GR}} \right)^2 + k_{p_0}}{1 + k_{p_0}} \quad (E.1.2)$$

En comparant avec l'Équation E.1.1, le taux d'alimentation du réservoir n'est plus égal à 100% des précipitations non interceptées lors d'une période extrêmement sèche, mais devient de plus en plus faible en fonction du paramètre k_{p_0} dépendant de l'intensité de précipitation seuil p_0 (en mm/h), selon :

$$k_{p_0} = 1 - \exp\left(-\frac{p_{th}}{p_0}\right) \quad (E.1.3)$$

Lorsque p_{th} est très importante par rapport à l'intensité seuil p_0 , $k_{p_0} \sim 1$ et $\frac{p_s}{p_{th}} = 50\%$ au lieu de 100% en période sèche (c'est-à-dire quand $\frac{Prod_{GR}}{X_{1,GR}} \sim 0$).

En intégrant sur un pas de temps de $\Delta t = 1$ h, la lame d'eau P_s qui alimente le réservoir de production devient :

$$P_s = \frac{X_{1,GR} \cdot \left(1 - \left(\frac{Prod_{GR}}{X_{1,GR}}\right)^2\right) \cdot \tanh\left(\frac{P_{th}}{X_1 \cdot (1+k_{p_0})}\right)}{1 + \frac{Prod_{GR}}{X_{1,GR}} \cdot \tanh\left(\frac{P_{th}}{X_1 \cdot (1+k_{p_0})}\right)} \quad (\text{E.1.4})$$

Cette formulation du réservoir de production rajoute un paramètre p_0 , l'intensité de précipitation seuil (en mm/h), qui doit être optimisé numériquement, en plus des paramètres $\{X_{i,GR}\}_{1 \leq i \leq 4}$, ce qui donne une structure à 5 paramètres à caler (**Modèle 40**).

Nous avons testé une autre formulation traduisant un comportement plus *nerveux* du réservoir de production, selon l'équation :

$$\frac{p_s}{p_{th}} = 1 - \frac{\sinh\left(\alpha \frac{Prod_{GR}}{X_{1,GR}} + \beta\right)}{\sinh(\alpha + \beta)} \quad (\text{E.1.5})$$

où α est un paramètre qui module le comportement du réservoir de production, et $\beta = 5.0$. Plus α est petit, plus le ruissellement pendant la saison sèche est important. Avec $\alpha \sim 2.86$, on revient à un comportement similaire au réservoir GR (Figure E.1).

Trois principales configurations ont été testées :

- $\alpha = 2.86$ (**Modèle 41**), ce qui permet de voir la différence entre cette formulation et celle de GR4H. Le nombre de paramètres pour cette structure est de 4 (c'est-à-dire les mêmes paramètres que la structure de départ).
- α change chaque heure en fonction de la lame d'eau brute moyenne horaire P (**Modèle 42**) : pour $P \leq 2.5$ mm, $\alpha = 4.0$; pour 2.5 mm $< P \leq 8.5$ mm, $\alpha = 2.86$; pour 8.5 mm $< P \leq 12.5$ mm, $\alpha = 2.0$; pour 12.5 mm $< P \leq 16.00$ mm, $\alpha = 1.33$; pour 16.0 mm $< P \leq 25.00$ mm, $\alpha = 1.00$; et finalement pour 25.00 mm $< P$, $\alpha = 0.5$. Le nombre de paramètres pour cette structure est 4.
- α est calé selon l'équation (**Modèle 43**) :

$$\alpha = K_{opt} \cdot \exp\left(-\frac{P}{25.00}\right)$$

avec P la lame d'eau brute moyenne horaire (mm), et K_{opt} un paramètre optimisé entre 0.01 et 100. En rajoutant K_{opt} aux quatre paramètres de départ, cette structure contient 5 paramètres à caler numériquement.

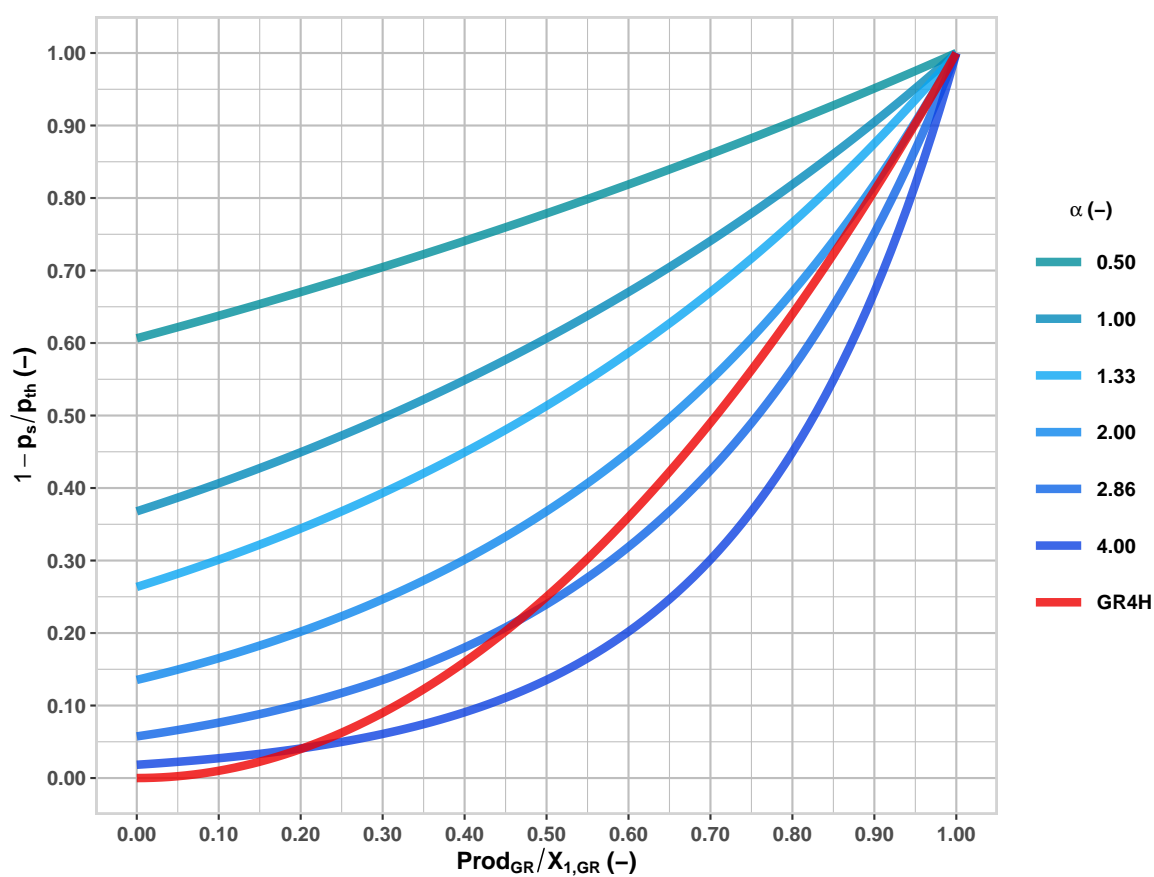


Figure E.1. Variabilité du taux de ruissellement $1 - \frac{p_s}{p_{th}}$ en fonction de l'état du réservoir de production $\frac{Prod_{GR}}{X_{1,GR}}$ pour différentes valeurs de α , avec $\beta = 5$.

Pour calculer P_s , la lame d'eau horaire qui alimente le réservoir de production, on doit intégrer l'Équation E.1.5 sur un pas de temps $\Delta t = 1$ h. Pour cela, posons $K_\alpha = \frac{1}{\sinh(\alpha+\beta)}$. On peut vérifier que la primitive de la fonction $f(x) = \frac{1}{1-K_\alpha \cdot \sinh(\alpha \cdot x + \beta)}$ est

$$F(x) = \frac{2}{\alpha \cdot \sqrt{K_\alpha^2 + 1}} \cdot \operatorname{arctanh} \left(\frac{K_\alpha \cdot \exp(\alpha \cdot x + \beta) - 1}{\sqrt{K_\alpha^2 + 1}} \right) + c \quad (\text{E.1.6})$$

où c est une constante.

L'estimation de P_s , la lame d'eau permettant d'augmenter le niveau du réservoir de production de $Prod_{GR}^h$ (au début du pas de temps) à $Prod_{GR}^{h+1}$ (à la fin du pas de temps), peut être effectuée grâce aux étapes suivantes :

$$\left\{ \begin{array}{ll} \beta & = 5.00 \\ \alpha & \text{imposé ou calé numériquement} \\ K_\alpha & = \frac{1}{\sinh(\alpha+\beta)} \\ c_\alpha & = \sqrt{K_\alpha^2 + 1} \\ Pent & = \min \left(\frac{\alpha \cdot c_\alpha \cdot P_{th}}{2 \cdot X_{1,GR}}, 15.00 \right) \\ FX^h & = K_\alpha \cdot \exp \left(\alpha \cdot \frac{Prod_{GR}^h}{X_{1,GR}} + \beta \right) - 1 \\ Prod_{GR}^{h+1} & = \frac{X_{1,GR}}{\alpha} \cdot \left(\ln \left(\frac{1}{K_\alpha} \left(\frac{c_\alpha \cdot FX^h + c_\alpha^2 \cdot \tanh(Pent)}{c_\alpha + FX^h \cdot \tanh(Pent)} + 1 \right) \right) - \beta \right) \\ P_s & = Prod_{GR}^{h+1} - Prod_{GR}^h \end{array} \right. \quad (\text{E.1.7})$$

E.2 Stratégie de calage et d'évaluation

Pour chaque bassin versant, la période de données hydroclimatiques est divisée en deux sous-périodes avec deux longueurs équivalentes P1 et P2. Chacun des 43 modèles est calé sur P1 et testé sur P2, puis calé sur P2 et testé sur P1. Deux fonctions objectif ont été considérées : le score de Kling-Gupta (Gupta *et al.*, 2009) appliqué sur les débits (KGE) et les racines carrées des valeurs de débit $KGESQ$ (voir l'Équation 6.3.5). Dans cette annexe, nous allons montrer uniquement les résultats pour un calage effectué en considérant $KGESQ$ comme fonction objectif. L'algorithme de calage est constitué de deux phases : (1) une recherche grossière sur un hypercube latin du meilleur point de départ, à partir duquel (2) on applique une optimisation par un algorithme de gradient (Edijatno *et al.*, 1999).

Pour l'évaluation, nous avons considéré 18 critères :

1. 4 critères d'évaluation sur toute la période de test (P2 en cas de calage sur P1, et *vice versa*), à savoir KGE , $KGESQ$, le critère de Nash-Sutcliffe (Nash et Sutcliffe, 1970) sur les débits (NSE) et sur les racines carrées des valeurs de débit ($NSESQ$).
2. Les mêmes 4 critères d'évaluation en focalisant sur la réponse en période sèche : KGE_{dry} , $KGESQ_{dry}$, NSE_{dry} et $NSESQ_{dry}$. La procédure permettant de pondérer les valeurs des chroniques simulées et observées est détaillée dans la Section 6.3.3.
3. Les mêmes 4 critères d'évaluation en focalisant sur la réponse en période humide : KGE_{wet} , $KGESQ_{wet}$, NSE_{wet} et $NSESQ_{wet}$. La procédure permettant de pondérer les valeurs des chroniques simulées et observées est détaillée dans la Section 6.3.3.
4. 3 critères se basant sur les courbes de débits classés observées et simulées. Le premier $ratQL$ focalise sur la simulation les bas débits, pris comme le ratio des quantiles $Q_{20\%}$ (le quantile $Q_{0\%}$ étant la valeur minimale de la distribution) des débits simulés et observés sur la période d'évaluation. Le deuxième $ratQH$ focalise sur la simulation les hauts débits, pris comme le ratio des quantiles $Q_{99\%}$ (le quantile $Q_{100\%}$ étant la valeur maximale de la distribution) des débits simulés et observés sur la période d'évaluation. Le troisième $SlopeBias$ calcule l'erreur dans l'estimation de la pente des débits classés, selon :

$$SlopeBias = \frac{\ln(Q_{0.2}^{sim}) - \ln(Q_{0.7}^{sim})}{\ln(Q_{0.2}^{obs}) - \ln(Q_{0.7}^{obs})} - 1$$

où Q_p^{obs} , Q_p^{sim} sont les débits observé et simulé correspondant à une probabilité de dépassement p .

5. Les 3 critères événementiels e_{Qp} , e_{tp} et VE , dont les équations sont données dans la Section 6.3.3.

Pour chaque critère, on applique la transformée suivante :

$$TCrit = 1 - |Crit_{opt} - Crit| \quad (E.2.1)$$

où $Crit_{opt}$ est la valeur idéale pour chaque critère. Les critères basé sur KGE et NSE , VE , $ratQL$ et $ratQH$ ont une valeur idéale de 1. Les critères e_{Qp} , e_{tp} et $SlopeBias$ ont une valeur idéale de 0.

Ensuite, on détermine pour chaque modèle et pour chaque critère la valeur médiane de $TCrit$ sur les 546 périodes de test en continu et en fréquentiel ($2 \cdot 273$) et les 45025 événements pour

les critères événementiels. Ces médianes sont utilisées pour classer les modèles selon chaque critère. Ainsi, à chacun des modèles est attribué un rang allant de 1 pour le meilleur modèle et 43 pour le pire. Enfin, pour chaque groupe de critères (c'est-à-dire test, test avec focalisation sur la période sèche ou humide, test en fréquentiel ou test en événementiel), la somme des rangs obtenus selon les critères d'évaluation appartenant à ce groupe traduit le rang moyen du modèle selon ce groupe de critères. En fonction de ce rang moyen, un rang final est attribué au modèle.

Exemple : On compare les Modèles 1, 2 et 3 selon les critères fréquentsiels; Modèle 1 a eu un rang de 1 selon *ratQL*, de 3 selon *ratQH* et de 2 selon *SlopeBias* (somme des rangs : 6); Modèle 2 a eu un rang de 2 selon *ratQL*, de 1 selon *ratQH* et de 1 selon *SlopeBias* (somme des rangs : 4); et finalement Modèle 3 a eu un rang de 3 selon *ratQL*, de 2 selon *ratQH* et de 3 selon *SlopeBias* (somme des rangs : 8). La somme des rangs pour les Modèles 1, 2 et 3 est respectivement 6, 4, et 8. Ainsi, Modèle 2 est classé premier, Modèle 1 est classé 2^{ème} et Modèle 3 est classé dernier.

E.3 Classement des 43 modèles

Tableau E.1 montre les rangs de chacun des 43 modèles testés selon les critères de calage, les critères de test continus (toute la période, focalisation sur la saison sèche et focalisation sur la saison humide), les critères de test fréquentsiels, les critères de test événementiels, et finalement le rang moyen en test (c'est-à-dire en prenant en compte tous les critères de test). On remarque l'amélioration des performances du modèle en prenant en compte *TIA* dans l'estimation des paramètres (de la fonction de production ou dans le partage entre écoulement lent et écoulement rapide, voir par exemple le Modèle 18). Les formulations alternatives du réservoir de production pour prendre en compte l'intensité de l'événement pluviométrique ont donné des performances modérées pour la formule proposée par Peredo (2019) et des rangs similaires à ceux de GR4H pour les modèles à formulation hyperbolique de la fonction de production.

Tableau E.1. Classement des 43 modèles testés. Rang en Calage est le rang du modèle selon les performances médianes en calage (critère *KGESQ*). Rang en Test est le rang moyen du modèle à partir des 4 critères de test (*KGE*, *KGESQ*, *NSE* et *NSESQ*). Rang en Test - Sec est le rang moyen du modèle à partir des 4 critères de test en focalisant sur la réponse hydrologique en conditions sèches (*KGESQ_{dry}*, *NSE_{dry}* et *NSESQ_{dry}*). Rang en Test - Humide est le rang moyen du modèle à partir des 4 critères de test en focalisant sur la réponse hydrologique en conditions humides (*KGE_{wet}*, *KGESQ_{wet}*, *NSE_{wet}* et *NSESQ_{wet}*). Rang en Fréquentiel est le rang moyen du modèle à partir des 3 critères fréquentiels (*ratQL*, *ratQH* et *SlopeBias*). Rang en Événementiel est le rang moyen du modèle à partir des 3 critères événementiels (*e_{qp}*, *e_{tp}* et *VE*). Rang Test Moyen est le rang moyen du modèle à partir de tous les critères d'évaluation en test (18 critères). CC signifie court-circuitage.

Modèle	# de paramètres	Modification	Rang en Calage	Rang en Test	Rang en Test - Sec	Rang en Test - Humide	Rang en Fréquentiel	Rang en Événementiel	Rang en Test Moyen
Structures conceptuelles non dédiées aux bassins urbanisés									
Modèle 01	4	GR4H	32	31	40	33	35	40	37
Modèle 02	6	IHACRES	34	40.5	16	42	15.5	8	28
Modèle 03	5	HYMOD	43	43	37	43	43	43	43
Modèle 04	5	Fonction de production de IHACRES et de routage de GR4H	35	38	25	39	11	26	34
Modèle 05	6	Fonction de production de IHACRES et de routage de MR5H (Modèle 07)	33	37	21	38	8	22	29
Modèle 06	5	Fonction de production GR4H et de routage de IHACRES	40	36	39	35	41	30	38

Suite à la page suivante...

Tableau E.1 suite : Classement des 43 modèles testés.

Modèle	# de paramètres	Modification	Rang en Calage	Rang en Test	Rang en Test - Sec	Rang en Test - Humide	Rang en Fréquentiel	Rang en Événementiel	Rang en Test Moyen
Structures issues de l'optimisation de paramètres initialement fixes et de la suppression de l'échange avec le réservoir souterrain									
Modèle 07	5	MR5H, GR4H avec optimisation de X_6	23	26	34	20	25	35.5	26
Modèle 08	5	MR5H sans échange sur la branche rapide	17	17.5	30.5	18.5	34	31	25
Modèle 09	5	MR5H sans échange sur la branche lente	25	27	32	23	31	38	27
Modèle 10	5	GR4H avec optimisation de la percolation	31	30	33	34	17	35.5	31
Modèle 11	6	MR5H avec optimisation de la percolation	15	22.5	23	26	10	29	22
Modèle 12	6	MR5H avec optimisation de la percolation et sans échange sur la branche rapide	12	17.5	22	24	15.5	27	20.5
Structures dont certains paramètres sont contraints par le taux moyen d'imperméabilité totale TIA									
Modèle 13	4	$X_5 = 0$ avec dépendant de TIA	36	35	41	36	40	42	40

Suite à la page suivante...

Tableau E.1 suite : Classement des 43 modèles testés.

Modèle	# de paramètres	Modification	Rang en Calage	Rang en Test	Rang en Test - Sec	Rang en Test - Humide	Rang en Fréquentiel	Rang en Événementiel	Rang en Test Moyen
Modèle 14	5	$X_5 = 0$ avec X_4 dépendant de TIA , X_6 optimisé, et sans échange dans la branche rapide	24	29	38	28	39	39	36
Modèle 15	4	$X_5 = TIA$ et $X_6 = 0.1$, CC de P à PR	29	24	17	27	14	7	17
Modèle 16	4	$X_5 = TIA$ et $X_6 = 0.1$, CC de P_{th} à PR (MU4H)	28	20	19	22	6	16	16
Modèle 17	5	$X_5 = TIA$ et $X_6 = X_8 \cdot TIA$ optimisé, CC de P à PR	26	15	13	17	19	2	13
Modèle 18	5	$X_5 = TIA$ et $X_6 = X_8 \cdot TIA$ optimisé, CC de P_{th} à PR	19	9	9	15	9	9.5	9
Modèle 19	5	$X_5 = TIA$ et X_6 optimisé, CC de P à PR	14	6	3	11.5	13	1	4
Modèle 20	5	$X_5 = TIA$ et X_6 optimisé, CC de P_{th} à PR (MU5H)	8	1	1	2	7	3	2
Modèle 21	5	$X_5 = TIA$ et X_6 optimisé, CC de P à $X_6 \cdot PR$ (branche rapide)	41	40.5	42	40	42	25	41

Suite à la page suivante...

Tableau E.1 suite : Classement des 43 modèles testés.

Modèle	# de paramètres	Modification	Rang en Calage	Rang en Test	Rang en Test - Sec	Rang en Test - Humide	Rang en Fréquentiel	Rang en Événementiel	Rang en Test Moyen
Modèle 22	5	$X_5 = TIA$ et X_6 optimisé, CC de P_{th} à $X_6 \cdot PR$ (branche rapide)	37	32	27.5	29	20	21	32
Modèle 23	5	$X_5 = TIA$ et X_6 optimisé, CC de P à $X_6 \cdot PR$ (branche rapide, sans échange)	42	42	43	41	38	24	42
Modèle 24	5	$X_5 = TIA$ et X_6 optimisé, CC de P_{th} à $X_6 \cdot PR$ (branche rapide, sans échange)	38	33.5	30.5	31	24	20	35
Structures avec optimisation de X_5 et X_6									
Modèle 25	5	X_5 optimisé et $X_6 = 0.1$, CC de P à PR	7	12.5	11	13.5	3	12	8
Modèle 26	5	X_5 optimisé et $X_6 = 0.1$, CC de P_{th} à PR (MUOpt5H)	11	11	10	8	1.5	15	7
Modèle 27	5	X_5 optimisé et $X_6 = 0.1$, CC de P à $X_6 \cdot PR$ (branche rapide)	18	22.5	20	21	33	13	18

Suite à la page suivante...

Tableau E.1 suite : Classement des 43 modèles testés.

Modèle	# de paramètres	Modification	Rang en Calage	Rang en Test	Rang en Test - Sec	Rang en Test - Humide	Rang en Fréquentiel	Rang en Événementiel	Rang en Test Moyen
Modèle 28	5	X_5 optimisé et $X_6 = 0.1$, CC de P_{th} à $X_6 \cdot PR$ (branche rapide)	6	8	7	11.5	30	5.5	10
Modèle 29	5	X_5 optimisé et $X_6 = 0.1$, CC de P à $X_6 \cdot PR$ (branche rapide, sans échange)	13	21	18	25	36	17	19
Modèle 30	5	X_5 optimisé et $X_6 = 0.1$, CC de P_{th} à $X_6 \cdot PR$ (branche rapide, sans échange)	5	12.5	8	13.5	27	11	11
Modèle 31	6	X_5 et X_6 optimisés, CC de P à PR	2	3	4	3	5	5.5	3
Modèle 32	6	X_5 et X_6 optimisés, CC de P_{th} à PR (MU6H)	1	2	5	1	4	4	1
Modèle 33	6	X_5 et X_6 optimisés, CC de P à Q_{sim}	20	16	29	16	28	32	24
Modèle 34	6	X_5 et X_6 optimisés, CC de P_{th} à Q_{sim}	21	14	26	9	22	33	20.5

Suite à la page suivante...

Tableau E.1 suite : Classement des 43 modèles testés.

Modèle	# de paramètres	Modification	Rang en Calage	Rang en Test	Rang en Test - Sec	Rang en Test - Humide	Rang en Fréquentiel	Rang en Événementiel	Rang en Test Moyen
Modèle 35	6	X_5 et X_6 optimisés, CC de PR à Q_{sim}	22	19	27.5	10	23	34	23
Modèle 36	6	X_5 et X_6 optimisés, CC de P à $X_6 \cdot PR$ (branche rapide)	10	10	14	7	21	18	14
Modèle 37	6	X_5 et X_6 optimisés, CC de P_{th} à $X_6 \cdot PR$ (branche rapide)	4	5	6	4	18	14	5.5
Modèle 38	6	X_5 et X_6 optimisés, CC de P à $X_6 \cdot PR$ (branche rapide, sans échange)	9	7	15	6	26	19	12
Modèle 39	6	X_5 et X_6 optimisés, CC de P_{th} à $X_6 \cdot PR$ (branche rapide, sans échange)	3	4	2	5	29	9.5	5.5
Formulations alternatives du réservoir de production									
Modèle 40	5	GR4H avec formulation dépendante de l'intensité de précipitation p_0	16	25	12	18.5	1.5	23	15

Suite à la page suivante...

Tableau E.1 suite : Classement des 43 modèles testés.

Modèle	# de paramètres	Modification	Rang en Calage	Rang en Test	Rang en Test - Sec	Rang en Test - Humide	Rang en Fréquentiel	Rang en Événementiel	Rang en Test Moyen
Modèle 41	4	GR4H avec formulation hyperbolique, $\alpha = 2.86$	30	28	35	30	37	37	33
Modèle 42	4	GR4H avec formulation hyperbolique, α fonction de l'intensité de précipitation	39	39	36	37	32	41	39
Modèle 43	5	GR4H avec formulation hyperbolique, α optimisé	27	33.5	24	32	12	28	30

Fin du Tableau E.1.

