

Thèse de doctorat

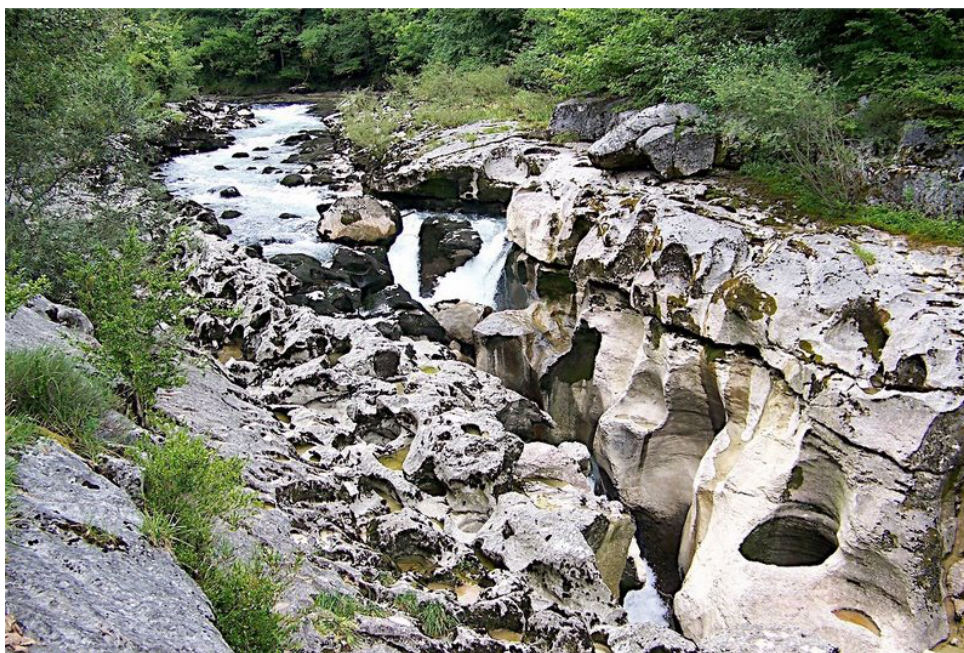
**Le bassin versant de surface vu par le souterrain :
une voie d'amélioration des performances
et du réalisme des modèles pluie-débit ?**

Nicolas Le Moine

Thèse préparée au CEMAGREF (UR HBAN, Antony)

Soutenue le 28 Novembre 2008 devant le jury composé de :

M. Pierre RIBSTEIN	UPMC	Président
M. Christian LEDUC	IRD	Rapporteur
M. Roger MOUSSA	INRA	Rapporteur
M. Ghislain DE MARSILY	UPMC	Examineur
M. Rémy GARÇON	EDF	Examineur
M. Hubert SAVENIJE	TU Delft	Examineur
M. Vazken ANDRÉASSIAN	CEMAGREF	Directeur de thèse
M. Éric PARENT	AgroParisTech	Invité



Bellegarde (01) — Pertes de la Valserine.

“ Toute ma vie est marquée par l’image de ces fleuves, cachés ou perdus aux pieds des montagnes. Comme eux, l’aspect des choses plonge et se joue entre la présence et l’absence. Tout ce que je touche a sa moitié de pierre et sa moitié d’écume.”

Jean TARDIEU (Saint-Germain-de-Joux, 1903 — Créteil, 1995),
Le Fleuve caché.

Remerciements

Une thèse est tout sauf une aventure individuelle : je ne saurais donc mettre un improbable point final à ce manuscrit sans remercier tous ceux qui, de près ou de loin, y auront contribué.

Je tiens tout d'abord à remercier les personnes qui me font l'honneur de juger ce travail : Christian Leduc et Roger Moussa, pour avoir accepté le rôle de rapporteurs du manuscrit, dans des délais restreints ; Hubert Savenije, professeur à l'Université de Delft, et Rémy Garçon, d'EDF-DTG, examinateurs, ainsi que Pierre Ribstein et Ghislain de Marsily, professeurs à l'Université Pierre et Marie Curie, que je remercie également pour avoir contribué, par leur enseignement, à mon enthousiasme hydro(géo)logique !

Mes remerciements vont bien sûr inmanquablement à mon directeur de thèse, Vazken Andréassian, dont la confiance et l'humour m'ont incité à me lancer dans cette aventure ; à Charles Perrin, pour ses conseils avisés, distillés au fil de nombreuses réunions informelles et de la relecture du manuscrit ; à Julien Lerat, pour les échanges passionnés et fructueux sur la semi-distribution des modèles hydrologiques. Je n'oublie pas non plus Claude Michel, parti à la retraite depuis peu mais dont le jugement plein d'expérience et d'humilité reste une valeur sûre et un modèle — dans le meilleur sens du terme. En fait, je crois que je vais remercier l'équipe d'hydrologie *globalement*, précisément pour ce qu'elle est : une *équipe*, dans laquelle j'ai pris un immense plaisir à mener mes recherches. Merci donc à celles et ceux que j'ai cotoyés depuis 2005 : Cécile, Marine, Jean-Louis, Mamoutou, Thib, Marie, Jean-Luc, Audrey, Lionel, Pierre, Maria-Helena, sans oublier tous les stagiaires passés par le bureau 129 entre temps !

En second lieu, j'adresse également mes remerciements aux personnes qui ont rendu possible la constitution de la base de données utilisée pour cette thèse : Laurent Franchisteguy et Jacques Parent du Châtelet, de Météo France, pour avoir mis respectivement à notre disposition l'archive pluviométrique SAFRAN et les données des postes pluviographiques horaires ; Laurent Cobos, du SCHAPI, qui nous a permis d'obtenir un crédit *subprime* auprès de la Banque HYDRO ; Francis Chiew, du CSIRO Land and Water, pour nous avoir gracieusement transmis les données de l'échantillon australien. Bien qu'il ait déjà été inclus dans mes remerciements « globaux », je remercie encore tout spécialement Jean-Louis Rosique, pour son efficacité et sa bonne volonté à toute épreuve dans la gestion des bases de données et la digitalisation des contours de bassins. Je lui souhaite une agréable retraite à venir !

Je remercie également Gilles Bonnet, du Centre de Calcul de Clermont-Ferrand, pour ses conseils dans l'utilisation du cluster. Un merci encore pour Valérie Dansin et Sophie Morin, qui ont assuré dans la bonne humeur le côté administratif de cette thèse.

Une mention un peu spéciale, maintenant. Au cours de ces trois années, j'ai eu l'occasion¹ — et le plaisir je dois dire — de dépoussiérer mes souvenirs de géologie, légués par quelques professeurs passionnés (parmi lesquels mes chers parents ;-) et de constater à quel point ils pouvaient être utiles à un hydrologue... J'ai une pensée particulière pour Daniel Gendre, mon professeur au Collège Jean-Zay de Chinon aujourd'hui retraité, à qui je dédie cette thèse et tout particulièrement le Chapitre 5 sur le karst de la Rochefoucauld. Je me souviens comme si c'était hier de cette descente de la Tardoire en kayak, il y a quinze ans de cela : à l'époque je maudissais chaque seuil de moulin, synonyme de chavirage potentiel en contrebas... Aujourd'hui ce serait plutôt : « *Tiens, une station hydrométrique potentielle!* »

Dans la catégorie « inclassables », je remercie également Jean Margat, hydrogéologue au BRGM et grand jocondoclaste devant l'Éternel, pour nous avoir transmis le « tuyau » initial sur la Charente ainsi que bon nombre de documents précieux, fruits de son impressionnant travail de synthèse sur les aquifères en France et à travers le monde. Un merci également à Stéphane Rénié, d'HydroInvest à Angoulême, pour son accueil et ses explications sur le terrain, ainsi que la journée d'échange avec les acteurs locaux.

Je remercie enfin tous ceux qui ont contribué à remplir ces trois années de bien d'autres choses que de millimètres de pluie virtuels. Je pense aux Hydrologues Sociales et Sociaux ; aux assidus de l'*Ultimate Frisbee* ; à la DreamTeamTM, qui se reconnaîtra (ce qui m'évite par là d'avoir à en dresser une liste « officielle » et, accessoirement, d'oublier des noms), et donc à nos hôtes réguliers du bar-restaurant *chez Mamane*, sis au numéro Cinq de la rue des Vingt-Trois Diamants², sur la Butte aux Cailles...

Je pense, enfin, aux amis de Paris et d'ailleurs ; je pense à ma famille, et ne résiste pas à la tentation d'adresser un clin d'œil à la p'tite Clémence... Me voilà arrivé au bas de ma deuxième page de remerciements, il est vraiment temps de s'arrêter !

¹à la fois dans les livres et sur le terrain, avec l'École Doctorale!

²ou le contraire, je ne sais plus...

Résumé

Les hydrologues dits « de surface » ont longtemps développé des représentations du fonctionnement des bassins versants basées sur une équation de continuité à trois termes : la pluie, l'évapotranspiration et le débit, un pas de temps suffisamment long permettant de négliger les variations de stock. L'hypothèse présente derrière les bilans hydrologiques ainsi calculés est que les infiltrations profondes, ne réapparaissant pas à l'exutoire des bassins topographiques, sont une quantité négligeable.

Dans cette thèse, nous avons cherché à remettre en cause cette hypothèse restée le plus souvent implicite. Les modèles hydrologiques disposent souvent (du fait d'un nombre important de paramètres) d'artifices numériques leur permettant d'adapter le bilan en eau, et les conséquences néfastes de cette hypothèse n'apparaissent que rarement au grand jour : en effet, tout déficit d'écoulement peut être imputé à une autre sortie du système, l'évapotranspiration, tout aussi difficile à estimer précisément. Notre étude partait du constat qu'une prise en compte conceptuelle des échanges surface-souterrain dans un modèle pluie-débit s'avère souvent profitable, quel que soit son pas de temps de fonctionnement (du mensuel à l'horaire). Nous avons donc axé nos recherches dans deux directions complémentaires :

- d'un côté, nous avons cherché à établir des méthodes et des outils pour évaluer le *réalisme* des flux souterrains — ou supposés tels — simulés par les modèles,
- de l'autre, nous avons cherché à modifier la formulation des échanges dans des modèles pluie-débit journaliers et horaires, mais dans le but principal d'améliorer la *capacité des modèles à reproduire le flux mesurable : le débit de surface*.

Ces travaux, basés sur un large échantillon de plus d'un millier de bassins versants français et australiens et menés au travers d'une approche comparative similaire à celle mise en oeuvre par Perrin (2000) et Mathevet (2005), nous ont permis d'aboutir à une structure (commune au pas de temps journalier et horaire) améliorant particulièrement la simulation des étiages. Toutefois, ce processus a également été l'occasion de soulever des questions relatives au choix d'un niveau de complexité.

L'évaluation du réalisme des flux souterrains calculés par un modèle pluie-débit global reste quant à elle question ouverte, de même que la recherche de déterminants physiographiques pour les paramètres correspondants. Des méthodes d'évaluation à la fois directes (dans des cas où une mesure de ces flux à l'échelle du système était possible) et indirectes (par des tests de semi-distribution) ont été mises en oeuvre avec un certain succès ; elles ne sauraient toutefois se substituer à la nécessité de vérifier que les modifications apportées se traduisent bien par amélioration de la capacité à simuler le débit. En effet, cette capacité reste la préoccupation première pour les applications d'hydrologie quantitative.

Abstract

For years, surface hydrologists have been assuming (at least implicitly) a three-term water balance equation as the basis of their models. This equation used to involve rainfall, evapotranspiration, and discharge, the variation in catchment moisture storage being eventually neglected for large time steps. The assumption lying behind this is that deep infiltrations that by-pass the outlet of the topographic catchment are of negligible quantity.

In this thesis, we tried to question this usually overlooked hypothesis. Indeed, hydrological models often have the ability to adjust the water balance due to their numerous degrees of freedom : a deficit in runoff can typically be dealt with by increasing the evapotranspiration rate, which is just as difficult to measure as a groundwater seepage.

Previous studies have shown the potential benefits of an explicit, though conceptual, representation of groundwater–surface water interactions in a rainfall-runoff model. Thus we followed two directions in this study :

- first, we tried to design tools and methods allowing to evaluate the realism of the simulated fluxes, supposedly exchanged between surface and groundwater bodies ;
- second, we tested a large number of formulations for this exchange fonction, but on the sheer point of view of the model's ability to reproduce the measurable flux : surface streamflow.

This work has been carried out on a large dataset (over a thousand catchments in France and Australia) using the same comparative approach as *Perrin (2000)* or *Mathevet (2005)*. Following this empirical approach, we proposed a structure suitable for both daily and hourly time steps and improving the simulation of low flows. However, this process raised important questions about the procedures and criteria used to select models and grade hypotheses.

Evaluating the realism of intercatchment groundwater flows simulated by a lumped rainfall-runoff model is still an open question, as well as the quest for physiographic determinants for the corresponding parameters. Direct evaluation methods (in case measurements of groundwater fluxes at the catchment scale were available) as well as indirect ones (through balance computations over nested catchments) have been developed with some success. However, they cannot be a substitute for the need to verify the improvement of streamflow prediction capabilities : these capabilities remain the main focus of attention for quantitative hydrological applications.

Introduction générale

Contexte de la thèse et problématique scientifique

Il y a une trentaine d'années, l'Américain *Nace* (1973) puis l'Espagnol *Llamas* (1974, 1975) diagnostiquaient un trouble psychique particulier chez les responsables de la planification hydraulique de nombreux pays. Ils forgèrent le terme d'*hydroschizophrénie* pour désigner une « *attitude mentale les conduisant à scinder ou séparer radicalement tout ce qui a trait aux eaux de surface de tout ce qui concerne les eaux souterraines* » (BRGM, 1975). *Llamas* proposa quelques éléments expliquant l'apparition de cette épidémie, due selon lui à la coexistence indépendante, « *d'un côté, d'ingénieurs hydrauliciens, plus ou moins en relation avec les hydrologues des eaux superficielles, de l'autre des hydrologues des eaux souterraines très fréquemment liés au domaine de la géologie universitaire ou des mines* ».

De fait, les hydrologues dits « de surface » ont longtemps développé des représentations du fonctionnement hydrologique d'un bassin versant basées sur une équation de continuité qui stipule que, sur le long terme (à partir du pas de temps annuel), le débit écoulé à l'exutoire d'un bassin versant est égal à la différence entre précipitation et évapotranspiration : $Q = P - E$ (un pas de temps large permettant de négliger les variations de stock). Les pertes (infiltrations profondes ne réapparaissant pas à l'exutoire du bassin) étaient généralement négligées, les bassins versants expérimentaux étant toujours décrits comme « reposant sur un substratum imperméable », et ce n'est que lorsque l'on rencontrait des exceptions criantes, comme dans *Nandakumar et Mein* (1997) par exemple, que l'on acceptait de remettre en cause le mythe du substratum imperméable.

Cette obstination à écarter les pertes fait sourire les hydrogéologues, dont l'objet d'étude, les nappes souterraines, est alimenté par ces pertes. Mais il faut avouer que dans le vocabulaire des hydrogéologues, ce que l'on a tendance à négliger et à appeler perte, c'est tout simplement... l'écoulement de surface. *Beven* (2001) rappelle que l'une des trois hypothèses des modèles hydrologiques actuels est que l'on peut équilibrer le bilan en eau. Mais il remarque qu'il n'existe aucun moyen de s'assurer par des mesures que « *le bassin est en réalité assis sur une couche imperméable. L'équation de continuité, fondamentale en hydrologie, ne peut tout simplement pas être vérifiée* ». Tout le problème vient de l'impossibilité, dans le cas général, de définir un système qui soit fermé à la fois du point de vue « de surface » et du point de vue hydrogéologique, comme nous le verrons plus loin : les lignes de crête topographiques ne correspondent pas nécessairement aux « lignes de crête » de la surface piézométrique (qui d'ailleurs ne sont pas stables au cours de l'année), et de multiples singularités ou discontinuités (fractures, vides de dissolution, etc.) peuvent encore ajouter aux pertes / apports. *Mouelhi et al.* (2006) donnent une preuve par l'absurde de cette non-superposition : si l'on suppose toujours nul le flux souterrain traversant la frontière

topographique lorsque l'on « déplace » virtuellement l'exutoire de l'amont vers l'aval, alors le bassin versant se réduit à une fine couche de matériau imperméable, ce qui est évidemment absurde.

Mais si les pertes souterraines sont si importantes, comment peuvent-elles être négligées par les modèles de bassin versant ? Tout simplement parce que ces pertes (ou apports) peuvent être maquillées sous l'apparence d'une augmentation (ou réduction) d'une autre variable très difficile à quantifier : l'évapotranspiration. Même le taux potentiel est difficile à estimer à l'échelle du bassin versant (*Oudin, 2004 ; Oudin et al., 2005a, 2005b*), et le nombre parfois excessif des paramètres des modèles hydrologiques leur laisse une souplesse suffisante pour adapter le bilan en eau. Cependant, ces adaptations sont souvent peu réalistes, elles engendrent des incertitudes dans la détermination des paramètres, et réduisent fiabilité et robustesse des simulations de débit.

L'objectif de cette thèse est de mettre au point une modélisation du comportement des bassins versants tenant compte explicitement du caractère ouvert du bassin topographique, afin d'aboutir à une représentation réaliste des flux échangés avec les hydrosystèmes souterrains, d'améliorer notre connaissance de la ressource en eau, et de réduire l'incertitude sur l'estimation des paramètres agissant sur les flux de surface. Dans cette brève introduction, nous avons abondamment employé le mot *perte* car il s'agit de loin de la situation la plus courante. Cependant, nous lui préférons désormais les termes plus généraux *d'échanges souterrains entre bassins*¹ — sous-entendu : *topographiques* — pour désigner les écoulements ne suivant pas la logique topographique. Comme nous le verrons au Chapitre 3, le fait qu'ils soient « orthogonaux » à la logique topographique ne signifie pas qu'ils soient orthogonaux à la logique pluie-débit (sinon, cela serait peine perdue de se lancer dans une telle étude...)

Organisation du mémoire

Ce mémoire s'organise en trois parties et onze chapitres.

La première partie est méthodologique (deux chapitres).

Le premier chapitre présente les différentes approches de la modélisation hydrologique, et plus particulièrement celle qui a été choisie pour cette thèse. Par des considérations aussi

¹Le terme communément utilisé dans la littérature est *Intercatchment Groundwater Flows* (IGFs).

bien pratiques que philosophiques nous tenterons de justifier le choix de notre système d'étude (le bassin versant dans son ensemble) ainsi que des méthodes d'évaluation des modèles hydrologiques.

Le deuxième chapitre présente l'échantillon de bassins versants constitué pour ce travail. Le cadre est essentiellement français, mais s'étend aussi à des bassins versants australiens pour le pas de temps journalier. Les différentes sources de données, les méthodes d'agrégation spatiale et de désagrégation temporelle des variables d'intérêt hydrologique (pluie, évapotranspiration potentielle, débit) y seront détaillées.

La deuxième partie (cinq chapitres) présente une approche « vérificationniste » du problème du bilan en eau à l'échelle du bassin versant, c'est-à-dire une approche visant à établir le réalisme quantitatif des flux simulés par un modèle hydrologique.

Dans le troisième chapitre, on établira une typologie des influences souterraines sur les débits de surface. Les différents processus à l'oeuvre seront détaillés, ainsi que leurs conséquences sur le fonctionnement des bassins versants à différentes échelles de temps.

Le quatrième chapitre est un article publié dans *Water Resources Research*. L'intérêt de ce chapitre est double : il présente la bibliographie sur la prise en compte des influences souterraines en modélisation pluie-débit globale, et montre que cette prise en compte est souhaitable au sein d'un modèle hydrologique si l'on veut se rapprocher d'une représentation réaliste des quatre termes du bilan hydrologique : pluie, évapotranspiration, débit et écoulement souterrain.

Le cinquième chapitre, également présenté sous forme d'article, détaille un cas particulier où les différentes composantes, notamment souterraines, du bilan hydrologique d'un groupe de bassins versants ont pu être confrontées à des mesures intégrées à l'échelle du système (karst de la Rochefoucauld, en Charente).

Au sixième chapitre, nous nous intéresserons à la possibilité de « vérifier » les flux souterrains simulés, au travers d'une approche semi-distribuée. Il s'agit d'évaluer, pour un modèle donné, le degré de cohérence entre les bilans simulés sur un ensemble de sous-bassins disjoints et le bilan calculé globalement.

Le septième chapitre est une conclusion à la deuxième partie. Nous y montrons les limites

de l'approche vérificationniste sous-entendue dans la formulation du sujet : les méthodes proposées dans les chapitres quatre à sept se heurtent *in fine* à l'impossibilité de confronter systématiquement les termes souterrains du bilan à des observations à l'échelle du système, et donc à l'impossibilité de se prononcer avec confiance sur leur réalisme. En outre, la pertinence de l'utilisation de la donnée piézométrique, comme entrée ou signal cible d'un modèle, sera discutée. A l'issue de cette partie, nous justifions le retour à une méthode plus rigoureuse de tests systématiques d'hypothèses.

La troisième partie (quatre chapitres) présente les résultats de ces tests systématiques sur les structures de modèles, destinés à améliorer la simulation des débits de surface.

Au huitième chapitre, nous présenterons les résultats obtenus au pas de temps journalier.

Les résultats obtenus au pas de temps horaire font l'objet du neuvième chapitre.

Au dixième chapitre, le problème de l'initialisation des modèles pluie-débit sera abordé. Du fait de la généralisation des modèles spatialisés, la spécification des conditions aux limites spatiales des hydrosystèmes fait l'objet d'une littérature bien plus abondante que ce qui concerne les conditions initiales. Pourtant, une mauvaise spécification de ces dernières peut affecter très sérieusement la qualité des simulations. Ce problème sera relié à celui des influences souterraines par le biais des dépendances de longue durée exhibées sur certains bassins emblématiques, tel que celui de la Somme.

Le onzième chapitre est davantage prospectif. Nous discuterons des problèmes spécifiques liés à la mise en oeuvre de tests sur de grands échantillons, et en particulier des critères d'évaluation et de sélection des structures de modèles. Le point central sera la question de la dépendance entre la complexité d'un modèle hydrologique et son pas de temps de fonctionnement. Les questions posées seront entre autres les suivantes : Pour chaque pas de temps (horaire, journalier, mensuel, etc.), peut-on identifier le nombre optimal de degrés de liberté sur une structure ? Les critères quadratiques sont-ils les plus pertinents pour cette identification ?

Enfin, au dernier chapitre, nous résumerons les principaux acquis de cette thèse, et nous donnerons quelques perspectives qui semblent s'inscrire dans son prolongement.

Table des matières

Remerciements	v
Résumé	ix
Abstract	xi
Introduction générale	xiii
Contexte de la thèse et problématique scientifique	xv
Organisation du mémoire	xvi

Partie I

MÉTHODES DE DÉVELOPPEMENT ET DE COMPARAISON DE MODÈLES PLUIE-DÉBIT

1 Modélisation hydrologique à l'échelle du bassin versant	3
1.1 Quel objet l'hydrologue étudie-t-il?	5
1.2 Le choix du système en modélisation hydrologique	5
1.3 Méthode de développement et d'évaluation des modèles	13
1.4 Conclusion	18
2 Constitution d'un échantillon de bassins versants	21
2.1 Introduction	23
2.2 Construction d'une base de données pluie-débit horaire française	23
2.3 Données étrangères	33
2.4 Résumé des échantillons	34
2.5 Faut-il faire le tri dans les données?	36
2.6 Test préliminaire : sensibilité des performances à l'entrée de pluie horaire	38

Partie II

COMMENT ÉVALUER LE RÉALISME DES FLUX SOUTERRAINS SIMULÉS PAR LES MODÈLES ?

3	Typologie des influences souterraines sur les débits de surface	45
3.1	Le problème de non-superposition des bassins topographiques et hydrogéologiques . . .	47
3.2	Approche descriptive	50
3.3	Approche fonctionnelle	60
3.4	Énoncé définitif des objectifs de la thèse	65
4	Comment les modèles hydrologiques peuvent-ils prendre en compte les flux souterrains inter-bassins ? Etude théorique sur 1040 bassins versants	69
	Résumé en français	71
	Article en anglais	73
5	Confrontation du bilan de surface et souterrain sur le système karstique de La Rochefoucauld-Touvre (Charente)	85
	Résumé en français	87
	Article en anglais	89
6	L'approche semi-distribuée comme test de la cohérence interne des modèles	99
6.1	Principe de l'étude	101
6.2	Le bassin intermédiaire : un système à plusieurs entrées	103
6.3	Choix des combinaisons de stations	103
6.4	Construction des pluies et ETP sur les bassins intermédiaires	107
6.5	Adaptation du modèle, méthodes de calage et de validation	107
6.6	Résultats	109
6.7	Un cas d'étude en semi-distribution	116
6.8	Synthèse	119
7	Conclusion à la deuxième partie	121
7.1	Les limites de l'approche vérificationniste	123
7.2	Peut-on utiliser la donnée piézométrique pour contraindre les simulations ? Un premier calcul de coin de table	123
7.3	Quelles sont les alternatives basées sur la mesure ?	127
7.4	N'est-il pas absurde de chercher à modéliser des influences <i>extérieures</i> au bassin topographique, sans modéliser cet extérieur ? Un second calcul de coin de table	128

Partie III**TESTS SYSTÉMATIQUES DE STRUCTURES
À DIFFÉRENTS PAS DE TEMPS**

8 Développements au pas de temps journalier	133
8.1 Point de départ	135
8.2 Canevas des modifications testées	136
8.3 Résultats	141
8.4 Vers une plus grande complexité?	152
8.5 Résultats complémentaires sur les bilans simulés	157
8.6 Un test extrême de robustesse	160
8.7 Synthèse	166
9 Développements au pas de temps horaire	169
9.1 Point de départ	171
9.2 Test au pas de temps horaire des structures identifiées au pas de temps journalier . . .	172
9.3 Recherche d'améliorations spécifiques au pas de temps horaire	182
9.4 Résultats détaillés sur quelques bassins versants	189
9.5 Synthèse	200
10 Stratégies d'initialisation des simulations pluie-débit	201
10.1 Problématique	203
10.2 Impact d'une mauvaise spécification des conditions initiales sur les performances des modèles pluie-débit	204
10.3 Obtention du niveau d'équilibre du réservoir sol par une analyse de son comportement asymptotique	209
10.4 Proposition et test d'une procédure d'initialisation alternative	215
10.5 Synthèse	217
11 Perspectives : Quel(s) critère(s) utiliser pour sélectionner un modèle?	219
11.1 Les stratégies de sélection d'un modèle et leurs limites	221
11.2 Tests empiriques de vraisemblances formelles	233
11.3 Synthèse	246

12 Conclusion générale	247
Références bibliographiques	253
 ANNEXES	
A Table de correspondance Critère de Nash – Critère C_{MM}	267
B Estimation de la significativité des écarts entre structures à l'aide d'un test non-paramétrique	269
B.1 Problématique	269
B.2 Description du test	270
B.3 Résultats	270
B.4 Synthèse	272
C Dépendance des paramètres de capacités au pas de temps de fonctionnement	275
C.1 Loi de vidange instantanée	275
C.2 Cas où $\alpha > 1$	275
C.3 Cas où $\alpha < 1$	277
C.4 Cas où $\alpha = 1$	277
D Méthode alternative d'identification de l'hydrogramme unitaire horaire	279
D.1 Objectifs	279
D.2 Principe (d'après <i>Croke, 2005</i>)	279
D.3 Adaptation à l'hydrogramme unitaire du modèle horaire	280
D.4 Synthèse	288
E Transformation en quantiles normaux	289
E.1 Origine et intérêt pour la transformation des débits	289
E.2 Mise en oeuvre	292
E.3 Modélisation des queues de distribution	293
F Drainage density and streamflow : a new look at an old debate	295
F.1 Introduction	297
F.2 The three flaws of the drainage density metric	298
F.3 Automatic computation of Effective Drainage Density	300
F.4 Confronting streamflow attributes to drainage density on a large catchment set	304
F.5 Detailed analysis of a case study in Normandy	306
F.6 Limits and perspectives of the use of drainage density in hydrology	307

Première partie

Méthodes de développement et de comparaison de modèles pluie-débit

1

Modélisation hydrologique à l'échelle du bassin versant

1.1 Quel objet l'hydrologue étudie-t-il ?

L'hydrologie¹ est, au pied de la lettre, la *science de l'eau*. Inévitablement, un programme aussi ambitieux devait se heurter à la réalité : il y a de l'eau à peu près partout, de l'atmosphère à l'océan en passant par le sol, le sous-sol, les rivières, les glaciers, et bien sûr les êtres vivants ; sa quantité autant que sa qualité dépendent de processus physiques, chimiques et biologiques innombrables. Par conséquent, il a fallu restreindre un peu cette définition étymologique.

On peut définir l'hydrologie moderne comme l'étude de la partie continentale du cycle de l'eau, d'un point de vue essentiellement quantitatif. Cela ne restreint que très partiellement le terrain de jeu des hydrologues, puisque ce cycle recouvre encore des phénomènes météorologiques, des écoulements à surface libre, en milieu poreux, des transferts de masse et d'énergie dans la biosphère, etc.

Comme le souligne *Klemeš* (1986), les hydrologues forment finalement une sorte de « famille recomposée », une communauté de spécialistes venant d'horizons très différents : météorologues, hydrauliciens, géologues et hydrogéologues, pédologues, agronomes ... ayant chacun leur vision du cycle de l'eau, généralement marquée par leur famille d'origine.

1.2 Le choix du système en modélisation hydrologique

1.2.1 Approches réductionniste, systémique et le problème des conditions aux limites

En pointant du doigt le « dilettantisme » des hydrologues, *Klemeš* constate une situation héritée de l'histoire... qui n'a rien de condamnable en soi ! Cependant, elle conduit rapidement à la vision du paysage présentée à la Figure 1.1 : un puzzle d'entités censé rendre compte exhaustivement de toute l'hétérogénéité existante, dont nous avons bien du mal à déterminer la part réellement pertinente à intégrer dans un modèle hydrologique.

En effet, l'hydrologie a ceci de particulier que son objet ne tient pas sur une paille, et que des expériences sur le système réel en faisant varier un seul facteur à la fois, *toutes choses égales par ailleurs*, ne sont tout simplement pas possibles. Le recours à des *modèles* est alors inévitable, et les développements de l'informatique au cours des dernières décennies ont

¹dont l'étymologie doit s'écrire ὕδωρ λόγος, précédée d'un esprit dur, afin de rappeler son appartenance aux sciences du même nom [...]

largement contribué à leur généralisation (*Kauark-Leite et Nascimento, 1993*).

Dans l'approche décrite précédemment, le modélisateur est confronté à la nécessité, pour pouvoir comprendre ou prévoir l'évolution de certaines grandeurs d'intérêt, de spécifier des relations entre chacune des petites entités et ses voisines immédiates : c'est à cela que s'emploient les tenants d'une vision strictement *mécaniste*, ou *réductionniste*. À la rigueur, comme pour un puzzle, il ne semble pas nécessaire de savoir ce que représente l'image globale : pour construire l'ensemble, il paraît suffisant de considérer les pièces une par une et de les relier de façon à ce qu'elles s'emboîtent, sans rupture dans la continuité du motif. Cependant, comme on le voit dès maintenant, cette approche nécessite un grand nombre d'hypothèses concernant les différentes pièces du puzzle et leurs relations : la manière dont les éléments intègrent les flux venant d'un élément voisin, la manière dont leurs variables d'état interagissent, etc. (*Michel et al., 2006*).

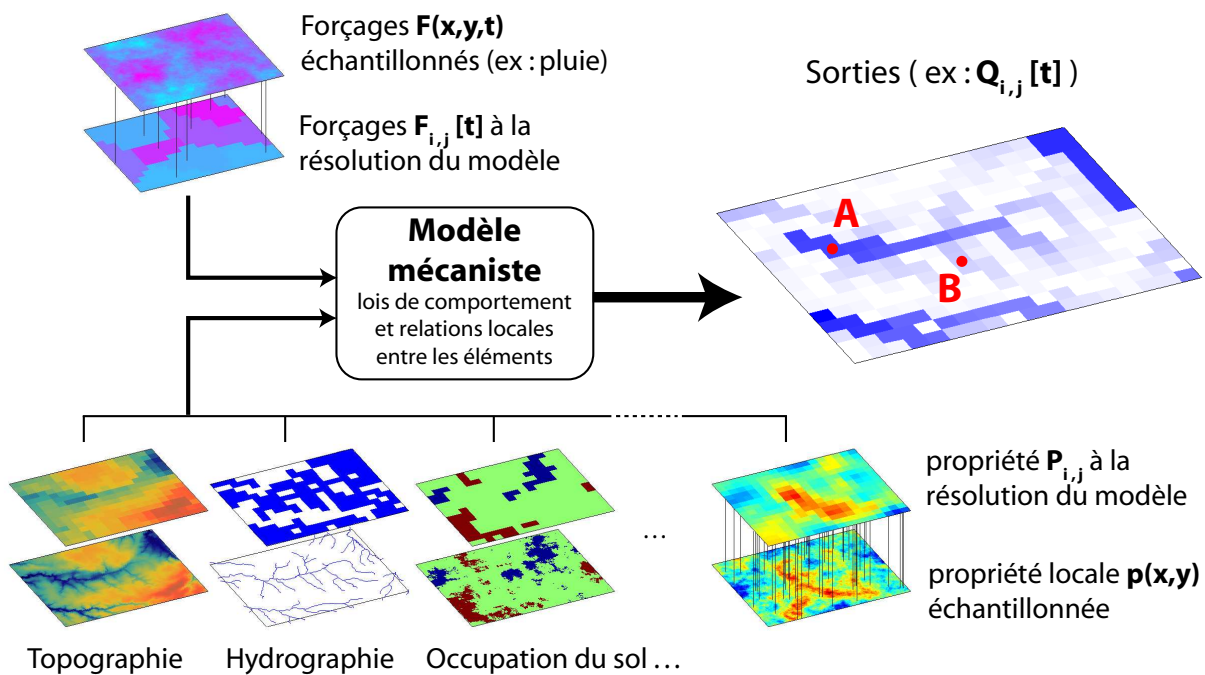


FIG. 1.1: Principe de l'approche mécaniste (ou réductionniste) en hydrologie. Cette approche repose sur la prescription *a priori*, dans un modèle maillé (différences ou éléments finis), de lois de comportement et de relations à l'échelle locale. Elle nécessite l'estimation de propriétés locales (perméabilité, épaisseur de sol, coefficient de Strickler, etc.) et de forçages (pluie, ETP, etc.) qui n'ont cependant plus nécessairement la même signification à la résolution de la maille du modèle. Elle permet théoriquement, à la résolution près, d'avoir accès aux variables de sortie en tout point du domaine décrit.

En cas de succès, nous serions en mesure de prédire l'évolution d'une grandeur donnée au cours du temps (humidité du sol, écoulement produit, etc.) en n'importe quel endroit du domaine décrit. Dans la suite, nous considérons que l'estimation qui nous intéresse est celle

du débit en un point A, ce « point » étant une section de cours d'eau.

Il y a une autre façon d'appréhender le problème de la simulation du débit en A. En effet, nous pouvons considérer qu'il existe une entité fonctionnelle (un volume de contrôle, pour parler comme les thermodynamiciens) privilégiée pour ce calcul : le bassin versant à l'amont de A, que nous pouvons définir comme l'ensemble des points de la surface topographique qui convergent vers la section A en suivant une ligne de plus grande pente (trajectoire orthogonale aux lignes de niveau)¹.

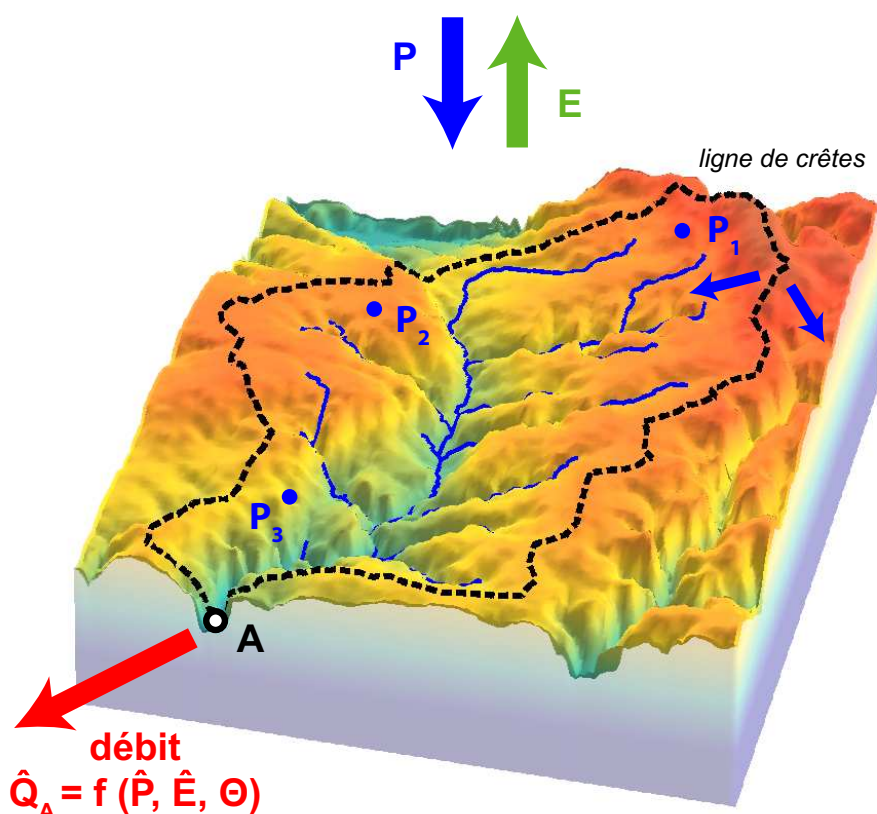


FIG. 1.2: Définition du système bassin versant. Le forçage principal de ce système est la pluie de bassin \hat{P} , grandeur qui existe objectivement : c'est le rapport entre le volume total de pluie qui s'est abattu sur le bassin au cours d'un pas de temps (journée, heure, ...), et la surface du bassin. En pratique, cette pluie de bassin est estimée par une moyenne pondérée des observations P_1, P_2, \dots, P_n aux postes pluviométriques couvrant le bassin.

Dans cette approche *systémique*, ou *globale*, le choix d'un volume de contrôle unique simplifie grandement le problème des conditions aux limites et des relations qu'il nous faut imposer, comme le montre la Figure 1.2 :

- Du point de vue des écoulements, la frontière de ce volume de contrôle est, en première approximation², à condition de flux nul,

¹Nous avons essayé de ne pas utiliser le mot « eau » dans cette définition.

²C'est précisément de cette approximation qu'il faudra sortir !

- Le débit mesuré à la station de jaugeage en A (s'il en existe une, bien sûr!) fournit une mesure de l'écoulement produit à l'échelle du système : ce type de mesure est généralement impossible si l'on prend comme volume de contrôle l'un des éléments du maillage précédent,
- Dans tous les cas, il nous faut estimer les forçages climatiques qui sont les entrées à l'échelle de la maille choisie : la pluie et l'évapotranspiration potentielle. Cependant, cette estimation sera nettement moins incertaine dans le cas global car le champ de pluie peut être échantillonné en plusieurs points du bassin versant, alors qu'on aura plus du mal (c'est un euphémisme) à avoir un poste pluviométrique par élément d'un modèle maillé. Pour prendre une analogie, il y a moins d'incertitude dans l'estimation de la moyenne du lancé simultané de n dés, que dans l'estimation du résultat du lancé d'un seul dé.

L'hydrologue « global » se reconnaît donc à ceci qu'il refuse de considérer tout autre système qu'un bassin (ou sous-bassin) versant. L'approche globale va alors consister à établir des relations entre les entrées, les états et les sorties de ce nouveau système. Étant donnée son échelle, il est clair que la relation fonctionnelle $\hat{Q} = f(\hat{P}, \hat{E}, \Theta)$ décrivant la transformation pluie-débit doit être établie *empiriquement*. Les liens entre les paramètres Θ de la relation d'un côté, et la réalité physique (à l'échelle de la parcelle) de l'autre, ne peuvent pas être définis *a priori* : ils restent implicites, à découvrir *a posteriori* (Michel et al., 2006).

Notons toutefois que ce problème du choix du système est distinct de celui de l'échelle. Certains modèles mécanistes, tels que les GCM (*Global Climate Models*), utilisent des mailles de plusieurs centaines de kilomètres carrés pour l'hydrologie : la plupart des bassins versants de l'échantillon décrit au chapitre suivant sont bien plus petits que cela. L'approche globale n'exclut pas non plus une forme de distribution (on parlera alors d'approche semi-globale, ou semi-distribuée) : simplement, le maillage sera construit sur une base topographique de telle sorte que chaque maille soit un sous-bassin versant, décrit par un jeu de paramètres non-réductionniste (le chapitre 6 présente une étude de ce type). Nous pouvons résumer la philosophie de l'hydrologue global par cette devise : *Il n'y a qu'une maille qui m'aille, le bassin versant.*

1.2.2 Discussion des principes

1.2.2.1 L'estimation des paramètres, un problème spécifique à l'approche descendante ?

Parmi les objections faites à l'approche globale, il en est une qui n'est pas négligeable : que faire lorsque l'on a vraiment besoin d'une information spatialisée¹, par exemple pour des

¹Ou que l'on n'en a pas besoin, mais que l'on est victime du « *Qui peut le plus peut le moins* »...

problèmes de qualité des eaux ? En effet, par construction du système (entrées-états-sorties), on ne peut pas faire de prédiction pour un point *intérieur* à celui-ci.

Supposons que nous ayons défini et étudié le fonctionnement du bassin versant en A, et que nous installions une nouvelle station de mesure (de débit, de concentration en nitrates, etc.) en un point B quelque part à l'amont. Dans une approche globale, il n'y a pas d'autre solution que de définir un nouveau système en B, et de recommencer l'étude empirique en ce point : au contraire, le modèle mécaniste donnait d'emblée accès au débit en A, B ou n'importe quel autre point, à la résolution près.

Si nous n'avons pas de station de mesure en B mais que nous voulons *dire quelque chose* au sujet du débit ou de la concentration en B, nous devons procéder de la même façon (isoler le bassin versant en B) mais, cette fois, ne devons estimer les paramètres de ce nouveau système « à l'aveugle » : c'est la problématique des bassins non-jaugés. Le problème est loin d'être simple, mais nous pouvons imaginer transposer en B les paramètres du bassin versant A, ou utiliser les relations statistiques établies *a posteriori* entre les paramètres d'autres bassins et certains descripteurs physiques, etc. Ceci paraît choquant, mais cette spécification *a priori* de paramètres probables pour des lois de comportement, est pourtant exactement ce que l'on a dû faire dans la construction du modèle mécaniste. Comment a-t-on affecté telle valeur de perméabilité à telle ou telle formation ? On peut avoir effectué un essai de pompage (i.e., avoir sollicité *in situ* le système pour étudier sa réponse et en déterminer la loi), ou avoir prélevé un échantillon, que l'on a testé dans un perméamètre (méthode *ex situ*), ce qui s'apparente dans un cas comme dans l'autre à un calage d'après des observations. On peut encore avoir choisi une valeur dans une bibliothèque, selon une typologie (taille des grains, etc.), ce qui s'apparente à la régionalisation des paramètres des modèles hydrologiques. Pourquoi les deux situations ne sont-elles pas perçues de la même façon ?

Fondamentalement, les modèles réductionnistes semblent plus satisfaisants pour l'esprit car ils font appel à des notions apparemment plus « objectives » que les modèles globaux : la « perméabilité » est un paramètre que nous rangeons inconsciemment dans la même case que « longueur » ou « masse » : nous l'identifions à une *propriété* du cylindre de sédiment, en comparaison de laquelle le « temps de base de l'hydrogramme unitaire » apparaît comme un infâme bricolage dépourvu de réalité objective¹. Pourtant, la physique ne manque pas d'exemples où des « lois » phénoménologiques macroscopiques (et les « propriétés » correspondantes) émergent statistiquement d'un niveau inférieur : le modèle du gaz parfait, par exemple, peut être retrouvé dans le cadre de la théorie cinétique des gaz. La plupart des lois classiques de transfert, de type « flux-gradient de potentiel » (Fick, Fourier, Ohm,

¹Ce qu'il est effectivement.

Darcy, etc.) sont elles aussi purement phénoménologiques et ne sont en rien irréductibles. Matheron (1984) a d'ailleurs proposé un modèle d'émergence de la loi de Darcy — loi dont la forme est très simple — dans un milieu poreux macroscopiquement homogène, à partir des lois d'écoulement dans les pores (équations de Navier-Stokes au niveau granulométrique, hétérogène). Mais il conclut sa démonstration par cette phrase : « *En contrepartie, [ce mode de raisonnement] nous dit seulement qu'il existe une relation linéaire entre flux et gradient macroscopiques, mais ne nous permet pas de calculer explicitement le tenseur K des perméabilités* »... Autrement dit, bien malin qui pourra déterminer la perméabilité macroscopique de notre cylindre de sédiment à l'aide d'un microscope ! Un hydrologue qui, suivant l'exemple du Docteur Faustroll (Jarry, 1898), se serait réduit, comme paradigme de petitesse, à la taille d'un pore, considérerait cette perméabilité comme un concept opératoire certes très commode pour calculer des flux à des échelles très grandes devant sa propre taille, mais dépourvu de toute réalité objective et très vaguement fiable aux « propriétés » de ce qui l'entoure ! En translatant les échelles et en remplaçant, respectivement, *échelle des pores* par *échelle de la parcelle* et *échelle de l'échantillon de sédiment* par *échelle du bassin versant*, nous nous ramenons à notre situation hydrologique : rechercher des lois de comportement à l'échelle « gigascopique » (i.e., l'échelle du bassin versant) et n'en rechercher le fondement macroscopique qu'*a posteriori* n'est pas moins scientifique que l'approche mécaniste, sauf à considérer que la loi de Darcy est moins scientifique que les équations de Navier-Stokes.

Si les caractéristiques du bassin versant (occupation du sol, etc.) changent, les paramètres du modèle global cessent d'être une représentation acceptable de son comportement, conformément au principe *d'unité d'action* énoncé par De Marsily (1994)... dans la même mesure que le paramètre de perméabilité d'un sol ou d'une roche cesse de représenter acceptablement son comportement si la formation est remaniée, soumise à compression ou dessiccation, gel et dégel, qui altèrent sa structure microscopique¹ ! En fait, pour reprendre la terminologie de De Marsily, nous faisons le choix délibéré de rechercher des *relations* entre différentes variables observées à l'échelle d'intérêt, plutôt que de chercher à identifier des *propriétés intrinsèques* des objets étudiés².

Ainsi, nous estimons que le principe des trois unités (temps, lieu, action) est exact, mais que s'interdire d'utiliser les modèles en extrapolation sous prétexte que « *pour un jeu de données d'entrée différent, dans des plages de valeurs différentes de celles utilisées pour l'apprentissage, les prévisions du modèle puissent être mises en défaut* » est contre-productif. Cette certitude

¹Dès lors, l'utilisation de certains modèles hydrogéologiques pour dimensionner et prévoir l'évolution de stockages de déchets nucléaires sur des millénaires devrait elle aussi être soumise, en toute justice, à des « *peines de réclusion à perpétuité* » !

²Outre le fait que c'est une régression à l'infini, on peut philosopher longtemps pour décider si, à la limite, de telles propriétés intrinsèques existent vraiment (voir par exemple Gröblacher et al., 2007)

de mise en défaut vaut pour n'importe quel type de modèle, et nous soutiendrons plus loin qu'elle doit même être recherchée activement. L'unité d'action est une contrainte indépassable : les modèles capables de prédire tous les phénomènes potentiellement d'intérêt — Matheron (1978) les qualifie de *panscopiques*¹ — sont une pure vue de l'esprit, à supposer même que nous connaissions quels sont *tous* ces phénomènes. Dans la pratique courante, les modèles utilisés sont au mieux *polyscopiques*, voire *monoscopiques* (c'est le cas de la plupart des modèles pluie-débit). Pour autant, l'utilisation d'un modèle monoscopique n'est pas incompatible avec la recherche d'une certaine objectivité, même si elle n'est pas *interne* au modèle. Comme le résume Matheron, « à strictement parler, il n'y a pas deux expériences identiques : elles diffèrent toujours l'une de l'autre par quelques facteurs accessoires (mais c'est nous qui les jugeons tels), et par les conditions de lieu ou de temps. Tout ce que l'on peut dire, c'est que les facteurs qui nous paraissent importants ont été rendus aussi semblables que nous le permettent nos moyens techniques, et les autres sont ce qu'ils sont » (p. 56). C'est finalement la sanction de la pratique qui fonde un critère *externe* d'objectivité, qui revient « à examiner si l'on a, ou non, "raison en moyenne" d'utiliser telle méthodologie pour tenter de résoudre telle catégorie de problème » (p. 57). Dans le même esprit, Beven (2001) dénonce la recherche de la *profondeur explicative* d'un modèle au détriment de son *pouvoir prédictif* (sur quelques grandeurs observables) : en effet, la première est une pure affaire de croyance, quand le second est testable. Un modèle, n'importe quel modèle, est par définition un outil d'extrapolation, une hypothèse constamment sur la sellette, une « *anticipation, téméraire et prématurée* » : le modélisateur est fondamentalement une sorte de pataphysicien, qui identifie — très temporairement et sans trop y croire, s'il est sérieux! — *quelques-unes* des propriétés du système réel avec celles de son modèle, ce dernier restant ce qu'il est : une solution imaginaire.

1.2.2.2 Limites pratiques à la spatialisation

À ce stade, il faut insister sur ce fait : il n'y a aucune supériorité de principe d'une approche sur l'autre à revendiquer. Il serait encore plus absurde, au nom d'une espèce de métaphysique « holiste », de nier que la réalité *soit* distribuée, que les processus et les propriétés des milieux soient effectivement hautement spatialisés, et que cela ait un effet sur le fameux « débit observé en A ».

Cependant, il y a une objection importante à l'usage généralisé de l'approche réductionniste pure et dure dans les sciences de la Nature, objection qui est de nature pratique : on n'en a tout simplement jamais fini de prendre en compte l'hétérogénéité de l'environnement. Les relations entre entités ont beau être fondées sur des lois physiques — avec toute l'ambiguïté

¹Du grec πᾶν, *tout*, et σκοπεῖν, *observer, viser*.

de ce mot, mentionnée plus haut —, ces lois ne valent que pour des milieux supposés homogènes. Dès lors, il n'y a que deux possibilités : soit on s'arrête à un certain point, et l'on accepte que le modèle soit global et conceptuel à l'échelle de la maille élémentaire, soit on s'engage dans le raffinement perpétuel de la description des propriétés et des relations au sein du système.

Pour formuler les choses de façon un peu polémique, le malheur est que la croissance continue des systèmes d'acquisition de données (télétection, RADAR, etc.) d'un côté, et de la puissance de calcul accessible de l'autre, donne l'illusion que ce projet est viable. Pour expliquer ce qu'il faut bien appeler un cercle vicieux, l'analyse la plus pertinente est celle d'un chercheur que l'on ne saurait soupçonner d'obscurantisme ou de haine des ordinateurs : Joseph Weizenbaum, professeur émérite d'informatique au M.I.T. et auteur du livre *Computer Power and Human Reason : from Judgement to Calculation* (Weizenbaum, 1976). Au chapitre 1, intitulé *On Tools*, Weizenbaum explique comment le développement de modèles de plus en plus complexes devient progressivement une fin en soi, un processus autonome :

“If a particular technique requires an enormous amount of computation and if only a limited computational effort can be devoted to it, then a failure of the technique can easily be explained away on the ground that, because of computational limitations, it was never really tested. The technique itself is immunized against critical examination by such evasions.”

Et Weizenbaum de décrire les errements dans lesquels tout modélisateur est tombé un jour ou l'autre :

“When the limitation seems to him to be entirely computational, and when a computer is offered to help remove it, he may well launch a program of intensive, time-consuming “research” aimed simply at “computerizing” his technique. Such programs usually tend, by virtue of their magnitude, to increasingly dominate the task and, unless great care is taken to avoid it, to eventually become the center of attention. As ever more investment is made in attacking these initially ancillary subproblems, and as progress is made in cracking them, an illusion tends to grow that real work is being done on the main problem. The poverty of the technique, if it is indeed impotent to deal with its presumed subject matter, is thus hidden behind a mountain of effort, much of which may well be successful in its own terms”

Ces deux paragraphes pourraient tout à fait s'appliquer à la modélisation hydrologique. Attacher de l'importance à « l'esthétique » d'un modèle n'est pas condamnable en soi,

mais à condition que cela ne fasse pas oublier la question centrale : progresse-t-on dans la prédiction du débit au point A? La réponse n'est pas évidente. Comme le soulignent également *Kauark-Leite et Nascimento* (1993), il ne s'agit pas de « *freiner le développement scientifique en interdisant toute complexité invérifiable à un moment donné* », mais de manier cette complexité « *avec prudence pour ne pas renvoyer sine die le jugement d'un modèle par l'expérimentation* ». Il s'agit d'éviter qu'un modèle reste le compendium perpétuellement inachevé de tous les processus hydrologiques possibles (*Michel et al.*, 2006), la Carte de l'Empire à l'échelle $\frac{1}{1}$ du conte de Borgès¹. Le réductionniste intransigeant risque de se heurter en pratique à un problème qu'il crée par son propre vocabulaire : on ne tombe jamais sur un élément irréductible (« *le réel est inépuisable* », comme l'écrit Matheron)!

Ce point nous permet d'introduire le paragraphe suivant, qui concerne la méthode de développement et d'évaluation des modèles.

1.3 Méthode de développement et d'évaluation des modèles

1.3.1 Règles du jeu

Le choix d'une méthode de travail fait nécessairement l'objet de conventions. Celle que nous présentons ici n'est rien d'autre qu'un ensemble de règles du jeu, que nous choisissons de suivre parce qu'elles nous semblent saines. En vertu du principe selon lequel rien ne saurait être soustrait à la critique, ces règles sont elles-mêmes susceptibles d'être critiquées (et nous ne manquerons pas de le faire au Chapitre 11).

1. Aucune loi, aucun modèle à quelque échelle que ce soit, et notamment aucune équation physique, n'est considérée comme valide *a priori* ni définitivement *vérifiée*, et surtout pas à une échelle autre que celle où elle a été établie,
2. Un modèle digne de ce nom doit pouvoir être soumis à des tests empiriques, c'est-à-dire confronté à des observations. Par conséquent, les conclusions tirées de cas « synthétiques » peuvent être utiles à la réflexion mais ont un intérêt scientifique limité, en l'absence de fondement empirique. Le modèle ne doit pas non plus nécessiter, pour fonctionner, de données qui ne soient pas couramment disponibles, car cela reviendrait à restreindre les possibilités de test,
3. Ces tests empiriques ne doivent pas avoir pour objectif de *confirmer* les modèles, mais au contraire de les mettre en défaut. Par conséquent, ils doivent être systématiques et de grande ampleur², et l'on ne doit procéder à des tris dans les observations qu'avec la plus grande précaution,

¹Suarez Miranda, *Viajes des Varones Prudentes*, Lib. IV, Cap. XIV, Lerida, 1658, cité par J.L. Borgès, *Histoire Universelle de l'Infamie*, 1935, et *Mathevet*, 2005.

²Popper parle de « *tentatives sincères et ingénieuses visant à les réfuter* ».

4. Les hypothèses supplémentaires ne doivent être introduites que pour des raisons de performance accrue, non pour des raisons « esthétiques ».

Ces règles nous incitent à privilégier des modèles *simples* au sens où l'entend *Popper* (1934, 1959), c'est-à-dire susceptibles d'être largement testés et réfutés :

“Nous pouvons répondre à toutes les questions épistémologiques que suscite le concept de simplicité si nous assimilons cette notion à celle de *degré de réfutabilité*.”

À mon point de vue, l'on doit dire qu'un système est *complexe au plus haut degré* si, conformément à la pratique conventionnaliste, on y tient comme à un système à jamais établi qu'on est résolu à sauver, chaque fois qu'il est menacé, par l'introduction d'hypothèses auxiliaires. En effet, le degré de réfutabilité d'un système ainsi protégé est égal à *zéro*. Ainsi sommes-nous ramenés, par notre concept de simplicité, [...] à cette règle ou principe qui nous interdit toute indulgence envers les hypothèses *ad hoc* et les hypothèses auxiliaires : au principe de parcimonie dans l'usage des hypothèses.”

Kirchner (2006) propose une démarche tout à fait similaire, insistant lui aussi sur le fait que les modélisateurs doivent rechercher activement les limites de leurs modèles :

“In my view, advancing the science of hydrology will require new hydrologic measurements, new methods for analyzing hydrologic data, and new approaches to modeling hydrologic systems. These three essential aspects of hydrology will all be advanced if we take full advantage of the linkages between them. Some promising directions forward, in my view, include (1) designing new data networks, field observations, and field experiments, explicitly recognizing the spatiotemporal heterogeneity of hydrologic processes, (2) developing “gray box” data analysis methods that are more compatible with the nonlinear, nonadditive character of hydrologic systems, (3) developing physically based governing equations for hydrologic behavior at the catchment or hillslope scale, recognizing that they may look different from the equations that describe the small-scale physics, (4) developing models that are minimally parameterized, and therefore stand some chance of failing the tests that they are subjected to, (5) developing ways to test models more comprehensively and incisively, given the intrinsic limitations of the available data.”

Incontestablement, ces principes conduisent à proposer des modèles qui paraissent « laids », c'est-à-dire dont la probabilité *logique* est faible : les modèles qui ne décrivent pas exhaustivement la réalité, en sciences de l'environnement, rentrent typiquement dans cette catégorie. Pourtant, si l'on suit *Popper*, ces modèles excessivement simples sont ceux qui ont

le plus grand contenu. Ce sont ceux qui *disent* le plus de choses au sujet du réel, parce qu'ils *interdisent* beaucoup (la gamme de comportements qu'ils autorisent est restreinte) : ils ont donc « *d'avantage d'occasions d'entrer en conflit avec des énoncés de base* », c'est-à-dire des compte-rendus d'observations (la classe de leurs « falsificateurs virtuels » est plus grande, pour utiliser la terminologie poppérienne). Ainsi, la loi de Darcy aurait été probablement d'un intérêt bien moindre si Darcy l'avait énoncée sous forme d'une relation polynomiale de degré huit entre le débit et le gradient de charge : hypothèse bien plus vague, donc bien plus probable sur un plan strictement logique, qu'une relation linéaire (cette forme étant incluse dans la première relation, on l'obtient en opérant une réduction formelle de sa dimension)... Bien sûr, nous ne comparons pas les modèles pluie-débit à la loi de Darcy : ils sont infiniment moins bien corroborés que cette dernière (ils sont même systématiquement falsifiés). Cependant, le principe empirique reste : la « laideur » des modèles est en un sens la condition nécessaire pour pouvoir trancher entre certaines alternatives et apprendre quelque chose sur la réalité. *A contrario*, on n'apprend rien d'une tautologie, ou d'un modèle qui nécessiterait, pour fonctionner, un niveau de description trop exhaustif. Pour reprendre l'analogie précédente, une carte à l'échelle $\frac{1}{1}$ ne nous serait absolument d'aucune utilité pour trouver notre chemin... Une carte est utile autant par ce qu'elle figure que par ce qu'elle omet, et il en va de même pour les modèles (ce n'est d'ailleurs pas une vision exclusivement utilitariste : cela nous semble vrai aussi bien pour les modèles de gestion que pour les modèles de pure compréhension).

Bien entendu, rien n'interdit de complexifier les modèles par la suite, s'ils s'avèrent définitivement trop squelettiques. La seule condition est que cette complexification donne lieu à une nouvelle, et encore plus large, mise à l'épreuve des observations... et ainsi de suite. Il ne s'agit pas de tomber dans les travers mentionnés précédemment pour l'exhaustivité, en faisant de la simplicité un critère esthétique, désirable en soi : elle n'est qu'un moyen.

1.3.2 Mise en oeuvre des tests

Nous pouvons maintenant présenter la façon concrète dont les modèles hydrologiques seront soumis au feu des tests empiriques, nos « *tentatives sincères et ingénieuses visant à les réfuter* »... Ces tentatives reposent sur quatre principes :

1. L'utilisation du plus large échantillon de bassins versants possible. Chaque bassin versant constituant *une* observation de la relation pluie-débit (en pratique, une observation de trois séries temporelles concomitantes : pluie, évapotranspiration potentielle et débit), l'utilisation de plusieurs centaines de bassins versants est nécessaire pour juger statistiquement de la qualité d'un modèle pluie-débit,

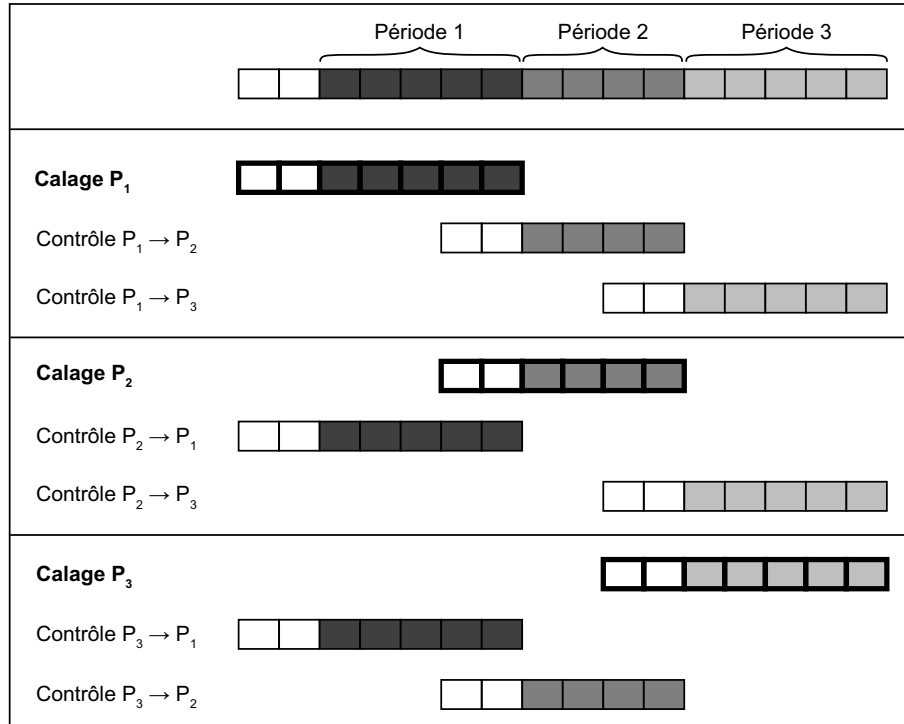


FIG. 1.3: Procédure de *split-sample test* symétrique sur un bassin versant (Klemeš, 1986) pour $n = 3$ périodes, donnant 3 critères de calage et 6 critères de contrôle. En calage comme en contrôle, le critère n'est calculé que pour les années « colorées », les années « blanches » étant des années d'initialisation où seuls les forçages climatiques sont utilisés.

2. Le recours, pour chaque bassin, à une procédure de *split-sample test* (Klemeš, 1986), décrite sur la Figure 1.3. Le jeu de données est découpé en n périodes ; sur chaque période, le modèle est *calé* (c'est-à-dire que l'on évalue les paramètres en minimisant une fonction objectif) puis *contrôlé* sur les $n - 1$ autres périodes. Nous obtenons donc $n(n - 1)$ mesures de la performance du modèle sur le bassin en contrôle, c'est-à-dire sur un jeu de données indépendant du jeu utilisé pour caler le modèle. Le plus souvent, nous utiliserons $n = 2$. Notons aussi que l'on permet au modèle une période d'initialisation (« préchauffage » utilisant les forçages climatiques réels) avant le début du calcul de la fonction objectif, pour limiter l'effet des conditions initiales sur la valeur du critère (en calage comme en contrôle). De façon à maximiser l'utilisation des données de débit, les périodes d'initialisation (où le débit observé n'est pas utilisé) sont choisies de façon à se superposer avec la période précédente.
3. La mesure, pour chaque période de calage ou de contrôle, de la performance du modèle au moyen d'un critère de type *Nash et Sutcliffe* (1970) :

$$C_{NS} = 1 - \frac{F}{F_0} \quad (1.1)$$

ou sa version bornée (Mathevet et al., 2006) :

$$C_{MM} = \frac{1 - \frac{F}{F_0}}{1 + \frac{F}{F_0}} \quad (1.2)$$

où F désigne l'erreur quadratique moyenne entre la série simulée et la série observée, et F_0 la variance de la série observée : ces séries peuvent être les débits, ou une transformation des débits (racine carrée, logarithme, etc.). Ces critères étant compris respectivement dans $] -\infty, 1]$ et $] -1, 1]$, on utilise couramment $100 \cdot C_{NS}$ et $100 \cdot C_{MM}$: nous éviterons autant que possible cette notation, qui induit en erreur (les critères C_{NS} et C_{MM} n'ont rien d'un pourcentage de variance expliquée). Par commodité, nous garderons tout de même le terme « *point* » pour désigner une variation de ± 0.01 sur ces critères. **Le lecteur peu familier du critère C_{MM} pourra au besoin se référer à la table de correspondance entre les valeurs de ce critère et celles du critère C_{NS} : cette table se trouve en Annexe A, page 267,**

4. La performance d'une structure est synthétisée par la distribution de ses critères de performance *en contrôle* sur l'échantillon de N bassins versants, soit $Nn(n-1)$ valeurs. Nous utiliserons couramment une représentation graphique de cette distribution sous forme de « boîtes à moustaches » (*boxplot*) :

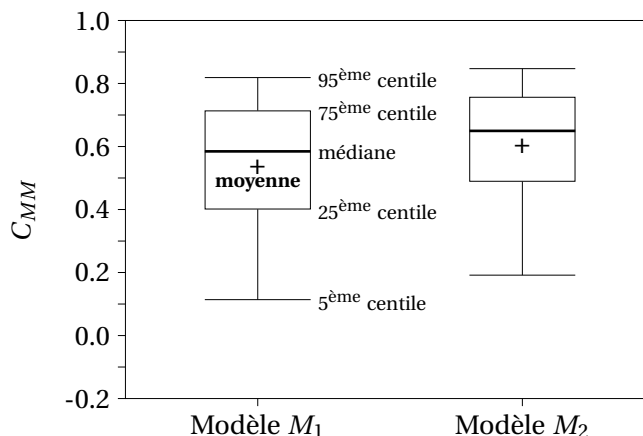


FIG. 1.4: Représentation synthétique de la distribution des critères de performance. La *boîte* contient 50% des valeurs et les *moustaches*, 90%.

En particulier, la moyenne du critère C_{MM} est utilisée comme critère synthétique (Mathevet et al., 2006) : ici, nous dirions que le modèle M_2 est meilleur que le modèle M_1 .

Le caractère « ingénieux » de ces tests, notamment l'utilisation d'un critère d'erreur quadratique et de sa moyenne sur un grand échantillon, sera discuté au dernier chapitre de la thèse. En revanche, leur sincérité semble indiscutable !

1.4 Conclusion

Comme on aura pu le constater, nous estimons que les oppositions classiques entre modèles globaux et distribués, « empiriques » ou « à base physique », ne sont ni intéressantes ni pertinentes : la seule opposition valable est celle qui distingue les modèles testables des modèles non-testables. Construire des modèles testables implique, entre autres, de renoncer à la croyance selon laquelle la description de chaque centimètre cube de la Terre est pour *demain*, et que les modèles ne peuvent donc pas vraiment être testés *aujourd'hui*. Cette attitude est assez compréhensible, comme le souligne Bergström (1991) : « *going from complex to simpler model structures requires an open mind, because it is frustrating to have to abandon seemingly elegant concepts and theories* ».

Nous pouvons conclure que le choix d'un modèle ne devrait pas être une question « idéologique » mais simplement résulter de l'analyse de l'échelle à laquelle se pose le problème. Si le problème nécessite une représentation à l'échelle de la parcelle, on choisira une discrétisation et des lois de comportement à l'échelle de la parcelle ; si l'échelle pertinente est le bassin versant, on choisira un modèle à cette échelle ! Nous pensons que le fait de prendre en compte *plus* de processus n'est en rien une excuse pour faire confiance à un modèle plutôt qu'à un autre ; cette attitude repose sur la *croyance* que si un modèle ne prend pas en compte explicitement les processus de petites échelles bien connus, ce modèle va perdre sa crédibilité, et que s'il lui arrive malgré tout de marcher, ce sera pour de « mauvaises » raisons : le raisonnement est circulaire (Michel *et al.*, 2006). Par exemple, si un modèle à l'échelle de la parcelle est moins performant qu'un modèle à l'échelle du bassin versant pour simuler les débits, pourquoi devrait-on faire d'avantage confiance au premier pour étudier l'impact d'un changement d'occupation du sol, sous prétexte que le second ne prend pas en compte ce facteur ? La vérité est que l'on ne devrait faire confiance ni à l'un, ni à l'autre : ni au modèle global, parce que la loi déterminée se réfère bien au système complet, mais que cette loi ne sera pas stationnaire (on ne vérifie pas *l'unité d'action*) ; ni au modèle parcellisé, parce que les lois identifiées pour les différents compartiments sont certes stationnaires, mais que leur inventaire est forcément incomplet (sinon, ce modèle serait au moins aussi bon que le modèle global pour simuler le comportement présent) : preuve que c'est, cette fois, *l'intégrité* de l'action qui pose problème !

Ainsi, nous nous efforcerons de travailler sur des structures simples en les complexifiant peu à peu au moyen d'hypothèses dont nous questionnerons *isolément* la pertinence. Comme le résume Michel *et al.* (2006), il n'y a pas d'autre raison pour préférer les constructions parcimonieuses : elles soulèvent des questions, laissent de la place au doute, et permettent de tester facilement des hypothèses.

Concluons ce chapitre sur une note historique, peut-être un peu pessimiste. Il se pourrait que cette activité relativement nouvelle qu'est la modélisation quantitative en sciences de l'environnement, et les obstacles auxquels elle est confrontée, soient en train de dissiper quelques illusions produites par les succès (incontestables) de la méthode réductionniste dans d'autres domaines, dont l'exemple typique est la mécanique classique. En plein XVIII^{ème} siècle, Buffon, pourtant newtonien convaincu, présentait déjà que cette méthode basée sur la géométrisation et la déglobalisation des problèmes se heurterait inévitablement à certaines limites, comme il l'écrit dans le Premier Discours de son *Histoire Naturelle* :

“Il est vrai que cette union des Mathématiques et de la Physique ne peut se faire que pour un très-petit nombre de sujets; il faut pour cela que les phénomènes que nous cherchons à expliquer, soient susceptibles d'être considérés d'une manière abstraite, et que de leur nature ils soient dénués de presque toutes qualités physiques, car pour peu qu'ils soient composés, le calcul ne peut plus s'y appliquer. La plus belle et la plus heureuse application qu'on en ait jamais faite, est au système du monde; et il faut avouer que si Newton ne nous eût donné que les idées physiques de son système, sans les avoir appuyées sur des évaluations précises et mathématiques, elles n'auroient pas eu à beaucoup près la même force; mais on doit sentir en même temps qu'il y a très-peu de sujets aussi simples, c'est-à-dire, aussi dénués de qualités physiques que l'est celui-ci; car la distance des planètes est si grande qu'on peut les considérer les unes à l'égard des autres comme n'étant que des points; on peut en même temps, sans se tromper, faire abstraction de toutes les qualités physiques des planètes, et ne considérer que leur force d'attraction: leurs mouvements sont d'ailleurs les plus réguliers que nous connoissons, et n'éprouvent aucun retardement par la résistance: tout cela concourt à rendre l'explication du système du monde un problème de mathématique, auquel il ne falloit qu'une idée physique heureusement conçue pour le réaliser; et cette idée est d'avoir pensé que la force qui fait tomber les graves à la surface de la terre, pourroit bien être la même que celle qui retient la lune dans son orbite. Mais, je le répète, il y a bien peu de sujets en Physique où l'on puisse appliquer aussi avantageusement les sciences abstraites, et je ne vois guère que l'Astronomie et l'Optique auxquelles elles puissent être d'une grande utilité; l'Astronomie par les raisons que nous venons d'exposer, et l'Optique parce que la lumière étant un corps presque infiniment petit, dont les effets s'opèrent en ligne droite avec une vitesse presque infinie, ses propriétés sont presque mathématiques, ce qui fait qu'on peut y appliquer avec quelque succès le calcul et les mesures géométriques.”

Les sciences dans lesquelles Buffon constate un succès de la géométrisation — sa liste, qui ne contient « *guère que l'Astronomie et l'Optique* », s'est avérée dans les faits un peu plus

longue! — sont celles qui ont pour objet des systèmes, très grands ou très petits, mais finalement assez simples, où un seul type d'interactions domine. Il serait peut-être plus sage de reconnaître que l'hydrologie ne rentre tout simplement pas dans cette catégorie, et de prendre conscience, malgré toute la puissance de calcul dont nous disposons,

“des inconvénients où l'on tombe lorsque l'on veut appliquer la Géométrie et le calcul à des sujets de Physique trop compliqués, à des objets dont nous ne connaissons pas assez les propriétés pour pouvoir les mesurer; on est obligé dans tous ces cas de faire des suppositions toujours contraires à la Nature, de dépouiller le sujet de la plupart de ses qualités, d'en faire un être abstrait qui ne ressemble plus à l'être réel, et lorsqu'on a beaucoup raisonné et calculé sur les rapports et les propriétés de cet être abstrait, et qu'on est arrivé à une conclusion toute aussi abstraite, on croit avoir trouvé quelque chose de réel, et on transporte ce résultat idéal dans le sujet réel, ce qui produit une infinité de fausses conséquences et d'erreurs.”

Ce paragraphe, écrit il y a deux cents cinquante ans, est à méditer!

2

Constitution d'un échantillon de bassins versants

“ Chaque fois que je vois le chiffre Un, j'ai envie de l'aider à s'échapper... Il a constamment à ses trousses, derrière, le zéro qui veut le rattraper et devant, toute la mafia des grands nombres qui le guette.”

Romain GARY.

2.1 Introduction

Nous avons exposé au chapitre précédent notre façon d'envisager la modélisation des systèmes naturels en général et des bassins versants en particulier. Nous avons insisté sur le fait que nous voyons les modèles comme de simples hypothèses, qu'il convient de soumettre à l'épreuve de tests empiriques. Cela suppose, dans notre cas, de confronter les modèles pluie-débit à une large gamme de comportements hydrologiques, résultant d'une diversité de conditions climatiques, géologiques, etc.

L'utilisation de grands échantillons de bassins versants est une pratique qui, sans être encore la règle, commence à rentrer dans les mœurs. Ce souci de généralité étant une constante de toutes les recherches effectuées dans l'équipe d'hydrologie du CEMAGREF d'Antony, cette thèse n'a pas échappé à la règle. Nous nous sommes donc attelés à la constitution d'un large échantillon de bassins versants, principalement français, pourvus de données de pluie et de débit au pas de temps horaire : en effet, sans abandonner le pas de temps journalier qui reste la référence pour de nombreuses applications, nous avons souhaité placer ces développements dans la suite du travail de *Mathevet* (2005).

2.2 Construction d'une base de données pluie-débit horaire française

Dans cette section, nous décrivons la constitution d'un échantillon de bassins versants français qui constituera la base de nos recherches.

2.2.1 Critères de sélection des débits

Nous avons cherché à constituer l'échantillon le plus large possible de bassins versants non-influencés. Météo France ayant mis gracieusement à notre disposition l'archive pluviométrique SAFRAN (cf. section 2.2.4) couvrant la France entière sur la période 1995–2005, la contrainte de disponibilité des données portait exclusivement sur les débits. La procédure suivie a donc été la suivante :

- Nous sommes partis d'une liste de 3054 stations hydrométriques répertoriées à **signification hydrologique** et **sans influence** dans la banque de données HYDRO.
- Nous en avons retenu 2950 ayant une surface supérieure à 10 km^2 — limite arbitraire pour l'approche globale —, et inférieure à 10000 km^2 — le travail au pas de temps horaire étant moins intéressant pour les très grands bassins¹,

¹De toutes façons, les très grands bassins ne posent généralement pas de problème en modélisation pluie-débit.

- Nous avons récupéré les données de **débites journaliers** (QJM) pour ces 3054 stations sur 1995–2005, de façon à faire une présélection sur les stations disponibles à un pas de temps plus fin tout en limitant la quantité de données extraites. La requête a donné un résultat positif pour **1284 stations** (i.e., 1284 stations disposaient d'au moins une donnée sur cette période).
- Une seconde vérification de l'absence d'influences anthropiques a été effectuée pour ces 1284 stations. Pour cela, nous avons utilisé la base de données de barrages constituée par *Payan* (2007) : nous avons inventorié les stations situées à l'aval de chaque ouvrage puis, pour chacune de ces stations potentiellement influencées, nous avons fait la somme des capacités de stockage (en Mm^3) à l'amont. Lorsque cette somme représentait plus de 10 mm de lame d'eau répartie sur le bassin versant, la station a été écartée (24 stations).
- La taille maximale des fichiers exportés depuis la B.D. HYDRO étant limitée, nous avons choisi de récupérer les données **au pas de temps variable** (QTVAR) pour les stations retenues, plutôt que le pas de temps horaire lui-même, de façon à limiter le nombre de requêtes et de fichiers à récupérer. Parmi la sélection faite d'après le pas de temps journalier, **1182 stations** disposaient effectivement de données QTVAR sur la période 1995–2005.

2.2.2 Constitution des séries de débits horaires

La Figure 2.1 illustre le type de données QTVAR récupérées depuis la Banque HYDRO. Les débits horaires ont été reconstitués par interpolation linéaire du pas de temps variable et intégration sur chaque heure.

Nous n'avons retenu que les stations pour lesquelles moins de 50% des données de débit étaient manquantes sur chacune des périodes 01/08/1995 – 31/07/2000 et 01/08/2000 – 31/07/2005¹, de façon à pouvoir utiliser sans problème la procédure de calage-contrôle décrite au Chapitre 1. Le chiffre de 50% peut paraître élevé — il faut bien fixer un seuil —, mais en pratique il est largement suffisant pour caler et évaluer un modèle ; de plus, dans 80% des cas, le taux de lacunes est bien moindre (moins de 10%), comme illustré sur la Figure 2.2. On a constaté que le taux de lacunes était un peu plus élevé pour la période 2000–2005 : cela s'explique par le fait que certaines données récentes n'avaient pas encore été « versées » dans la Banque HYDRO à la date de l'extraction (Janvier 2006). Ce délai de versement, différent selon les DIREN, explique aussi la sous-représentation regrettable des stations de l'Arc Méditerranéen.

¹On notera que cette condition est plus exigeante que celle d'avoir moins de 50% de lacunes sur la période totale 1995–2005.

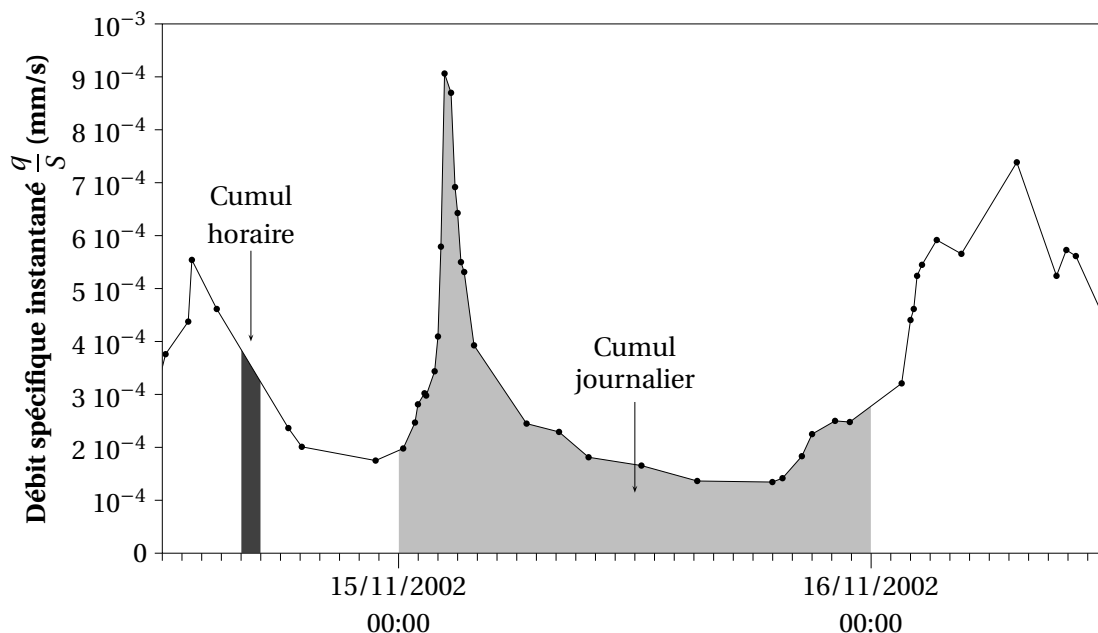


FIG. 2.1: Méthode de passage des débits instantanés aux cumuls par pas de temps : interpolation linéaire du pas de temps variable (QTVAR) puis intégration sur les intervalles choisis. Exemple de la Barberolle à Barbières (V4015030), 10 km^2 ($10^{-3} \text{ mm/s} \Leftrightarrow 10 \text{ m}^3/\text{s}$).

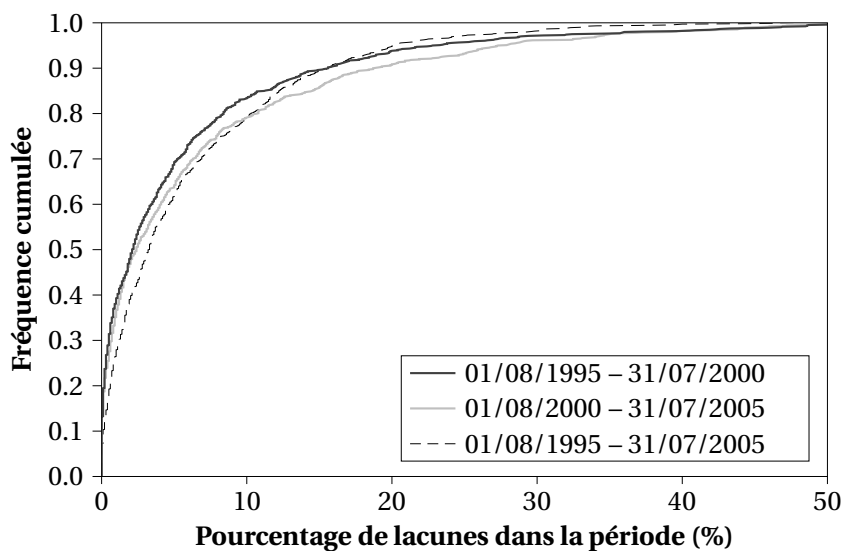


FIG. 2.2: Taux de lacunes des séries de débits horaires par périodes (fréquence sur les 1040 stations de l'échantillon).

2.2.3 Extraction des contours des bassins versants

Une fois fixée la liste des stations choisies, les contours des bassins versants associés ont été extraits d'un Modèle Numérique de Terrain (MNT) au pas de 75 m. Ce travail colossal a été réalisé par J.L. Rosique, du CEMAGREF d'Antony. Les stations et contours des 1040 bassins versants retenus sont tracés sur la carte de la Figure 2.3, avec le MNT en fond.

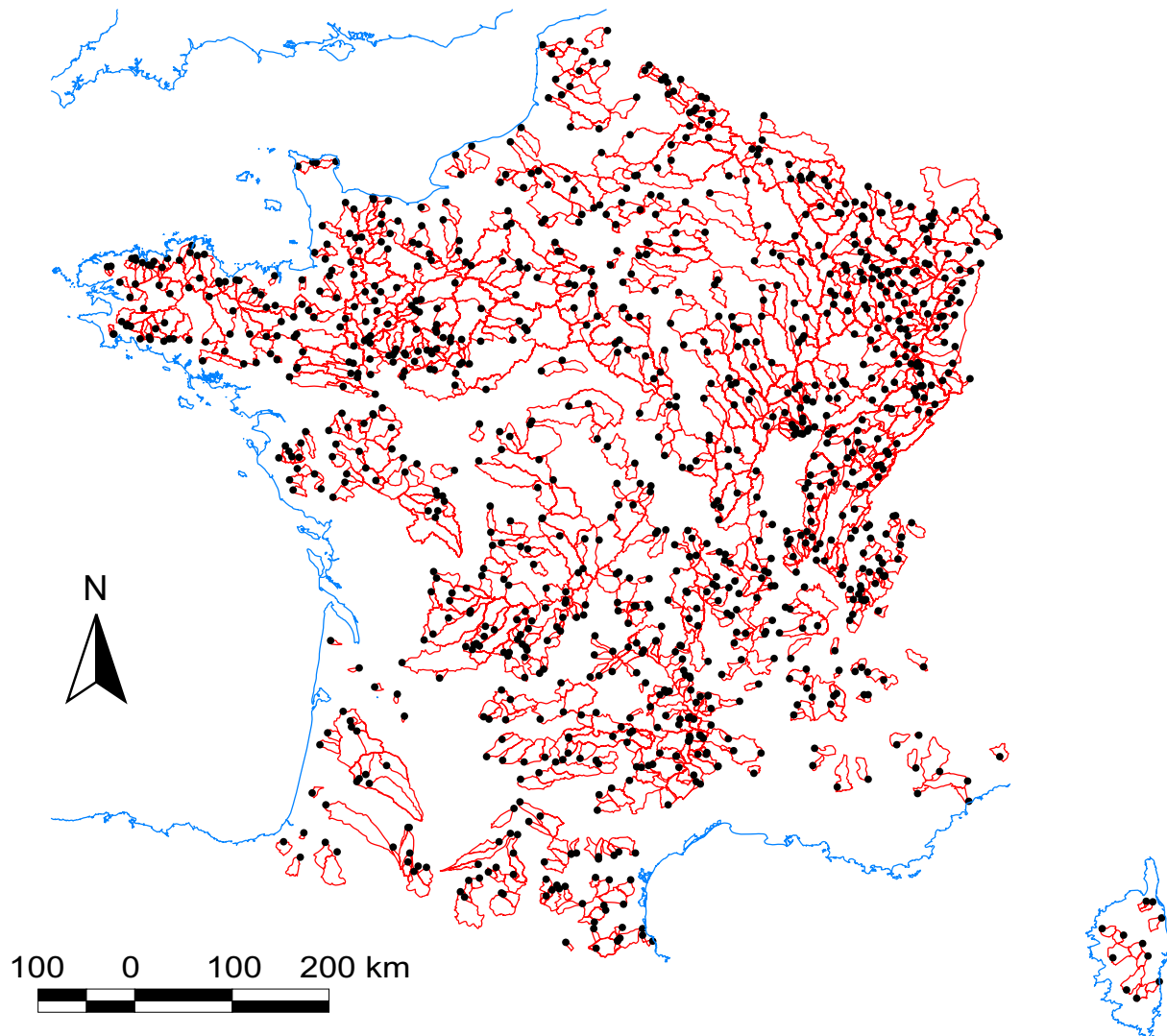


FIG. 2.3: Échantillon de 1040 bassins versants français de surface comprise entre 10 et 10000 km² : contours des bassins et localisation des stations hydrométriques.

2.2.4 Constitution des séries de pluies horaires

2.2.4.1 Construction des pluies journalières de bassin

Les données de l'analyse SAFRAN (Système d'Analyse Fournissant des Renseignements Atmosphériques pour la Nivologie) ont été mises à notre disposition par Météo France sur la période 1995–2005. Historiquement, comme son nom le rappelle, ce système a été conçu par le Centre d'Étude de la Neige (*Durand et al.*, 1993; *Durand et al.*, 1995) pour améliorer l'estimation des précipitations en altitude. Il est aujourd'hui utilisé plus généralement pour analyser et interpoler les observations journalières de précipitations du réseau pluviométrique au sol de Météo France. Ce réseau est formé par les 1200 postes du réseau synoptique principal, auxquels s'ajoutent les nombreux postes du Réseau Climatique d'État (*Le Moigne*, 2002). Ce sont ainsi en moyenne 4100 observations journalières de pluie qui sont collectées sur l'ensemble du territoire (Figure 2.4), soit une observation pour 135 km² en moyenne.

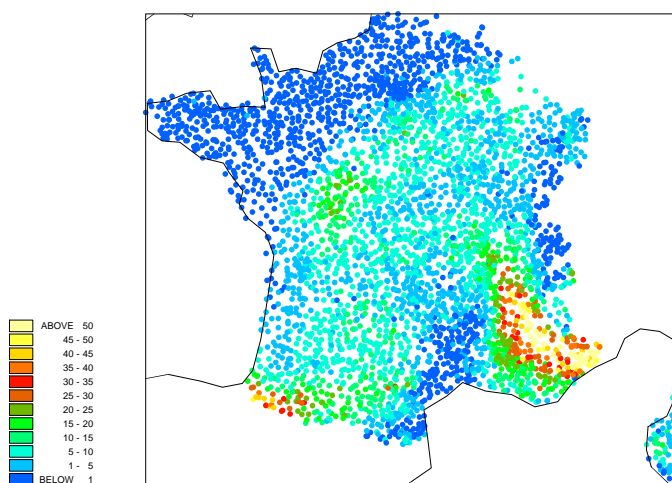


FIG. 2.4: Cumul de pluie (mm) sur 24 h (du 14/02/2002 06TU au 15/02/2002 06TU) observé aux postes synoptiques et stations automatiques, y compris le RCE (*Le Moigne*, 2002).

Les observations de ce réseau très dense (cumuls pluviométriques entre 6h TU du jour et 6h TU du lendemain) sont ensuite interpolées par zones climatiques sensiblement homogènes (découpage SYMPOSIUM II, illustré Figure 2.5), en prenant en compte un gradient orographique dans chaque zone (par tranches de 300 m d'altitude). Les résultats sont ensuite moyennés sur une grille de 8 × 8 km² (Figure 2.6), qui est la résolution utilisée par le modèle de surface ISBA.

À partir de ce type de cartes de précipitations journalières, nous pouvons construire des pluies de bassin à partir de l'intersection de la grille SAFRAN et du contour de chaque bassin versant, et en affectant comme poids à chaque maille sa fraction à l'intérieur du bassin.



FIG. 2.5: Découpage SYMPOSIUM II de la France en 615 zones climatiquement homogènes (*Le Moigne, 2002*).

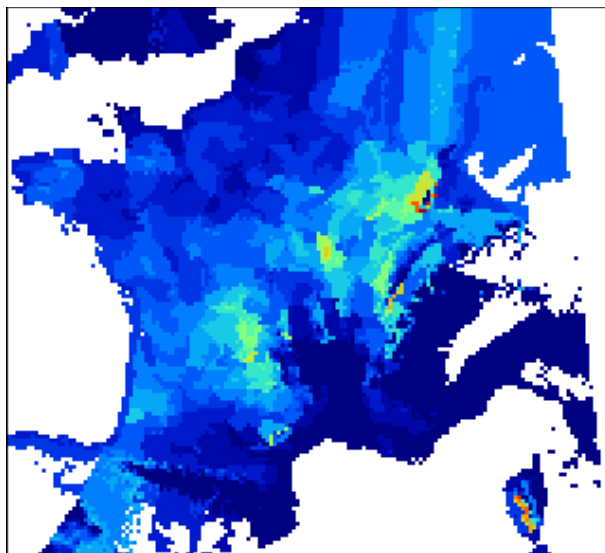


FIG. 2.6: Format des sorties de l'analyse SAFRAN (grille de $8 \times 8 \text{ km}^2$).

L'analyse SAFRAN fournit aussi des précipitations horaires mais elles sont obtenues indirectement : les cumuls journaliers sont désagrégés en fonction d'un modèle d'humidité du sol. Nous préférons, pour notre étude, une désagrégation basée sur des observations issues du réseau de pluviographes automatiques, dont la procédure est décrite au paragraphe suivant.

2.2.4.2 Désagrégation horaire des pluies de bassin journalières

Nous avons décrit jusqu'à présent la procédure nous permettant d'obtenir des pluies de bassin journalières. Pour les désagréger en pluies de bassin horaires, il faut avoir, pour

chaque journée, une fonction de répartition sur les 24 heures. Nous obtenons cette fonction de répartition à partir du réseau de pluviographes automatiques de Météo France.

On pourra se demander pourquoi nous n'avons pas utilisé directement ce réseau pour calculer les pluies de bassin horaires. Il y a deux raisons à cela : d'une part, ce réseau automatique télétransmis est nettement moins dense que le réseau journalier de la Figure 2.4. D'autre part, la corrélation spatiale entre les cumuls horaires décroît beaucoup plus rapidement avec la distance que la corrélation entre cumuls journaliers (i.e., même à densité du réseau constante, l'interpolation des cumuls horaires est souvent bien plus hasardeuse que celle des cumuls journaliers). Ainsi, pour éviter des problèmes de bilan en eau sur le long terme, nous préférons prendre comme cumuls journaliers les valeurs données par SAFRAN, et n'utiliser que la *dynamique* donnée par le réseau horaire.

Nous avons donc sélectionné 580 postes pluviographiques présentant moins de 15% de lacunes sur la période 1995–2005. Ce réseau, ainsi que les polygones Thiessen associés, sont présentés sur la Figure 2.7 : on notera la différence flagrante de densité avec le réseau de la Figure 2.4.

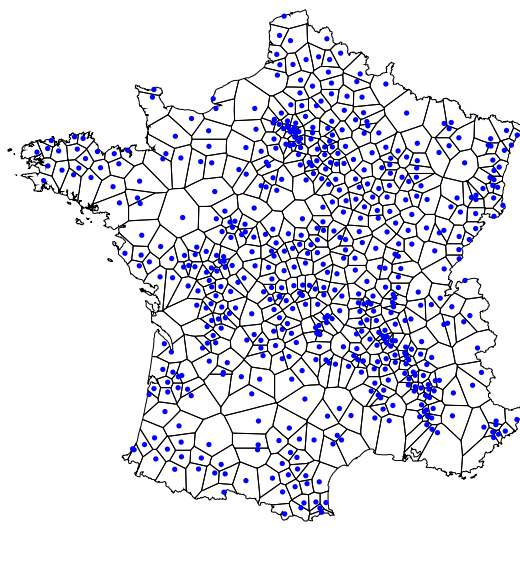


FIG. 2.7: Réseau de 580 postes pluviographiques horaires utilisés pour désagréger les cumuls journaliers issus de SAFRAN, et polygones de Thiessen associés.

Sur la base de ce réseau, nous calculons pour chaque bassin une moyenne pondérée de chroniques pluviographiques horaires, les poids étant définis par la surface de l'intersection entre chaque polygone de Thiessen et le contour du bassin (Figure 2.8, premier graphe en partant du haut). Ensuite, nous calculons la fonction de répartition pour chaque journée entre 06 h TU et 06 h TU du lendemain (deuxième graphe). Cette fonction de répartition est ensuite utilisée pour désagréger la chronique des pluies de bassin journalières (présentée

sur le troisième graphe), ce qui permet d'obtenir la série horaire finale (en trait plein sur le quatrième graphe).

Lorsque tous les pluviographes couvrant un bassin versant étaient indisponibles à une date donnée, nous avons gardé la répartition horaire donnée par SAFRAN (trait discontinu sur le quatrième graphe). Cette situation reste marginale.

Globalement, cette méthode tend à lisser la série initiale : les évènements « sous-captés » par le réseau automatique horaire sont amplifiés sur la base du cumul journalier vu par le réseau journalier, tandis que les évènements éventuellement « sur-captés » sont atténués. Bien évidemment, la prise en compte des effets d'advection entre postes (décalages temporels) dans la construction de la série pondérée permettrait d'améliorer encore la méthode.

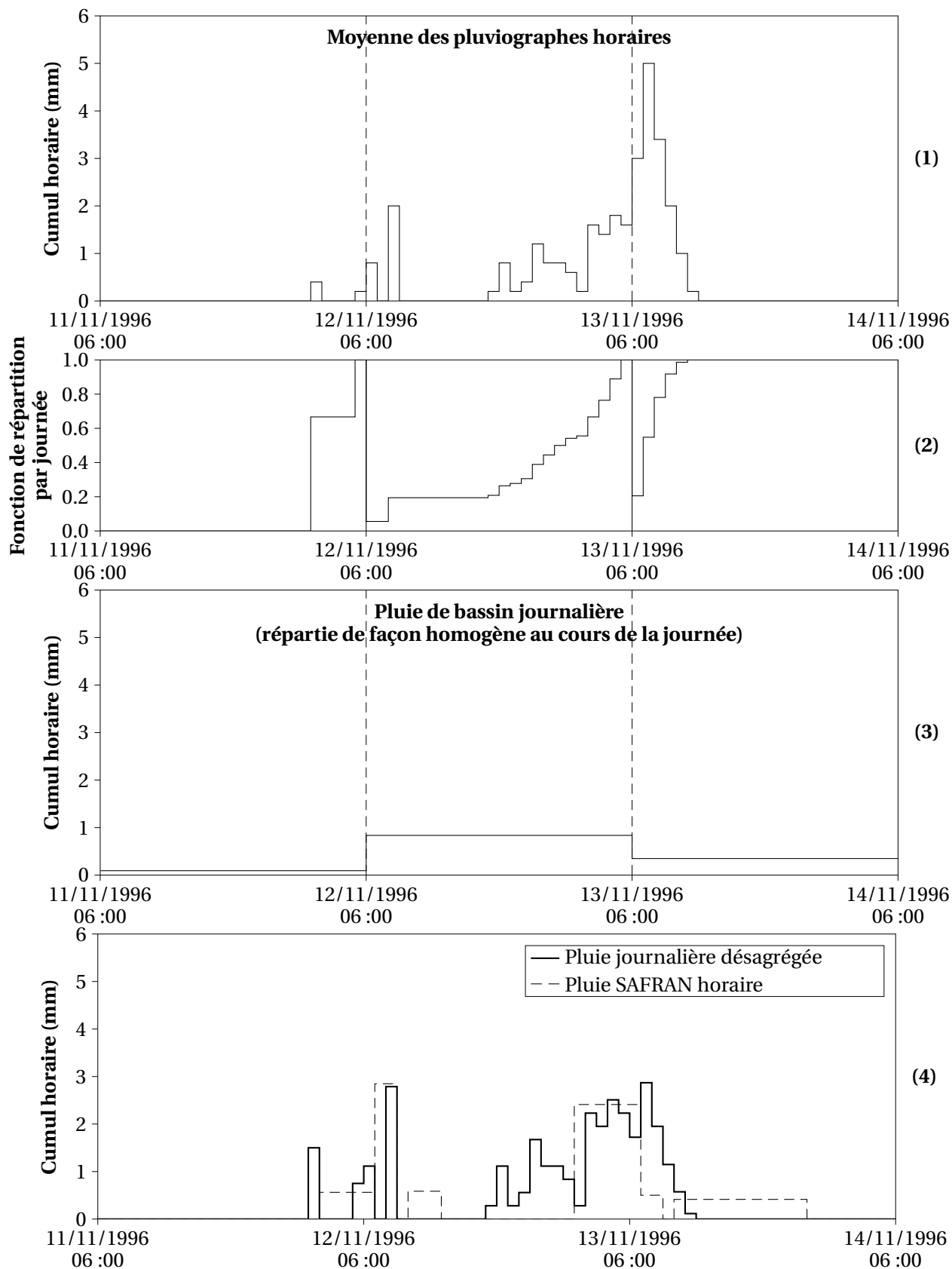


FIG. 2.8: Méthode de désagrégation des pluies de bassins journalières : (1) Calcul de la moyenne des pluies horaires observées par les postes pluviographiques automatiques; (2) Calcul de la répartition de ces pluies à l'intérieur de chaque journée; (3) et (4) Utilisation de cette fonction de répartition pour désagréger la pluie de bassin journalière. Le graphe tout en bas montre la série obtenue comparée à la série horaire de SAFRAN (désagrégée d'après un modèle d'humidité du sol).

2.2.5 Constitution des séries d'ETP de bassin

Pour chaque bassin versant, trois séries d'évapotranspiration potentielle (ETP) ont été construites :

- une série d'ETP datées, calculées par la formule de *Oudin* (2004),
- une série d'ETP non datées (interannuelles), calculées avec cette même formule,
- une série d'ETP de Penman non datées, construites d'après le réseau de 88 stations synoptiques de Météo France.

En pratique, nous utiliserons systématiquement les chroniques d'ETP non datées calculées par la formule de L. Oudin, qui est la plus robuste des trois.

2.2.5.1 Construction des ETP de Oudin datées et interannuelles

Oudin (2004) a proposé la formule suivante pour l'ETP journalière :

$$ETP(J) = \begin{cases} \frac{Re(\varphi, J) \left(\frac{T_a(J) + 5}{100} \right)}{\lambda \rho} & \text{si } T_a(J) > -5 \\ 0 & \text{sinon} \end{cases} \quad (2.1)$$

où T_a désigne la température de l'air (en °C), Re le rayonnement extraterrestre ($\text{MJ} \cdot \text{m}^{-2} \cdot \text{jr}^{-1}$), λ la chaleur latente de vaporisation de l'eau ($\text{MJ} \cdot \text{kg}^{-1}$) et ρ sa masse volumique ($\text{kg} \cdot \text{m}^{-3}$).

Cette ETP n'est fonction que de la température de l'air T_a et du rayonnement extraterrestre Re (lui-même fonction de la latitude φ et du jour julien J), par les formules suivantes (*Allen*, 1998) :

$$\lambda = 2.501 - 0.002361 T_a \quad (\text{en } \text{MJ} \cdot \text{kg}^{-1})$$

$$Re = \frac{24 \times 60}{\pi} G_{sc} d_r \left[\omega_s \sin(\varphi) \sin(\delta) + \cos(\varphi) \cos(\delta) \sin(\omega_s) \right]$$

où G_{sc} constante solaire = $0.0820 \text{ MJ} \cdot \text{m}^{-2} \cdot \text{min}^{-1}$

d_r distance relative inverse Terre-Soleil

ω_s angle horaire au coucher de Soleil (rad)

φ latitude (rad)

δ déclinaison solaire (rad)

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

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

$$\omega_s = \arccos\left[-\tan(\varphi) \tan(\delta)\right]$$

L'archive SAFRAN sur 1995–2005 comportant également les données de températures diurnes, nous avons pu calculer des cartes d'ETP à la même résolution ($8 \times 8 \text{ km}^2$). Les ETP interannuelles ont été obtenues en calculant, pour chaque maille, les 365 températures moyennes interannuelles et en appliquant la même formule.

2.2.5.2 Construction des ETP de Penman interannuelles

Les ETP de Penman journalières interannuelles ont été obtenues par moyennes pondérées des ETP calculées aux 88 stations synoptiques de Météo France, selon la méthode des polygones de Thiessen décrites plus haut pour la pluie.

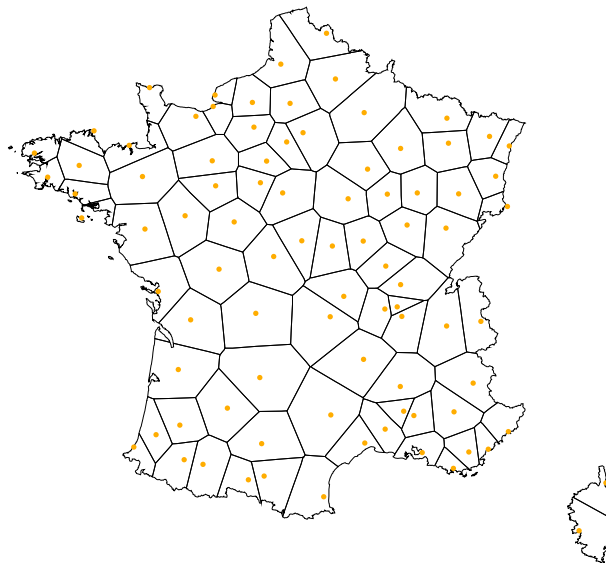


FIG. 2.9: Réseau Météo France de 88 stations synoptiques d'ETP.

2.3 Données étrangères

Au pas de temps journalier, nous avons également utilisé une base de données de 331 bassins versants australiens supplémentaires de façon à augmenter la diversité climatique de l'échantillon. Cet échantillon nous a été fourni par Francis Chiew du CSIRO Land and Water.

La localisation de ces bassins versants, situés principalement dans le Sud-Est de l'Australie (États de Nouvelle-Galle du Sud, Victoria et Tasmanie), est visible sur la Figure 2.10.

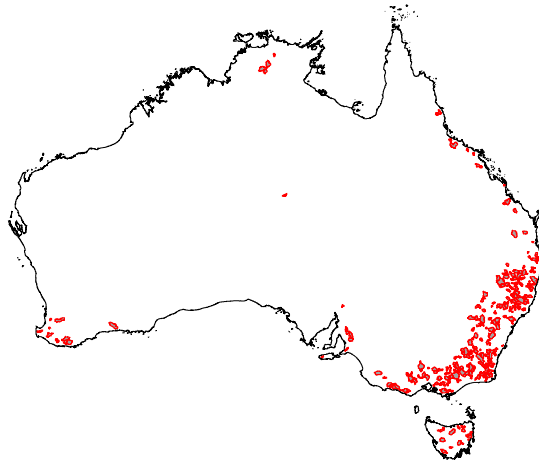


FIG. 2.10: Carte des 331 bassins versants australiens disponibles au pas de temps journalier

2.4 Résumé des échantillons

La Figure 2.11 donne un aperçu de la diversité hydroclimatique de l'échantillon, par les distributions de cinq grandeurs :

- la surface des bassins versants,
- la pluie annuelle moyenne,
- l'ETP annuelle moyenne,
- le module (lame d'eau moyenne écoulée annuellement),
- la pluie journalière décennale

Les échantillons apparaissent très contrastés, l'échantillon australien amenant une diversité bienvenue pour compenser la sur-représentation des bassins humides dans l'échantillon français (malheureusement pas au pas de temps de horaire). Si les distributions de la pluviométrie annuelle sont à peu près comparables pour les deux échantillons (1020 mm en moyenne pour les bassins français, 980 pour les bassins australiens), la distribution des pluies journalières décennales traduit en revanche une bien plus grande irrégularité des précipitations pour les bassins australiens, ce que confirme la différence entre les cumuls d'ETP annuelle. Ces contrastes climatiques se reflètent logiquement dans la distribution des débits : la médiane des modules vaut 360 mm pour les bassins français, contre 170 pour les bassins australiens. Le phénomène d'intermittence est également beaucoup plus répandu dans l'échantillon australien : 60% des bassins y sont soumis, contre seulement 10% des bassins français.

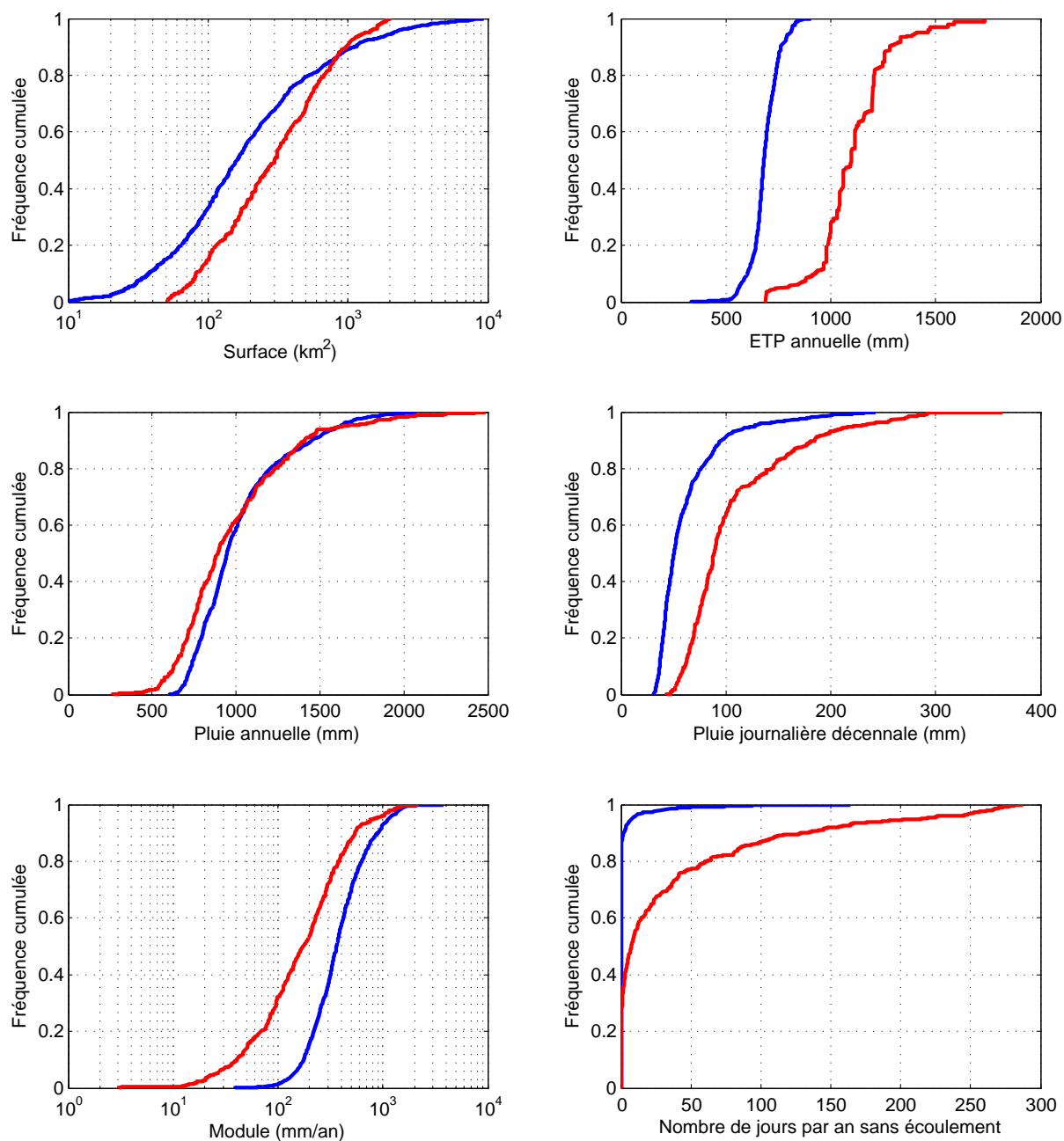


FIG. 2.11: Résumé des caractéristiques des bassins versants étudiés. Les distributions en bleu concernent les 1040 bassins versants français, celles en rouge concernant les 331 bassins australiens.

Nous verrons au chapitre suivant que la présence des bassins arides est importante pour le sujet, la recharge des aquifères par infiltration à la faveur d'une concentration de ruissellement étant un processus fréquent.

2.5 Faut-il faire le tri dans les données ?

Lorsque l'on présente des résultats sur un groupe de bassins versants que l'on n'a pas étudiés individuellement très en détail, il n'est pas rare d'avoir à faire face à des remarques du genre : « *Vous avez regardé les chroniques ?! Vous avez un Nash à 80, mais ça ne signifie rien, la station hydrométrique sur la rivière Machin a été détruite par la crue du 31 Février 2002 et les crues précédentes sont bien plus faibles !* » ou, à l'inverse : « *Votre modèle donne de mauvais résultats, mais c'est normal, le réseau de pluviomètres au sol capte mal les évènements (vous devriez utiliser le radar)* », etc., etc. Lorsque l'on est décidé à défendre, ou inversement à discréditer un modèle, la question de la qualité et de la représentativité des données utilisées est une arme de tout premier choix, un argument massue.

Pourtant, nous pensons que dans le cadre méthodologique que nous nous sommes fixé, ce problème n'en est pas vraiment un.

Si nous acceptons le fait qu'un modèle puisse donner de bons résultats pour de mauvaises raisons (cela fait partie des règles du jeu de l'approche empirique), nous devons en contrepartie accepter les cas où ce même modèle donne de mauvais résultats pour de probablement « bonnes » raisons. En effet, par définition, nous ne pouvons pas savoir dans lequel des deux cas nous nous trouvons. Il n'y a pas d'inconvénients majeurs à présenter des distributions de critères dont une petite partie sont négatifs :

- soit ces cas sont effectivement dûs à des données douteuses, et dans ce cas tous les modèles sont logés à la même enseigne dans des tests comparatifs. Puisque ces échecs n'ont pas ou peu à voir avec le phénomène modélisé (la relation pluie-débit), ils n'ont probablement aucune incidence sur le classement statistique des modèles,
- soit ils correspondent à des déficiences du modèle, et les retirer de l'évaluation serait risqué.

Bien entendu, cela n'a de sens que dans un cadre de *comparaison* de modèles, et ce sur de grands échantillons de bassins versants. Il faut d'ailleurs noter que, contrairement à une opinion couramment répandue, les modèles simples n'ont pas plus de raisons de « marcher pour de mauvaises raisons » que n'importe quel autre modèle. Comme le dit *Linsley* (1982) :

“At this point one frequently is told that it is better to use some simpler method since the data are not good enough for the model. It may be true that the data are not adequate but if one reflects on my comments regarding accuracy it will be noted that what I have said applies equally to all hydrologic models. If the data are too poor for the use of a good simulation model they are also inadequate for any other model. Nothing in the simpler models substitutes for the accuracy of the data.”

Ainsi, un tri dans les données peut être effectué avant le début de l'exercice de modélisation¹. Nous ne pousserons pas la provocation jusqu'à nier les énormes problèmes de mesure de la pluie et du débit auxquels les hydrologues doivent faire face : erreurs aléatoires dues à des dysfonctionnements des appareils de mesure, biais systématiques dans l'estimation des pluies de bassin du fait des problèmes de captation ou de représentativité des pluviomètres, erreurs d'extrapolation et non-stationnarités des courbes de tarage, etc. Inutile d'insister, cet inventaire déprimant a déjà été fait des centaines de fois (voir par exemple *Moulin, 2007*). Pourtant, nous estimons que *sur un grand échantillon de bassins versants*, toutes ces erreurs se ramènent plus ou moins à des erreurs aléatoires : tel bassin voit sa pluie systématiquement sous-estimée, mais elle est sur-estimée sur le bassin d'à côté... Il ne s'agit pas tant de rappeler que le passage en revue systématique des données pour 1000 bassins versants serait un exercice très fastidieux, voire inhumain : le réel problème est que nous manquerions de toute façon de règles objectives pour éliminer des données *a priori*. À partir de combien de postes pluviométriques a-t-on une « bonne estimation » de la pluie de bassin ? Comment décide-t-on qu'une chronique de débit est « manifestement non stationnaire »?... Nous sommes ouverts à ce genre de questions et de critiques, mais elles ne nous semblent pas prioritaires pour faire progresser la modélisation pluie-débit.

En revanche, éliminer des observations sur la base d'échecs de modélisation nous semble être une faute professionnelle. Ce n'est rien d'autre qu'une pétition de principe de la validité du modèle : de l'apriorisme de la pire espèce. De plus, un autre modèle aurait peut-être donné, sinon de *bons* résultats, du moins de *meilleurs* : c'est un élément de comparaison.

Il y a peut-être un inconvénient en terme d'image à présenter des résultats dont une partie est calamiteuse. Cet inconvénient, très subjectif et sans importance scientifique, est minime au regard des risques encourus à faire le tri dans les données : à long terme, l'accroissement artificiel de la performance qui en résulte a des conséquences infiniment plus graves sur la crédibilité des modèles (*Andréassian et al., 2007 ; Le Moine et al., 2007*), notamment telle qu'elle est perçue de la part des utilisateurs.

¹bien que nous pensions, pour notre part, que c'est le travail du producteur de données, et que le modélisateur utiliserait mieux son temps à critiquer son propre travail plutôt que celui des autres...

2.6 Test préliminaire : sensibilité des performances à l'entrée de pluie horaire

Dans le but d'illustrer les propos du paragraphe précédent, nous présentons dès ce chapitre quelques résultats de modélisation au pas de temps horaire. Nous comparons quatre structures, que nous noterons de façon anonyme M_1 à M_4 car notre but n'est pas pour l'instant de les expliciter (nous le ferons au Chapitre 9). Ces quatre structures sont chacune testées avec quatre entrées de pluies détaillées ci-dessous.

2.6.1 Entrées de pluies testées

Les quatre entrées de pluie horaire utilisées pour ce test sont :

- P_{sol} , les pluies horaires données par une moyenne de pluviographes au sol (réseau peu dense),
- P_{SAFRAN} , les pluies horaires données par SAFRAN : cumuls journaliers des pluviomètres et pluviographes (réseau dense), interpolés puis désagrégés temporellement d'après un modèle d'humidité (et non des observations),
- $P_{\text{desagr.}}$, les pluies horaires de notre méthode : cumuls journaliers interpolés avec le réseau sol complet (SAFRAN), puis désagrégation temporelle avec la dynamique de la solution P_{sol} ,
- $c \cdot P_{\text{sol}}$, les pluies horaires de la solution P_{sol} multipliées par un facteur correctif (unique par bassin) permettant d'avoir le même cumul que les séries P_{SAFRAN} et $P_{\text{desagr.}}$ sur la totalité de la période 1995–2005.

2.6.2 Résultats

Sur la Figure 2.12, nous avons tracé les distributions du critère $C_{MM}(Q)$ pour les quatre modèles combinés aux quatre entrées de pluie.

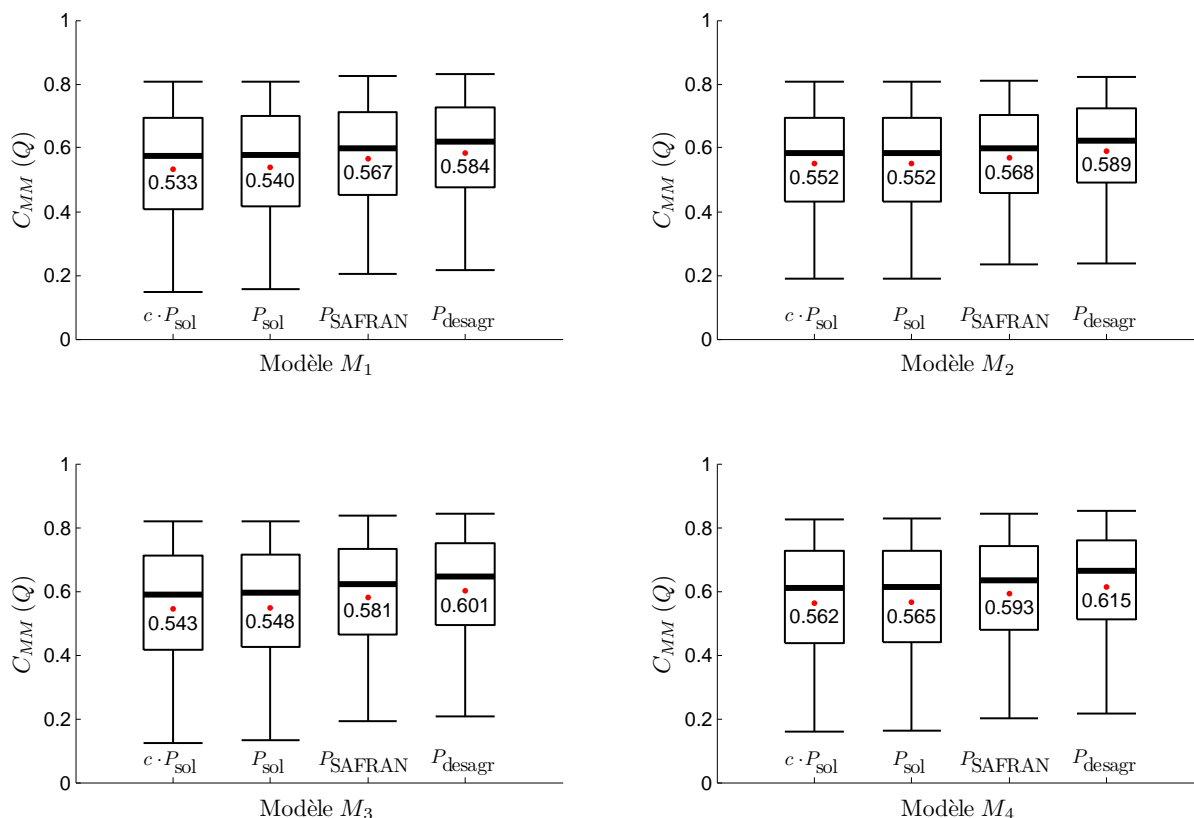


FIG. 2.12: Influence de l'entrée de pluie horaire sur les performances de quatre modèles pluie-débit. Résultats en critères C_{MM} sur le débit non-transformé Q , en contrôle sur l'échantillon de 1040 bassins versants français.

Les graphiques de la Figure 2.13 synthétisent les résultats de manière plus parlante, en trois dimensions :

- selon le premier axe horizontal, nous trouvons le type d'entrée de pluie,
- selon le second axe horizontal, nous trouvons le modèle pluie-débit,
- selon l'axe vertical, nous trouvons la performance de la combinaison des deux, exprimée par la moyenne en contrôle du critère $C_{MM}(Q)$ (graphique du haut, équivalent aux seize « boîtes à moustaches » précédentes), $C_{MM}(\sqrt{Q})$ (graphique du milieu) ou $C_{MM}(\ln Q)$ (graphique du bas).

Nous voyons clairement l'importance de la qualité des données, comparativement au choix d'un modèle : l'écart-type entre les performances des différentes entrées de pluie, à modèle donné, est de l'ordre de 0.02 à 0.025 (2 à 2.5 points, comparable pour les trois critères), tandis que l'écart-type des performances des différents modèles, à entrée de pluie fixée, est deux fois plus faible (1.3 point en moyenne, à nuancer par le fait que tous les modèles testés ici sont relativement bons). En revanche, nous constatons que la hiérarchie des modèles est assez indépendante de la qualité des données : la dégradation de cette qualité affecte globalement tous les modèles de la même façon. Nous pouvons noter trois points importants sur la Figure 2.13 :

- Le gain à la transition notée **A**, qui consiste à passer d'une moyenne de pluviographes horaires aux pluies horaires SAFRAN ; bien que la désagrégation des pluies journalières en pluies horaires dans SAFRAN soit assez rudimentaire (elle n'utilise pas de données pluviographiques horaires), nous voyons que l'estimation de cumuls journaliers corrects, grâce à la densité du réseau et à l'expertise météorologique, constitue une amélioration majeure, avec en moyenne pour les quatre modèles un gain de 2.6 points (0.026) sur le $C_{MM}(Q)$ moyen, et de 4.5 points sur le $C_{MM}(\ln Q)$ moyen,
- L'échec de la normalisation **B**, qui consiste à multiplier la moyenne des pluviographes horaires par une constante de façon à avoir la même pluie de bassin interannuelle que celle donnée par SAFRAN, montre que les biais d'échantillonnage du réseau horaire sont bien évènementiels, et non constants. On notera que le modèle M_2 (qui possède six paramètres) n'est pas concerné par cette transition (**B'**), car il utilise un coefficient correctif constant de la pluie. Nous voyons cependant que le modèle M_3 (cinq paramètres), dépourvu d'un tel coefficient, « rattrape » son retard lorsque la qualité de l'entrée de pluie augmente. Notons d'ailleurs que le coefficient correctif du modèle M_3 ne se comporte pas du tout comme un coefficient variant aléatoirement autour de 1 : il joue un rôle implicite de correction du bilan qui, dans d'autres structures, est rempli par une fonction d'échanges « souterrains »,
- Le gain pour la simulation des crues à la transition **C**, qui consiste à passer de la décomposition SAFRAN des cumuls journaliers à la décomposition donnée par la dynamique du réseau horaire. Cela permet en moyenne un gain de 2.0 points de $C_{MM}(Q)$ (du même ordre de grandeur que la transition **A**). Cette transition est sans effet sur le critère $C_{MM}(\ln Q)$: ce critère est plus sensible aux erreurs sur les faibles débits, et il semble logique que la simulation des étiages ne soit pas très affectée quand on ne fait que changer la dynamique des pluies horaires à l'intérieur de chaque journée, sans changer leur cumul.

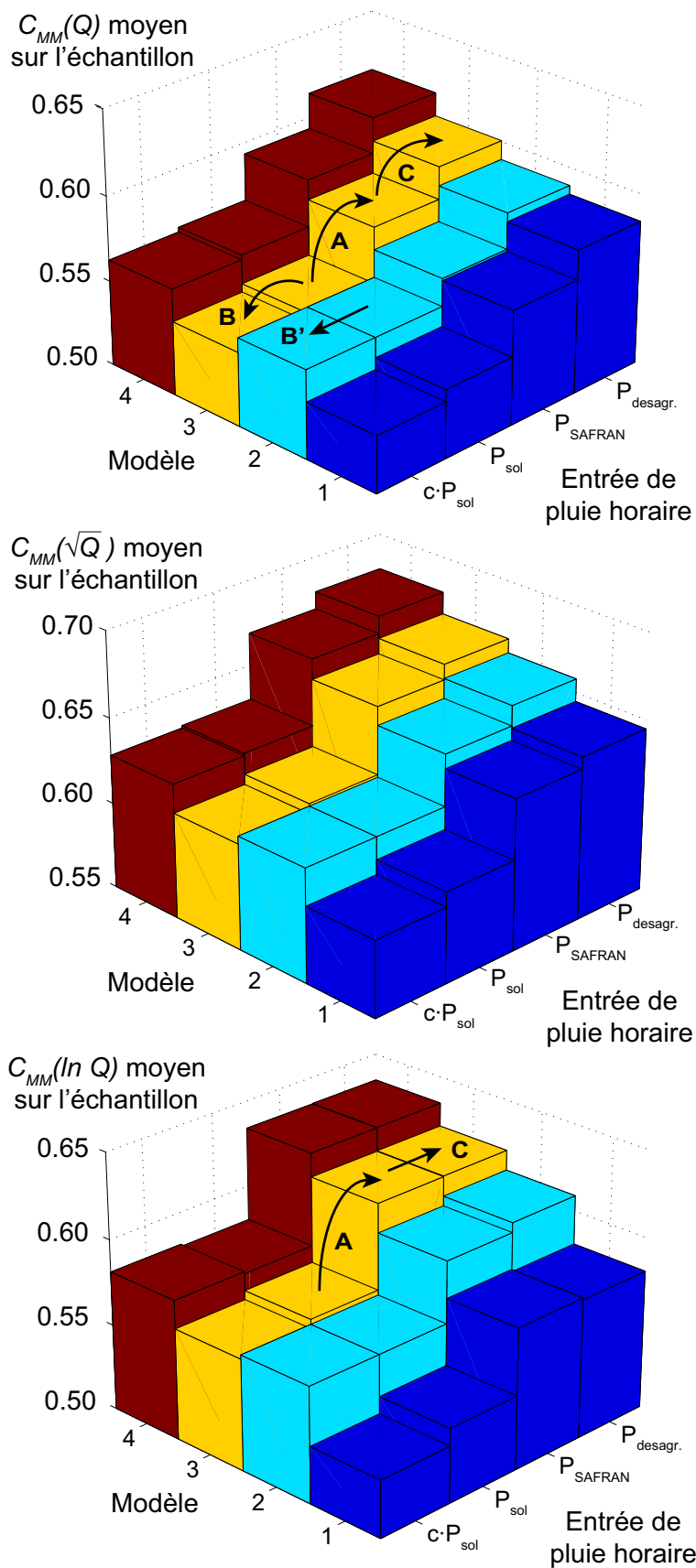


FIG. 2.13: Comparaison des performances moyennes de quatre modèles hydrologiques combinés à quatre entrées de pluie horaire, sur l'échantillon de bassins versants français.

2.6.3 Conclusion

L'étude précédente, portant sur quatre structures, quatre entrées de pluie et trois critères, permet de justifier le parti pris dans cette thèse : ne pas « trier » les données. En effet, nous pouvons constater deux choses :

- Les *performances absolues* des modèles sont indéniablement affectées par la qualité des données d'entrée : il est donc bien évidemment souhaitable d'apporter le plus grand soin à leur collecte et à leur préparation (en particulier pour l'estimation des cumuls événementiels)...
- En revanche, le *classement relatif* des modèles n'est pas notablement modifié lorsque l'on change la qualité des données. Cela s'explique par ce que nous avons dit plus haut, à savoir que les problèmes de mesures agissent statistiquement (i.e., d'un bassin à un autre) comme un bruit indépendant du signal à modéliser (la relation pluie-débit).

Finalement, travailler sur un large échantillon permet de *séparer* ce qui semblait à première vue inextricablement *lié*, à savoir la qualité du modèle, la qualité des données, et la qualité des simulations qui résultent de leur combinaison. Le progrès considérable, du point de vue du modélisateur, est donc de pouvoir se consacrer à l'amélioration des modèles *avec les données telles qu'elles existent*, sans se soucier outre mesure de les analyser en détail : il peut laisser ce travail à l'utilisateur final, qui aura à appliquer le modèle sur un bassin particulier¹.

¹Remarquons que nous pourrions appliquer un raisonnement symétrique en nous plaçant du côté du producteur de données : celui-ci peut très bien s'appuyer sur les modèles hydrologiques existants pour évaluer la pertinence de certains des traitements systématiques qu'il applique aux données brutes, sans se soucier outre mesure du caractère foncièrement « inachevé » de ces modèles...

Deuxième partie

**Comment évaluer le réalisme des flux
souterrains simulés par les modèles ?**

3

Typologie des influences souterraines sur les débits de surface

3.1 Le problème de non-superposition des bassins topographiques et hydrogéologiques

Au Chapitre 1, nous avons brièvement exposé le type de problèmes auquel est confronté l'hydrologue lorsqu'il s'agit de définir son système d'étude. Nous avons vu que cette question du *volume de contrôle* pertinent divise les hydrologues en deux écoles :

- d'un côté, les tenants d'une approche *réductionniste*, qui tend à privilégier des volumes de contrôle aussi petits que possible de façon à capturer le maximum d'hétérogénéité du système et à le décrire de la façon la plus exhaustive possible,
- de l'autre les tenants d'une approche *globale* qui ne reconnaissent comme volumes de contrôle pertinents que des bassins versants (ou éventuellement sous-bassins), privilégiant ainsi le caractère fonctionnel des entités plutôt que l'homogénéité de leurs propriétés.

Dans notre cas, le choix d'une approche globale soulève un deuxième problème passé sous silence au Chapitre 1. En effet, le cycle de l'eau ne se limite pas à sa partie superficielle (pluie, ruissellement et évapotranspiration) mais comporte bien évidemment une partie souterraine avec, dans certains cas, des transferts à grande échelle (aquifères des grands bassins sédimentaires par exemple). Le problème fondamental est que rien ne garantit la superposition des bassins versants définis d'après la topographie, avec ceux définis d'après les directions des écoulements souterrains données par le gradient piézométrique. Si la correspondance est à peu près assurée pour les très grands bassins (Seine, Loire, Garonne, Rhône, Rhin, etc.), cette hypothèse peut devenir localement très fautive pour des bassins versants de quelques dizaines à quelques centaines de kilomètres carrés.

La Figure 3.1 illustre le problème. La définition topographique du bassin versant permet d'assurer que sa frontière (ligne de crêtes) est à condition de flux (de surface) nul : on n'a bien sûr jamais vu une goutte d'eau remonter une pente. Il faut noter que, en comparaison des mailles définies dans une approche réductionniste, la donnée de cette condition aux limites est un atout considérable, d'autant plus que sur le même système nous connaissons aussi le débit de surface total (mesuré à l'exutoire) et, approximativement, la pluie. Le problème subsiste cependant, car cette condition de flux nul ne vaut que pour la *ligne* frontière et non pour la *surface* fictive que l'on obtient par « extrusion » verticale sous la surface du sol. Une partie des précipitations continentales (un quart à un tiers en moyenne mondiale, soit environ 10 000 km³/an, selon Margat, 2008) s'infiltré à travers la zone *non-saturée* (dans laquelle les flux sont essentiellement verticaux) vers la zone *saturée*, qui permet l'existence de systèmes d'écoulements locaux et régionaux (Tóth, 1963, 1995). Une partie variable de cette recharge peut être restituée de façon différée au cours d'eau (constituant ce que l'on

nomme *débit de base*) ; une autre partie peut, éventuellement, alimenter des bassins versants adjacents et / ou aval ; enfin, le reste, qui forme les écoulements souterrains à l'échelle des très grands bassins sédimentaires tels que le bassin de Paris, s'écoule vers la mer (un dixième du volume total selon *Margat, 2008*).

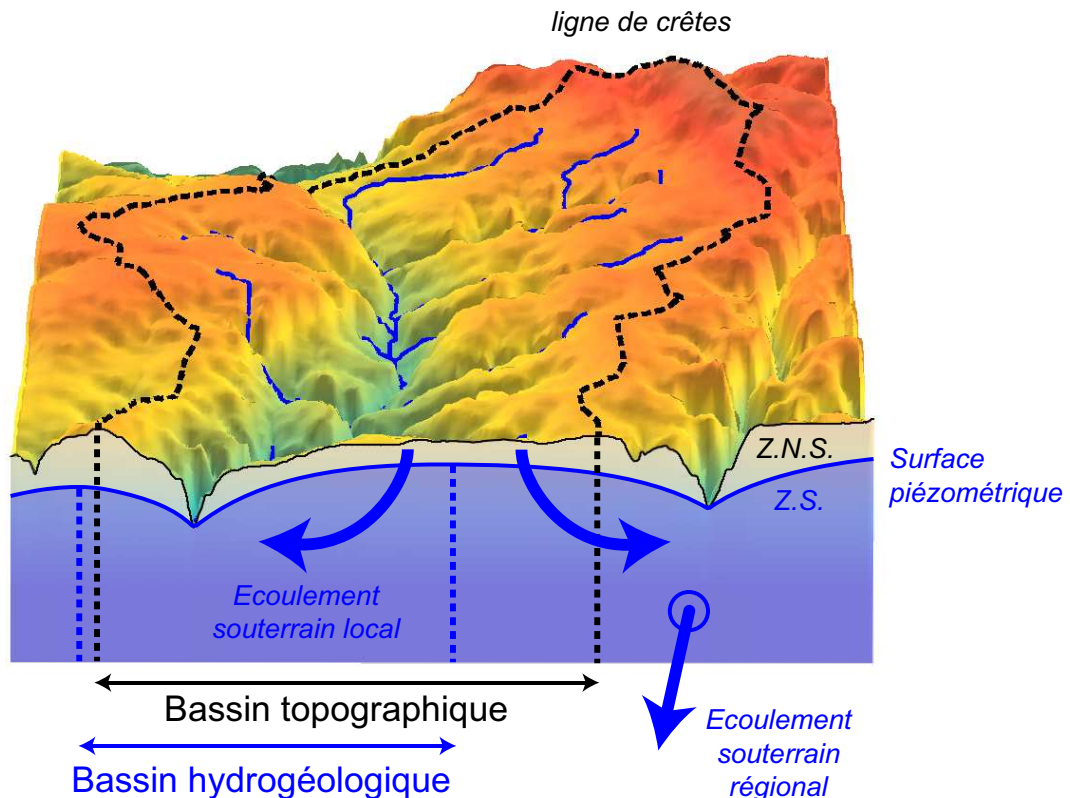


FIG. 3.1: Illustration du caractère ouvert du bassin versant topographique. Z.N.S. = zone non-saturée ; Z.S. = zone saturée.

Le problème de la non-superposition des bassins versants topographiques et hydrogéologiques se pose donc à une grande variété d'échelles : c'est un ensemble de systèmes d'écoulements « gigognes » qu'il faut imaginer, comme le montre la Figure 3.2 (le nombre de niveau n'étant bien sûr pas limité à deux !). L'écoulement de surface apparaît finalement autant comme la résultante de la topographie (ce qui, aux pas de temps qui nous concernent, est le postulat hydrologique de base) que comme sa cause, les formes de relief traduisant souvent la compétition entre les écoulements de surface et les écoulements souterrains. Celle-ci est essentiellement contrôlée par le climat et la géologie (voir par exemple *Chorley, 1964 ; Winter, 2001 ; Sanford, 2002 ; Gleeson et Manning, 2008*). Qu'on songe au contraste existant par exemple entre les régions crayeuses de la Normandie ou de la Picardie, faites de plateaux très faiblement inclinés, plus ou moins entamés par un réseau de vallées sèches et profondément incisés par les rares cours d'eau importants ; et les régions de socle telles que

le Massif Armoricain ou le Massif Central, aux vallées bien incisées et où les écoulements souterrains, sans être inexistants, sont formés de systèmes très locaux indépendants les uns des autres dont le développement dépend du degré de fracturation et d'altération (Lachassagne et Wyns, 2006).

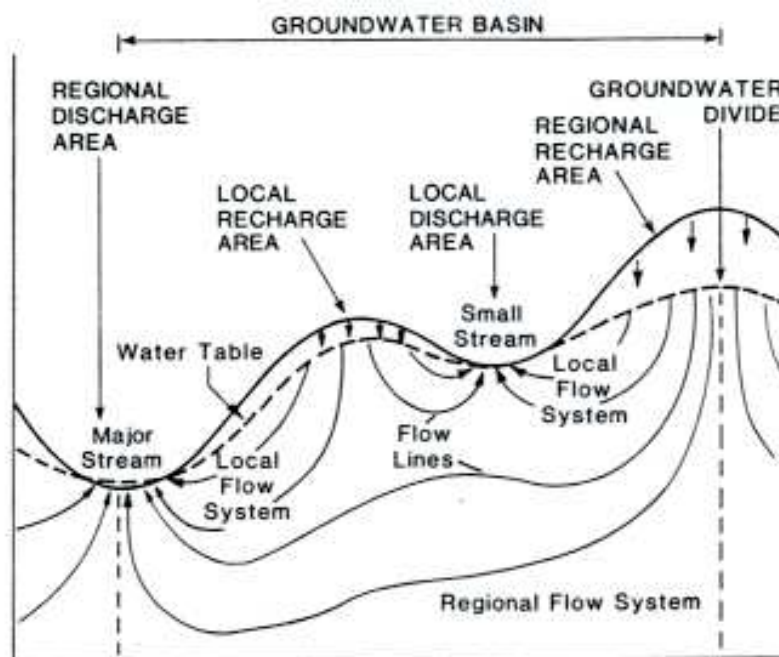


FIG. 3.2: Continuité hydraulique à l'échelle d'un grand bassin sédimentaire (Van Der Heijde, d'après Tóth, 1963). Le développement relatif des systèmes d'écoulements locaux, intermédiaires et régionaux, ainsi que la morphologie de surface, résultent d'un équilibre dont les contrôles sont essentiellement climatiques et géologiques.

Dans des formations hétérogènes, les effets locaux peuvent être encore accentués par l'existence d'écoulements préférentiels s'organisant indépendamment du réseau de surface : c'est par exemple le cas des aquifères karstiques (cf. 3.2.2.1) ou volcaniques. Ainsi, si l'on veut garder une approche générale et ne pas avoir à se poser systématiquement la question de la superposition des bassins topographiques et hydrogéologiques, il faut considérer que la masse d'eau souterraine (volume bleuté de la Figure 3.1), dans le cas où elle existe, *n'appartient pas* à notre système, et qu'elle échange avec lui.

Après ce bref aperçu, nous allons dans un premier temps examiner d'où peut provenir cette non-superposition, d'un point de vue des processus : cela constituera l'approche *descriptive* du problème. Ensuite, nous verrons quels sont les effets possibles de ces processus sur le régime des cours d'eau, c'est-à-dire leurs conséquences *fonctionnelles* sur la relation pluie-débit.

3.2 Approche descriptive

Dans cette section, nous essayons de donner un aperçu aussi exhaustif que possible des processus qui font que le bassin versant topographique est un système ouvert échangeant avec les hydrosystèmes souterrains. Nous distinguerons pour cela deux types d'échanges : les échanges au niveau des points bas du bassin versant, c'est-à-dire les échanges à l'interface nappe-rivière ; et les échanges au niveau des points hauts du bassin versant, c'est-à-dire les flux traversant les lignes de crêtes topographiques.

3.2.1 Échanges à l'interface nappe-rivière

Même dans le cas où les bassins versants topographiques et hydrogéologiques coïncident localement au niveau des lignes de crêtes, l'existence (fréquente) d'une nappe d'accompagnement de la rivière, en étroite connexion hydraulique avec celle-ci, rend nécessairement le système de surface ouvert. En effet, comme le rappellent *Mouelhi et al.* (2006), si l'on suppose par l'absurde que le flux de subsurface traversant la frontière topographique est toujours nul quand on déplace l'exutoire du bassin vers l'amont ou vers l'aval, alors le bassin versant est réduit à une tranche de matériau imperméable, et tout écoulement souterrain régional est impossible : ceci est en contradiction évidente avec l'expérience... Ceci tient au fait que les écoulements dans un milieu poreux relativement homogène, tel que des alluvions, ne sont pas chenalés, au contraire des écoulements dans un réseau hydrographique. Il existe donc toujours nécessairement une *perte* (plus ou moins importante) à l'exutoire, c'est-à-dire une composante longitudinale sortante de l'écoulement souterrain.

La Figure 3.3 illustre les différents types de connexité nappe-rivière et les contributions au débit qui en résultent.

1. Dans la situation du haut, de loin la plus fréquente en climat tempéré, la rivière draine l'hydrosystème souterrain : le cours d'eau est une condition de potentiel imposé à la nappe (condition de type Dirichlet, ou Fourier si le lit du cours d'eau est un peu colmaté). Il en résulte deux composantes à l'exutoire : une composante de perte dans le sens de l'écoulement, déjà mentionnée (notée Q_L sur le schéma de gauche), et une composante transversale d'apport à la rivière. Bien qu'ils soient dus à la configuration locale à l'interface nappe-rivière, ces flux jouent sur le bilan à l'échelle du bassin versant puisque l'écoulement souterrain est alimenté par l'infiltration sur l'ensemble du bassin. Ainsi, la composante transversale restituée à l'aval une partie de cette infiltration à l'écoulement de surface. Si, par ailleurs, le bassin versant topographique et le bassin hydrogéologique coïncident, c'est-à-dire si l'hydrosystème souterrain n'est alimenté que par l'infiltration sur le bassin topographique, la résultante de ces flux est nécessairement une perte nette, d'autant plus importante que l'aquifère

- accompagnant est transmissif et le colmatage du lit de la rivière, faible. Les affluents du Rhin, dans la plaine d'Alsace (Ill, ...) sont des exemples typiques de cette situation ;
2. La deuxième configuration traduit l'absence de connexité hydraulique, localement, entre l'aquifère et le cours d'eau. Cela peut être le résultat d'un fort colmatage du lit de la rivière, ou encore de la couverture du niveau aquifère par un niveau imperméable. Dans ce cas, il n'y a pas de composante transversale. À l'échelle du bassin, il y a de fortes chances que cette situation traduise une perte nette, l'aquifère étant généralement alimenté par l'infiltration plus à l'amont ;
 3. Le troisième schéma illustre un cas plus rare où la cote de la rivière est plus élevée que le niveau piézométrique mais où la connexité hydraulique est maintenue. Cela peut se produire lorsque l'on passe d'une situation 1 à la situation 4 décrite au point suivant ;
 4. Le schéma du bas correspond au cas d'une rivière « perchée » par rapport à la nappe : le cours d'eau « fuit » par son lit et impose à l'aquifère une condition de flux (condition aux limites de type Neumann). Cette situation est très fréquente pour des cours d'eau dont le lit est peu colmaté et l'encaissant, perméable. *Besbès (2006)* rappelle que ce type de pertes (*transmission losses* dans la terminologie anglo-saxonne) est le principal processus de recharge en zone aride, où « *l'infiltration directe des précipitations ne constitue pas une source déterminante d'apports aux nappes souterraines. Du fait d'un déficit hydrique permanent, l'infiltration n'est possible qu'à la faveur des concentrations du ruissellement dans le lit des cours d'eau* ». En Normandie crayeuse, l'Iton (affluent de l'Eure) fournit un exemple typique de décrochement par rapport à la nappe sur certaines portions de son cours (notamment le *Sec Iton*). Dans tous ces cas, les écoulements sont souvent intermittents : la rivière ne coule que lorsque le débit est supérieur à la capacité d'infiltration du lit.

Les quatre configurations précédentes peuvent bien sûr se retrouver à différentes sections d'un même cours d'eau, ou en une même section à différents moments de l'année.

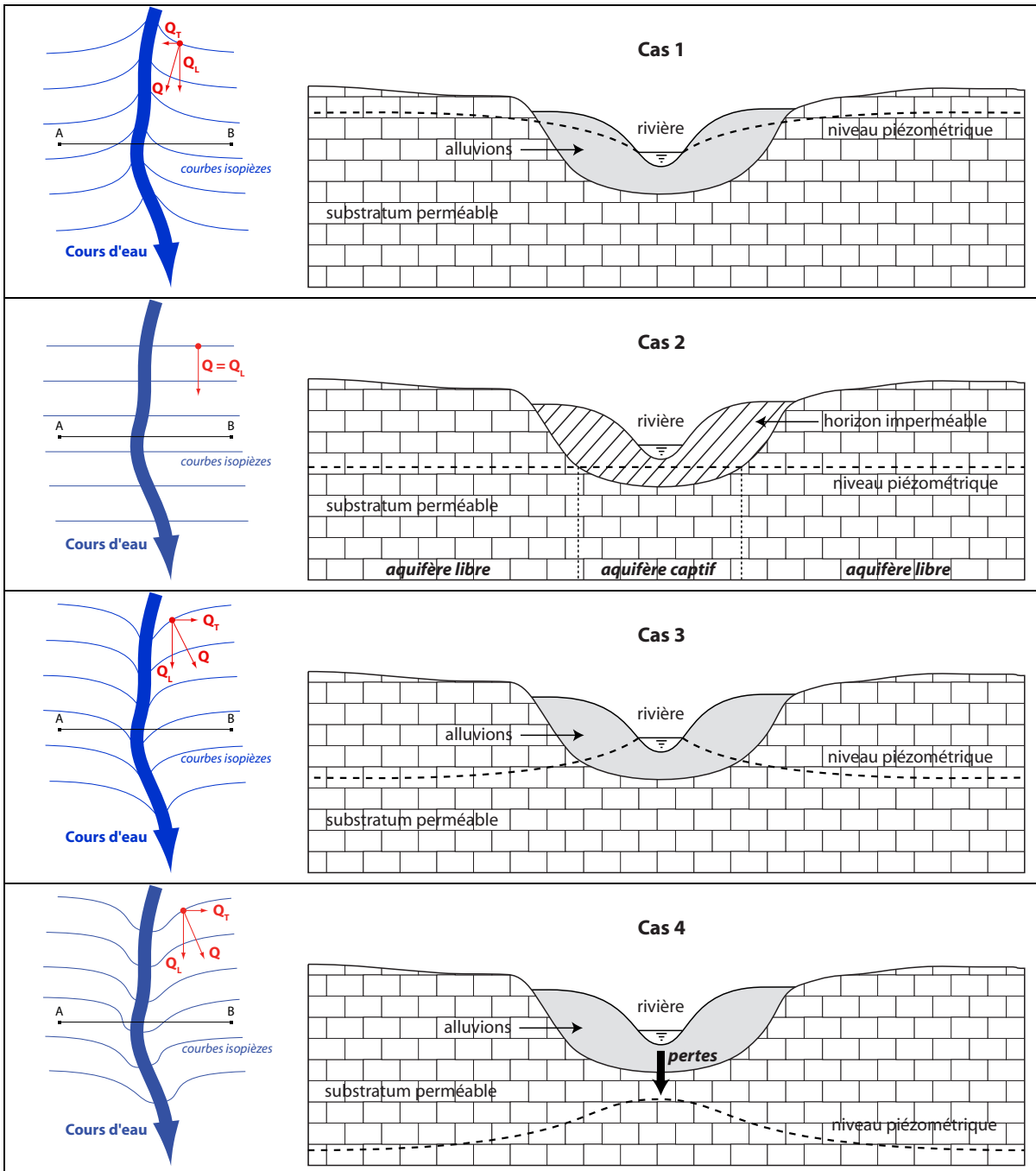


FIG. 3.3: Différents types de relations à l'interface nappe-rivière (d'après MEDD / BRGM, 2003). La colonne de gauche présente une vue dans le plan avec les courbes équipotentiels de la surface piézométrique et les composantes de flux (Q : vecteur flux total, orthogonal aux équipotentiels; Q_L composante longitudinale dans le sens d'écoulement de la rivière; Q_T composante transversale d'échange entre la nappe et la rivière). La colonne de droite est une coupe verticale transversalement à l'écoulement.

Cas 1 : Rivière drainant la nappe (limite à condition de potentiel)

Cas 2 : Absence de connexion hydraulique

Cas 3 : Rivière alimentant la nappe, avec connexion (limite à condition de potentiel)

Cas 4 : Cas d'une rivière perchée par rapport à la nappe avec pertes et recharge par la zone non-saturée (limite à condition de flux)

3.2.2 Échanges avec les bassins voisins

3.2.2.1 Pertes ou apports localisés : le milieu karstique

Dans ce paragraphe, nous considérons le cas où le système de surface échange de l'eau avec les bassins voisins par l'intermédiaire d'écoulements localisés, à travers un réseau de cavités et de véritables rivières souterraines. Cette forme d'écoulement souterrain, qui se rencontre principalement en milieu karstique, correspond à l'image que se fait souvent le grand public de la notion d'aquifère, alors qu'elle en est un cas plutôt extrême¹. La Figure 3.4 illustre les principales caractéristiques d'un système karstique. En particulier, le développement du système est contrôlé par le niveau de base : le réseau karstique, formé par dissolution de la matrice carbonatée, peut être vu comme une dérivation, un « court-circuit » de tout ou partie du réseau de surface vers ce niveau, indépendamment des frontières topographiques.

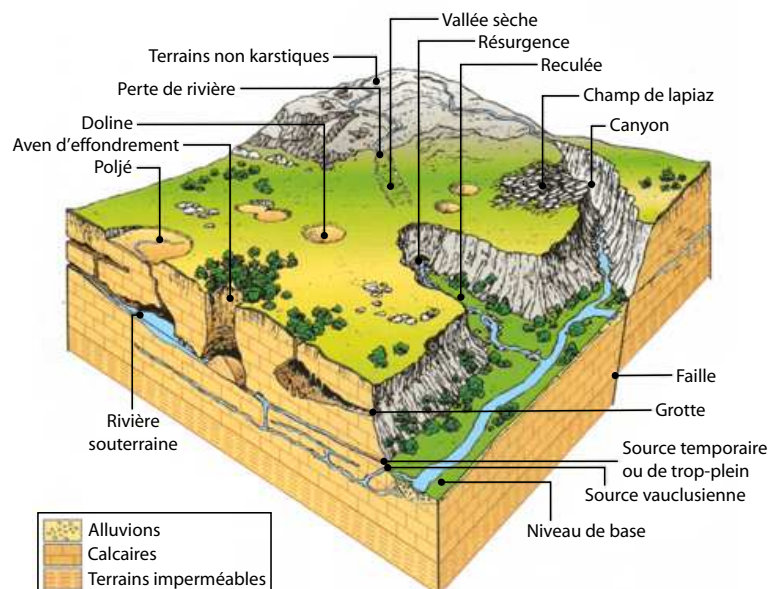


FIG. 3.4: Modèle conceptuel de karst (Bakalowicz, 1999).

Pour le sujet qui nous intéresse, une distinction mérite d'être faite entre karsts *unaires* et *binaires* :

- un karst unaire s'étend exclusivement sur des formations karstifiées, dans lesquelles se développe un réseau de conduits aboutissant à la source et drainant les zones de stockage. Les eaux recueillies sur la surface calcaire sont absorbées directement par l'intermédiaire de dolines et/ou lapiaz,
- un karst binaire inclut également un ensemble de bassins versants situés à l'amont sur des terrains non karstifiés, imperméables, mais qui aboutissent à des pertes lorsque les cours d'eau entrent en domaine karstique.

¹Il faut reconnaître que dans les autres cas, comme le souligne fort bien le Pr. Pierre Hubert, « les eaux souterraines ne sont pas très photogéniques ».

Il est souvent plus facile d'étudier le bilan hydrologique des systèmes karstiques que celui d'autres systèmes aquifères, car les écoulements souterrains y sont majoritairement concentrés dans des conduits (chenalisation analogue à celle de l'écoulement de surface), et non pas diffus dans une matrice poreuse. Ainsi, ces écoulements souterrains peuvent être jaugés lorsqu'ils réapparaissent à la surface sous forme d'un nombre limité de résurgences : le Chapitre 5 présente un cas de karst binaire où ce type de mesures peut être mis à profit pour tester expérimentalement les bilans hydrologiques simulés à l'échelle d'un bassin versant.

3.2.2.2 Pertes ou apports par écoulement diffus

Le cas que nous allons considérer maintenant est probablement le moins intuitif. Il concerne les cas où des écoulements régionaux (aquifères continus) traversent les frontières topographiques du bassin même au niveau des points hauts : c'est-à-dire les cas où les lignes de crêtes de la surface piézométrique sont complètement distinctes des lignes de crêtes de la surface topographique. La géométrie des aquifères, leurs conditions aux limites, ainsi que des variations de perméabilité dans l'espace peuvent en effet aboutir à des « solutions » du champ de potentiel (la charge hydraulique h) assez différentes de la morphologie de surface.

Dans l'hypothèse de Dupuit-Forchheimer (équipotentiels verticales), l'équation de continuité locale dans un milieu poreux en régime permanent s'obtient en écrivant que la divergence du flux ($\vec{\phi} = -T\vec{\nabla}h$ d'après la loi Darcy, où $T(x, y)$ est la transmissivité locale de l'aquifère) est égale au terme de recharge q [$L \cdot T^{-1}$] :

$$\vec{\nabla} \cdot (-T\vec{\nabla}h) = q \quad (3.1)$$

Pour une nappe libre, la transmissivité est égale au produit de la perméabilité $K(x, y)$ par la hauteur mouillée : $T(x, y) = K(x, y) \cdot (h(x, y) - z_0(x, y))$ où $z_0(x, y)$ est la cote du substratum. En supposant que ce substratum est horizontal et en prenant la cote z_0 comme origine des altitudes, l'équation de continuité précédente s'écrit :

$$\begin{aligned} \vec{\nabla} \cdot (-Kh\vec{\nabla}h) &= q \\ \vec{\nabla} \cdot \left(-\frac{1}{2}K\vec{\nabla}h^2 \right) &= q \\ \vec{\nabla} \cdot (K\vec{\nabla}h^2) &= -2q \end{aligned} \quad (3.2)$$

La Figure 3.5 montre quelques solutions analytiques de cette équation de diffusion, en régime permanent et à une dimension, obtenues en faisant varier divers facteurs. La géométrie considérée est un interfluve de largeur L (nous supposons que la frontière topographique est localisée en $x = \frac{L}{2}$, mais on ne change en rien les conclusions en la

positionnant ailleurs), constitué de deux versants A et B dont les cours d'eau imposent des conditions de charge $h(0) = h_A$ et $h(L) = h_B$. Ces exemples ont pour l'instant un caractère illustratif mais nous les réutiliserons dans les discussions du Chapitre 7.

Les facteurs contrôlant la position de la ligne de partage des eaux sont essentiellement :

1. les conditions aux limites h_A et h_B , la recharge q et la perméabilité K étant considérées comme uniformes. L'équation de continuité se résume alors à :

$$\begin{aligned}\frac{d^2 h^2}{dx^2} &= -\frac{2q}{K} \\ \frac{dh^2}{dx} &= -\frac{2q}{K}x + C \\ h^2(x) &= -\frac{q}{K}x^2 + Cx + D\end{aligned}\quad (3.3)$$

où les constantes C et D sont identifiées à l'aide des conditions aux limites. En climat humide, les conditions aux limites imposées par les cours d'eau sont généralement des charges imposées (*Winter*, 2001) mais d'autres conditions existent (cf. section 3.2.1),

2. la répartition spatiale de la recharge, résultant par exemple d'un gradient pluviométrique ou de l'hétérogénéité de l'occupation du sol. Dans le schéma 2, nous imposons une recharge linéairement décroissante entre A et B , la perméabilité K étant uniforme, conduisant à l'équation de continuité :

$$\frac{d^2 h^2}{dx^2} = -\frac{2}{K}q(x) = -\frac{2}{K}\left[(q_B - q_A)\frac{x}{L} + q_A\right]$$

3. la variabilité spatiale de la perméabilité, illustrée sur le schéma 3 par un gradient constant entre A et B sur le logarithme de la perméabilité : $\ln K(x) = (\ln K_B - \ln K_A)\frac{x}{L} + \ln K_A$, soit, en posant $\beta = \frac{1}{L} \ln\left(\frac{K_B}{K_A}\right)$: $K(x) = K_A e^{\beta x}$. L'équation de continuité s'écrit alors :

$$\begin{aligned}\vec{\nabla} \cdot (K \vec{\nabla} h^2) &= -2q \\ \frac{d}{dx} \left(K \frac{dh^2}{dx} \right) &= -2q \\ K \frac{d^2 h^2}{dx^2} + \frac{dK}{dx} \cdot \frac{dh^2}{dx} &= -2q\end{aligned}$$

En remplaçant K par son expression, on obtient :

$$e^{\beta x} \cdot \frac{d^2 h^2}{dx^2} + \beta e^{\beta x} \cdot \frac{dh^2}{dx} = -\frac{2q}{K_A}$$

équation que l'on résout en cherchant h^2 sous la forme $h^2(x) = A(x)e^{-\beta x}$.

4. l'existence de singularités ou de discontinuités hydrauliques. Dans le schéma 4, nous

considérons une fracture importante (ou éventuellement une faille) située en x_F autorisant un débit de fuite Q_F fonction de la hauteur $h(x_F)$, par exemple $Q_F = \alpha h^2(x_F)$. À gauche et à droite de la discontinuité, la charge vérifie la même équation que dans le cas 1 (Équation 3.3) :

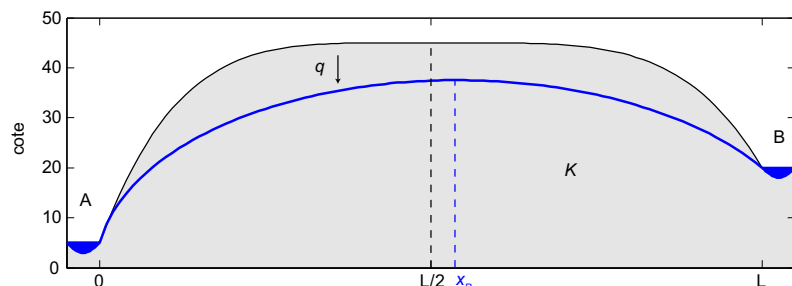
$$\begin{cases} h^2(x) = -\frac{q}{K}x^2 + C_g x + D_g & \text{pour } x \in [0, x_F], \\ h^2(x) = -\frac{q}{K}x^2 + C_d x + D_d & \text{pour } x \in [x_F, L], \end{cases}$$

Les quatre constantes C_g , D_g , C_d et D_d sont identifiées à l'aide des conditions aux limites :

$$\begin{cases} h^2(x=0) = h_A^2 \\ h^2(x=L) = h_B^2 \\ h^2(x=x_F^-) = h^2(x=x_F^+) \\ -\frac{K}{2} \frac{dh^2}{dx} \Big|_{x=x_F^-} + \frac{K}{2} \frac{dh^2}{dx} \Big|_{x=x_F^+} = Q_F(h(x_F)) \end{cases} \iff \begin{cases} D_g = h_A^2 \\ -\frac{q}{K}L^2 + C_d L + D_d = h_B^2 \\ -\frac{q}{K}x_F^2 + C_g x_F + D_g = -\frac{q}{K}x_F^2 + C_d x_F + D_d \\ \frac{K}{2} (C_d - C_g) = \alpha \left(-\frac{q}{K}x_F^2 + C_d x_F + D_d \right) \end{cases}$$

La situation 4 est typique des aquifères à double porosité, tels que celui de la Craie. À la différence des aquifères karstiques classiques (cf. section 3.2.2.1), dont la perméabilité est faible en dehors des conduits de dissolution, l'aquifère de la Craie a la particularité de développer *à la fois* des écoulements dans la matrice poreuse *et* des écoulements préférentiels dans la fracturation. Cette propriété peut provoquer des cas assez spectaculaires de non-superposition des bassins topographiques et souterrains, tels que la capture du bassin hydrogéologique d'Étretat par le bassin d'Yport, en Pays de Caux (cf. Figure 3.6 pour les conditions aux limites de la nappe de la Craie dans ce secteur). Sur la Figure 3.7, on a représenté les quatre systèmes « globaux » que l'on peut définir : on constate qu'il n'y a aucun moyen de faire coïncider l'un ou l'autre système de surface avec son « homologue » souterrain, ce qui complique considérablement l'écriture d'un bilan entrée / sortie. Pour prédire le débit des sources d'Yport, il semble évidemment préférable de considérer le système souterrain ($\mathbf{Y_P}$), la surface du bassin topographique $\mathbf{Y_T}$ étant nettement insuffisante pour expliquer ce débit (en effet, la modification de la taille de l'impluvium souterrain se répercute sur le débit à l'exutoire du bassin topographique par l'intermédiaire des flux à l'interface nappe-rivière). Mais dans ce cas, il nous faut estimer la recharge en entrée de ce système $\mathbf{Y_P}$, la zone de recharge s'étendant à cheval sur les deux bassins topographiques $\mathbf{E_T}$ et $\mathbf{Y_T}$...

(1) Influence des conditions aux limites h_A et h_B

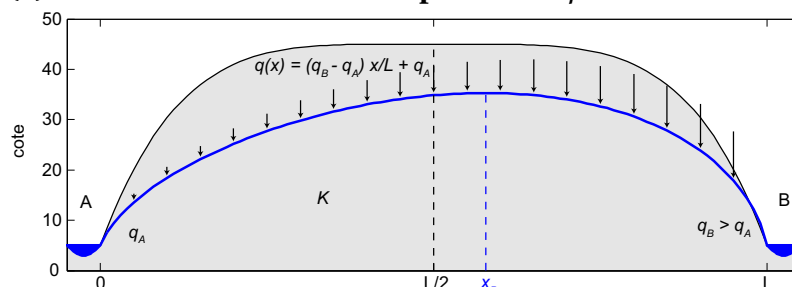


$$h(x) = \sqrt{-\frac{q}{K}x^2 + Cx + D}$$

$$D = h_A^2$$

$$C = \frac{1}{L} \left(h_B^2 - D + \frac{q}{K}L^2 \right)$$

(2) Influence de la variabilité spatiale de q

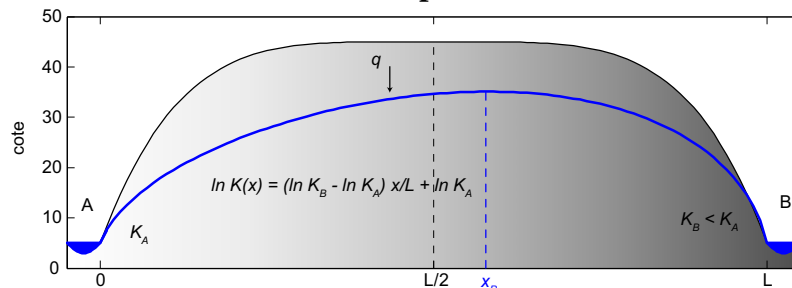


$$h(x) = \sqrt{\left[(q_A - q_B) \frac{x^3}{3KL} - \frac{q_A}{K}x^2 \right] + Cx + D}$$

$$D = h_A^2$$

$$C = \frac{1}{L} \left(h_B^2 - D + \frac{L^2}{3K} (q_B + 2q_A) \right)$$

(3) Influence de la variabilité spatiale de K



$$h(x) = \sqrt{\left[\frac{2qL}{\beta K_A}x + C \right] e^{-\beta x} + D}$$

$$C = \frac{h_B^2 - h_A^2 - \frac{2qL}{\beta K_B}}{\frac{K_A}{K_B} - 1}$$

$$D = h_A^2 - C$$

(4) Influence des discontinuités / singularités

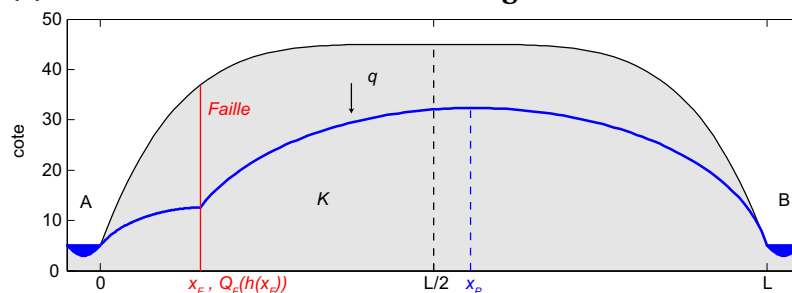


FIG. 3.5: Influence des différents facteurs sur la position de la ligne de partage des eaux souterraines dans un interfluve.

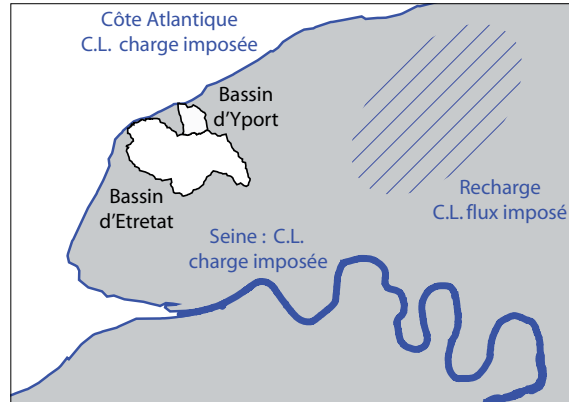


FIG. 3.6: Conditions aux limites de la nappe de la Craie au niveau du bec de Caux.

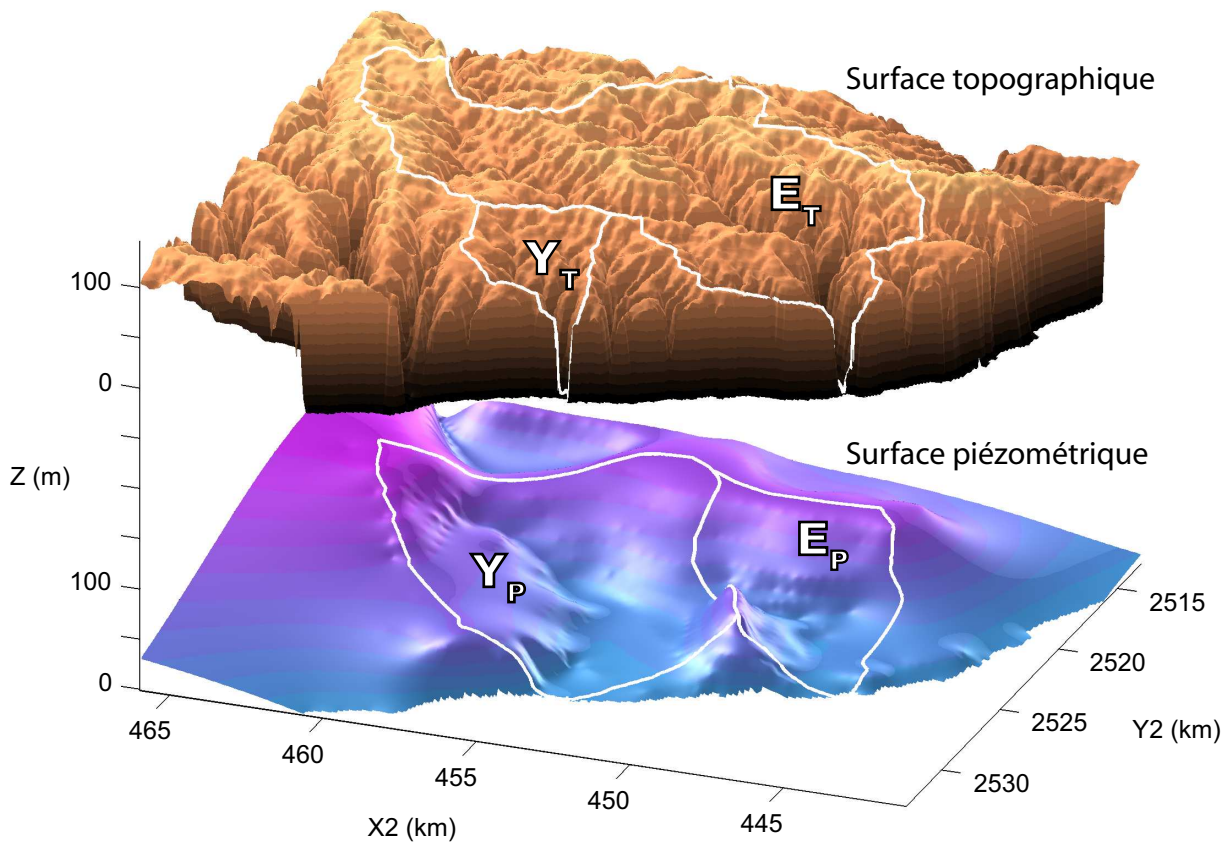


FIG. 3.7: Capture d'une partie du souterrain d'Étretat par le bassin d'Yport (carte piézométrique numérisée d'après Delu *et al.*, 1978). L'échelle verticale est exagérée 30 fois, et la topographie a été « soulevée » de 300 m pour laisser apparaître la surface piézométrique.

- Y_T : bassin topographique d'Yport (17 km^2)
- Y_P : bassin piézométrique d'Yport (104 km^2)
- E_T : bassin topographique d'Étretat (145 km^2)
- E_P : bassin piézométrique d'Étretat (46 km^2)

Outre ce problème de non-superposition avec le bassin topographique, il faut encore ajouter le fait que le bassin hydrogéologique est variable dans le temps. En effet, la surface piézométrique n'est pas stationnaire et par conséquent ses lignes de crête peuvent se déplacer au cours de l'année, comme illustré sur la Figure 3.8 (bassins versants du Pang et du Lambourn, situés dans le bassin de Londres, du côté britannique de l'aquifère de la Craie). De ce fait, l'infiltration à un endroit donné d'un bassin versant topographique n'alimente pas toujours le même bassin hydrogéologique au cours de l'année...

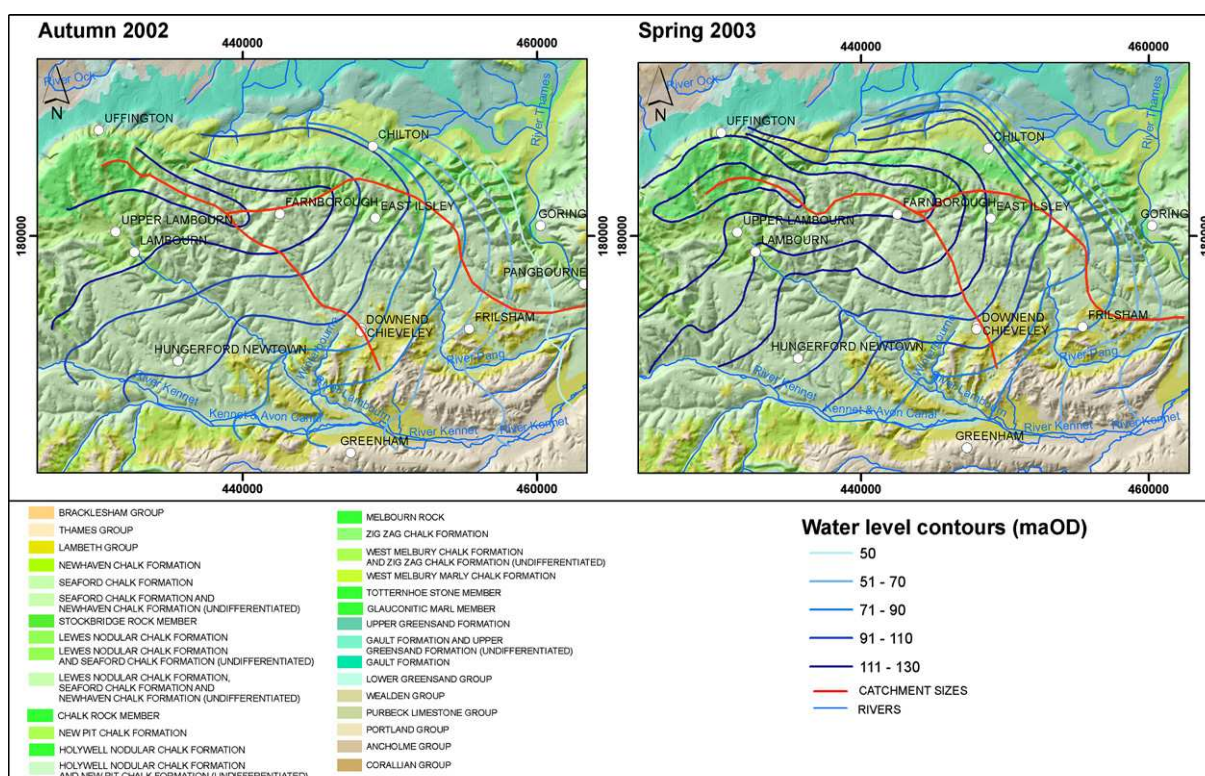


FIG. 3.8: Variabilité saisonnière des lignes de crête de la surface piézométrique (en rouge) : cas des bassins du Pang et du Lambourn (Wheater et al., 2006, d'après Peach et al., 2006, British Geological Survey. Reproduit avec l'aimable autorisation des auteurs et du journal).

C'est donc l'estimation du terme de recharge qui constitue le noeud du problème. De façon synonyme, ce terme est l'une des sorties du bassin topographique à l'issue de la transformation pluie-débit (les autres sorties étant le débit et l'évapotranspiration). Il constitue le forçage d'entrée du bassin hydrogéologique : on peut résumer cela en disant que la recharge est à l'hydrogéologue ce que la pluie est à l'hydrologue, à la différence majeure que la recharge n'est pas mesurable directement. En effet, la recharge se traduit par une variation de stock ΔS dans l'hydrosystème souterrain (exprimée par exemple en mm) mais cette variation de stock n'est pas directement observable : la grandeur qui s'en rapproche est la variation de charge (hauteur piézométrique), que l'on peut échantillonner de façon analogue au champ de pluie. Toutefois, convertir cette variation de charge en variation de stock nécessite de connaître le coefficient d'emmagasinement (ou porosité efficace pour

une nappe libre), dont l'estimation n'est pas aisée. La difficulté de la tâche croît encore lorsque l'on doit estimer une recharge à l'échelle globale, ce coefficient n'étant pas une propriété homogène de l'aquifère dans l'espace. Ainsi, la moyenne des variations de charge sur un grand nombre de piézomètres ne reflète qu'imparfaitement une variation de stock à l'échelle du bassin... L'hydrologue a plus de chance avec la pluie : une variation de 10 mm dans un seau équivaut à la même quantité d'eau quel que soit l'endroit où est posé le seau¹, et la moyenne des cumuls pluviométriques sur un grand nombre de pluviomètres commence à ressembler à une pluie de bassin. Nous reviendrons sur cette différence cruciale au Chapitre 7, lorsque nous discuterons de l'intérêt d'utiliser, entre autres, des données piézométriques pour améliorer la simulation des débits de surface².

Dans le cas des bassins d'Yport et d'Étretat, la situation est tellement caricaturale qu'elle simplifie le problème : la recharge est presque égale à la pluie nette (solde entre la pluie et l'évapotranspiration), car les écoulements de surface sont quasi-inexistants. Les vallées des deux bassins topographiques sont « sèches », la craie y étant extrêmement perméable.

3.3 Approche fonctionnelle

Dans cette section, nous allons nous placer du point de vue décrit au Chapitre 1, et étudier quelles peuvent être les conséquences des différents processus décrits précédemment sur la transformation pluie-débit. Pour cela, nous prendrons comme illustration sept bassins versants situés eux aussi en Seine-Maritime (cf. Figure 3.9) :

- six d'entre eux sont situés sur la craie : l'Andelle, le Cailly, le Commerce, le Durdent, la Ganzéville et l'Yères. Ils sont tous influencés par la nappe de la Craie ;
- le septième est situé dans le pays de Bray, région correspondant à un anticlinal érodé où des niveaux plus anciens et imperméables ont été mis à jour (marnes du Crétacé inférieur et calcaires du Jurassique supérieur). Ce bassin ne subissant que très peu, voire pas du tout, d'influences souterraines, nous le prendrons comme bassin « témoin ».

¹La porosité efficace de l'air est égale à 1, si l'on ose dire!

²Mentionnons dès à présent qu'il existe de nombreuses méthodes d'estimation de la recharge, basées sur le suivi de profils d'humidité, de température, de concentration de divers isotopes, etc. mais que la question de leurs incertitudes et de leur représentativité spatiale est encore plus problématique que dans le cas de la pluie.

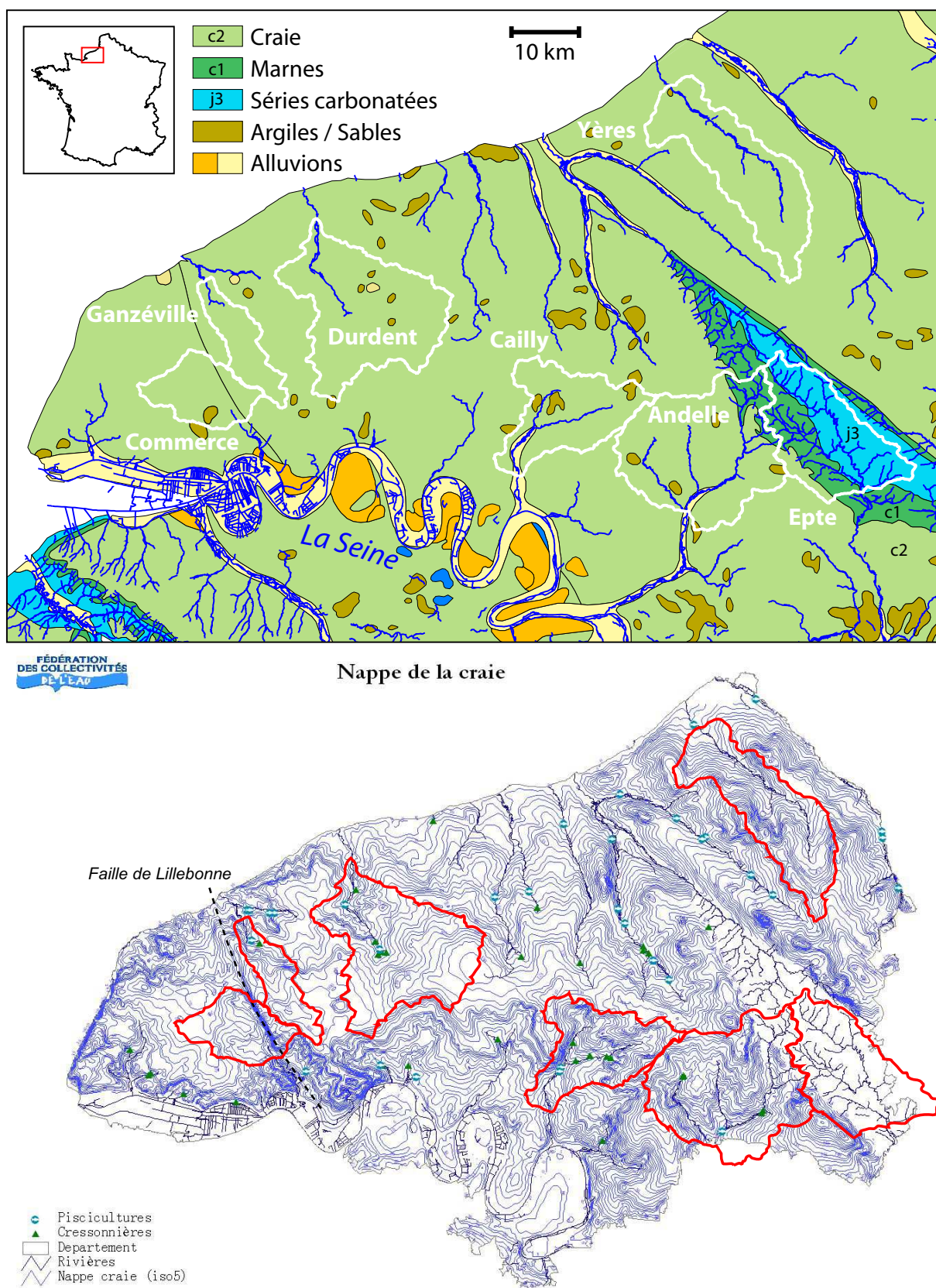


FIG. 3.9: Localisation de six bassins versants influencés par la nappe de la Craie (Andelle, Cailly, Commerce, Durdent, Ganzéville et Yères) et d'un septième bassin témoin situé sur un substratum imperméable (Epte). En haut : carte géologique simplifiée; En bas : carte piézométrique moyenne, isopièzes 5 m (source : <http://www.sidesa.fr/Upload/medias/nappecraie.jpg>).

3.3.1 Influences sur le bilan interannuel

L'indicateur le plus simple pour détecter d'éventuelles contributions souterraines au débit des bassins versants est le rapport y entre le module Q_m (i.e., le débit annuel moyen calculé sur un grand nombre d'années), exprimé en mm, et la pluie annuelle moyenne P_m s'abattant sur le bassin topographique. Si ce rapport est élevé, par exemple proche ou supérieur à 1, ce sera l'indication que la pluie seule ne peut pas expliquer le débit : à moins qu'il ne s'agisse d'erreurs de mesure, une contribution souterraine nette apparaît nécessaire. À l'inverse, si le bassin écoule une très faible part de la pluie, on suspectera une perte nette vers l'hydrosystème souterrain.

Nous pouvons préciser un petit peu plus avant ce que nous entendons par « faible » et « fort » en faisant intervenir une autre grandeur interannuelle, E_m , l'ETP de bassin. À supposer que le bassin soit conservatif, le débit écoulé se situe entre deux comportements extrêmes :

- à une extrémité, l'évapotranspiration réelle ne représente qu'une toute petite fraction de l'ETP et en conséquence le débit écoulé est égal à la pluie : $Q_m \simeq P_m$, soit $y = \frac{Q_m}{P_m} = 1$,
- à l'autre extrémité, l'évapotranspiration réelle atteint le taux potentiel E_m , la quantité d'eau restant pour l'écoulement étant alors $Q_m \simeq P_m - E_m$.

En notant $x = \frac{P_m}{E_m}$ nous obtenons donc l'encadrement suivant pour le rendement y des pluies sous l'hypothèse de conservativité (Mouelhi, 2006) :

$$1 - \frac{1}{x} \leq y \leq 1 \quad (3.4)$$

Cette méthode de représentation adimensionnelle du bilan interannuel sera abondamment employée au Chapitre 5. Les erreurs de mesures et les variations de stock mises à part, nous pouvons d'ores et déjà dire qu'à chaque fois que la valeur de y sera située nettement en dehors de cet intervalle, une influence souterraine est certaine (apport si $y > 1$, perte si $y < 1 - \frac{1}{x}$, aux problèmes de mesure près), la réciproque n'étant pas systématique.

Le Tableau 3.1 résume le bilan annuel moyen pour les sept bassins versants décrits précédemment.

Bassin versant	P_m (mm)	E_m (mm)	Q_m (mm)	y	$1 - \frac{1}{x}$
Andelle	903	631	357	0.40	0.30
Cailly	923	635	415	0.45	0.31
Commerce	978	647	97	0.10	0.34
Durdent	950	639	372	0.39	0.33
Ganzéville	969	644	298	0.31	0.34
Yères	918	627	322	0.35	0.32
Epte	899	631	244	0.27	0.30

TAB. 3.1: Termes du bilan annuel moyen pour les sept bassins de la Figure 3.9.

Pour des conditions climatiques très similaires, nous pouvons constater une grande disparité du rendement des pluies y sur ces sept bassins. L'Andelle et le Cailly ont un rendement nettement supérieur à celui de l'Epte, qui indique très certainement un apport net de la nappe à ces deux rivières. À l'inverse, le Commerce (perché sur une bonne partie de son cours) a un rendement très faible¹, inférieur à 10%, ce qui est même moins que la différence entre la pluie et l'évapotranspiration potentielle (i.e., $y \ll 1 - \frac{1}{x}$).

3.3.2 Influences sur la dynamique

Les échanges entre système de surface et système souterrain dépendent des états de chacun des deux systèmes et ne sont donc pas stationnaires, comme nous l'avons déjà noté au paragraphe 3.2.1. En conséquence, le fait que le débit spécifique moyen du Commerce soit deux fois et demi plus faible que celui de l'Epte ne signifie pas que l'hydrogramme du Commerce soit simplement une homothétie de celui de l'Epte.

Pour montrer cela, nous avons tracé sur la Figure 3.10 les hydrogrammes au pas de temps horaire pour l'Epte et pour la Ganzéville, qui ont un module et un rendement assez similaires (244 mm, 27% et 298 mm, 31% respectivement). Ces valeurs annuelles moyennes masquent la dynamique des échanges entre la nappe de la Craie et la Ganzéville : les étiages de la rivière sont massivement soutenus par l'aquifère, tandis que ses crues sont très atténuées par l'infiltration et la recharge. Il en résulte une variance des débits bien plus faible pour la Ganzéville que pour l'Epte.

Nous pouvons synthétiser ces différences de dynamique au moyen d'un graphique quantile-quantile (*QQ-plot*), représenté Figure 3.11, dont le principe est le suivant :

- on construit la courbe des débits horaires classés pour chaque bassin,

¹Pour être exact, le Commerce détient le record du plus petit module sur tout l'échantillon français. Ceci tient entre autres au fait qu'il est localisé sur la faille de Lillebonne (faille Fécamp-Rouen), qui crée une condition aux limites du type (4) de la Figure 3.5 (page 57).

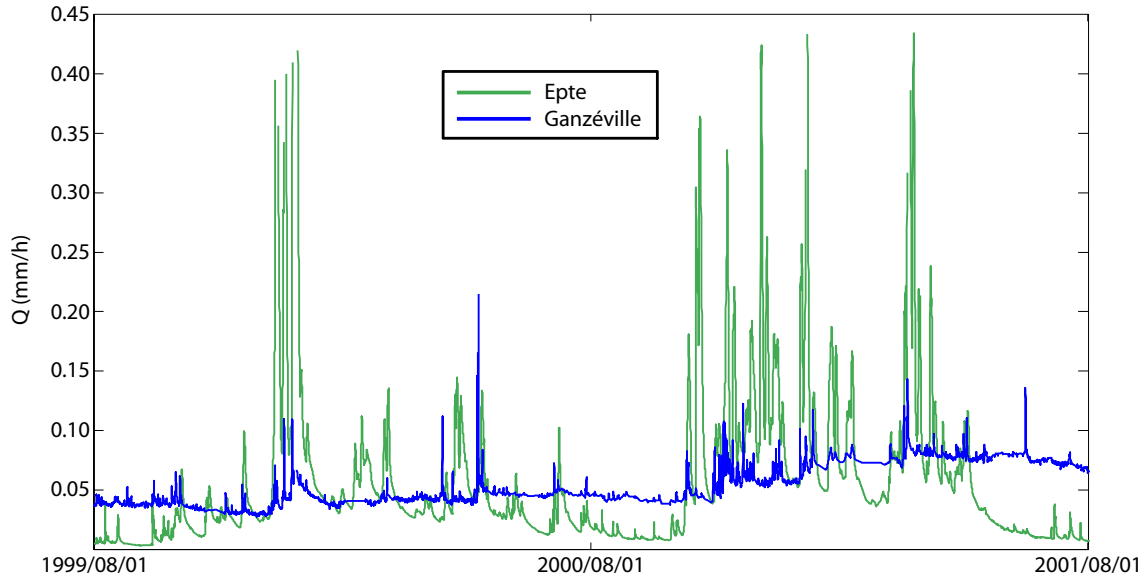


FIG. 3.10: Hydrogrammes horaires de l'Epte et de la Ganzéville sur deux années hydrologiques (1999–2000). Les débits sont exprimés en mm/h (débits spécifiques).

- on échantillonne cette courbe pour l'Epte en n valeurs de fréquences $\{f_1, \dots, f_n\}$, ce qui nous donne n valeurs de débit $\{q_1, \dots, q_n\} = X$ (série d'abscisses),
- pour chacun des six autres bassins, on trouve les quantiles de mêmes fréquences $\{f_1, \dots, f_n\}$ ce qui nous fournit six séries d'ordonnées Y_i ,
- on trace les six séries de points définies par (X, Y_i) .

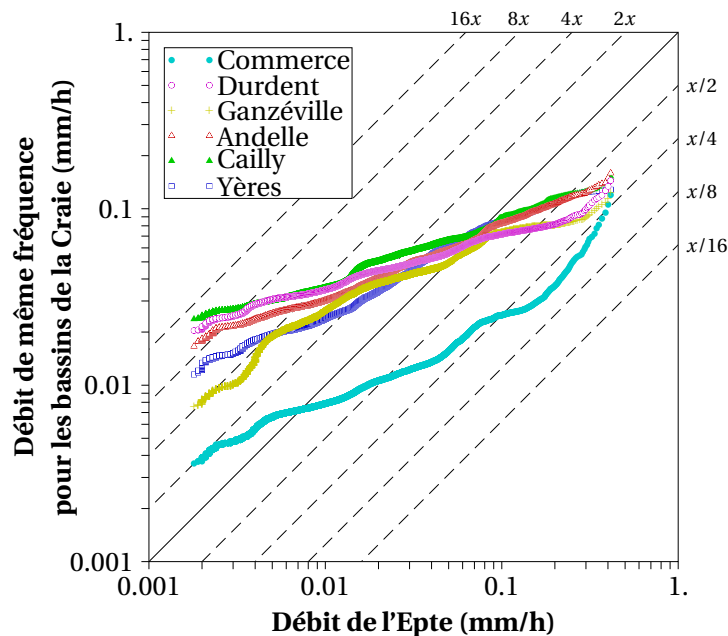


FIG. 3.11: Graphique quantile-quantile des débits horaires, avec comme référence les débits de l'Epte. Les échelles sont logarithmiques et les droites en pointillés indiquent des rapports de 2 successifs.

On observe ainsi quantitativement ce que nous avons décrit plus haut :

- À l'étiage (gauche du graphique), les bassins de la Craie ont un débit spécifique systématiquement supérieur à celui de l'Epte, même le Commerce. Malgré tout, ces débits d'étiage sont très variables d'un bassin à l'autre (de 0.005 mm/h pour le Commerce à 0.03 mm/h pour le Cailly en étiage extrême),
- En crue, les débits sur les bassins de la Craie sont nettement inférieurs à ceux de l'Epte, même ceux du Cailly (les surfaces étant suffisamment proches pour ne pas influencer notablement le temps de réaction). De plus, contrairement aux étiages, il semble y avoir un comportement convergent des six bassins de la Craie, les débits de pointe étant compris entre la moitié et le quart des débits de même fréquence sur l'Epte¹.

Ces conclusions sont principalement valables en domaine de nappe continue. En domaine karstique, les transferts, qui sont beaucoup plus rapides et représentent des volumes beaucoup plus importants, jouent souvent systématiquement dans le même sens en crue et à l'étiage. Ainsi les principales résurgences (Fontaine de Vaucluse, Touvre, Loue, etc.) sont-elles continuellement alimentées par des pertes de surface ; ces bassins de surface voient quant à eux leur débit réduit en crue comme à l'étiage (il n'est pas rare que certains cours d'eau disparaissent intégralement dans le karst en été).

Dans tous les cas, les échanges surface-souterrain affectent non seulement le débit moyen annuel mais également la dynamique des bassins versants. En outre, ils peuvent affecter toute la gamme des débits, jusqu'aux crues, et pas uniquement les débits d'étiage : cela montre qu'il n'est pas du tout absurde de travailler à un pas de temps fin (horaire) pour l'étude des bassins influencés par les eaux souterraines.

3.4 Énoncé définitif des objectifs de la thèse

Dans ce chapitre, nous avons mis en évidence de façon schématique les différents processus pouvant lier les hydrosystèmes de surface et les hydrosystèmes souterrains. Dans la réalité, ces processus se combinent tantôt en s'amplifiant, tantôt en s'annulant mutuellement. Nous nous proposons donc de repartir du schéma 1.2 du Chapitre 1, en le complétant de façon à résumer les relations surface-souterrain en un seul terme $g(t)$, résultante de tous les flux échangés entre ces deux systèmes (cf. Figure 3.12). Ce terme, que nous appellerons désormais « échange », peut être d'amplitude et de signe variable au cours du temps, $g(t) > 0$ signalant un apport net et $g(t) < 0$ une perte nette pour le système de surface.

Le bilan hydrologique s'écrira donc en termes instantanés :

$$p(t) + g(t) = q(t) + e(t) + \frac{\partial S}{\partial t} \quad (3.5)$$

¹La vérification statistique de cette tendance fait, entre autres, l'objet de l'Annexe E

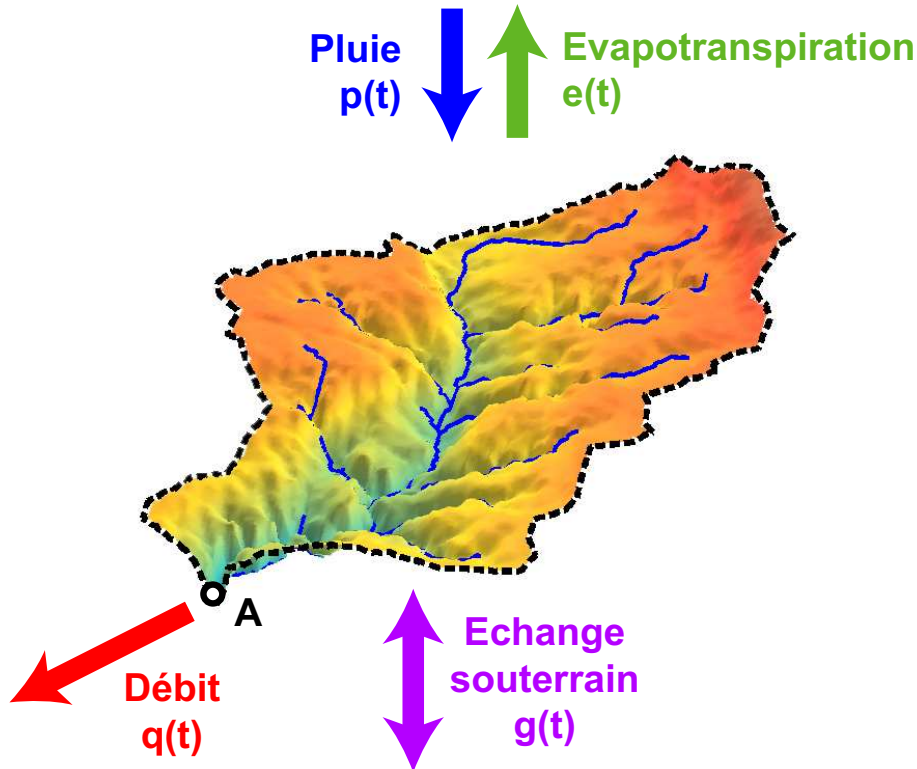


FIG. 3.12: Écriture du bilan hydrologique global sur le système de surface.

où S désigne le stock à l'échelle du bassin. Soit, en intégrant sur un pas de temps Δt :

$$P_{\Delta t} + G_{\Delta t} = Q_{\Delta t} + E_{\Delta t} + \int_t^{t+\Delta t} \frac{\partial S}{\partial t} dt \quad (3.6)$$

Nous pouvons maintenant énoncer clairement l'objectif de cette thèse :

Notre objectif est de trouver une expression fonctionnelle du terme d'échange $g(t)$, dépendant de certains paramètres et états d'un modèle représentatif du comportement du bassin versant topographique, et permettant :

- 1. d'améliorer la simulation des débits de surface,**
- 2. d'assurer autant que possible le réalisme des différents termes du bilan, et notamment la partition du déficit d'écoulement entre recharge et évapotranspiration.**

Nous chercherons à introduire empiriquement, dans des modèles simples, des modifications conceptuelles permettant de reproduire les dynamiques observées directement au niveau d'intérêt (le bassin versant). Les différentes solutions testées seront évaluées suivant la procédure décrite au Chapitre 1.

Ce nouvel énoncé doit nous faire prendre conscience que, vu la manière dont le terme d'échange g est défini, il est irrémédiablement voué à prendre *aussi* en compte, de manière

implicite, les erreurs de mesure sur les autres termes du bilan hydrologique ainsi qu'une partie des erreurs structurelles du modèle. À cette critique parfaitement valable, nous opposerons à nouveau l'argument statistique : nous estimons qu'au delà d'un certain seuil, tout progrès d'une structure sur un large échantillon de bassins versants (par exemple en terme de C_{MM} moyen, défini au Chapitre 1) doit nécessairement traduire une amélioration de la représentation du phénomène étudié (la relation pluie-débit).

Notre étude aura comme point de départ le modèle journalier GR4J. *Nascimento* (1995) a proposé une formulation de l'échange pour cette structure, constatant l'apport de cette fonction notamment pour les bassins arides. Nous repartirons donc de cette formulation en cherchant quels peuvent être les moyens d'estimer le réalisme des flux calculés (Chapitres 4 à 6). Ensuite, dans la troisième partie, nous étendrons nos tests à d'autres formulations et d'autres pas de temps, notamment le pas de temps horaire, en poursuivant le travail de *Mathevet* (2005).

4

**Comment les modèles hydrologiques
peuvent-ils prendre en compte les flux
souterrains inter-bassins ? Etude
théorique sur 1040 bassins versants**

Nicolas Le Moine, Vazken Andréassian, Charles Perrin, et Claude Michel
Article publié dans *Water Resources Research*, vol. 43, 2007

Résumé en français

L'article qui suit examine les différentes possibilités qui pourraient permettre à un modèle pluie-débit de prendre en compte l'existence de flux souterrains inter-bassins sur un bassin versant donné. Sur la base de l'échantillon présenté au chapitre 2, nous comparons quatre versions des modèles hydrologiques journaliers GR4J et SMAR qui diffèrent dans leur façon d'utiliser l'un de leurs paramètres pour ajuster le bilan en eau à l'échelle du bassin versant. Nous montrons que, du point de vue du réalisme hydrologique aussi bien que de celui de l'efficacité, il est préférable de représenter explicitement les transferts souterrains. Les solutions de substitution (facteurs correctifs des entrées climatiques ou de la surface) parfois utilisées en pratique peuvent se justifier du point de vue de la seule efficacité, mais nous montrons qu'elles devraient être évitées car elles conduisent à des corrections manifestement irréalistes et par conséquent à une partition irréaliste entre le débit écoulé, l'évapotranspiration et les flux souterrains.



How can rainfall-runoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments

Nicolas Le Moine,¹ Vazken Andréassian,¹ Charles Perrin,¹ and Claude Michel¹

Received 10 October 2006; revised 3 January 2007; accepted 14 March 2007; published 29 June 2007.

[1] This paper examines the possible solutions that may allow a rainfall-runoff model to cope with the existence of unknown intercatchment groundwater flows over a given catchment. On the basis of a large catchment set we compare four versions of the GR4J and the SMAR rainfall-runoff models that differ in the way they use one of their parameters to adjust catchment-scale water balance. We show that from both the hydrological likelihood and the modeling efficiency point of view it is preferable to explicitly represent intercatchment groundwater transfers. The surrogate corrective solutions tested in this paper (correcting or scaling factors applied to the climatic input data or to the catchment area) that are sometimes used in practice could be used on the sole grounds of streamflow simulation efficiency, but we show that they should be avoided since they may lead to obviously unrealistic corrections and consequently yield a similarly unrealistic distribution between evaporation streamflow and underground fluxes.

Citation: Le Moine, N., V. Andréassian, C. Perrin, and C. Michel (2007), How can rainfall-runoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments, *Water Resour. Res.*, 43, W06428, doi:10.1029/2006WR005608.

1. Introduction

1.1. Why Should General Model Structures Be the Goal?

[2] Rainfall-runoff (RR) models can either be considered as purely research objects or as tools for hydrological analysis and engineering applications. Here we look at RR models from the second point of view, one where models must be as robust and generic as possible because often the time and input-data resources available to apply them are scarce. We strongly believe in the need for generality, a characteristic promoted by the father of rainfall-runoff modeling himself, *Linsley* [1982]. Alluding to the great variety of hydrological processes which do control the rainfall-runoff relationship over different catchments, he objects that “these differences do not mean that a single model cannot be applied in all cases” (p. 14) since “a new model for every application would eliminate the opportunity for learning that comes with repeated applications of the same model” (p. 15). Practicing hydrologists (and, in our opinion, research hydrologists as well) cannot afford to learn or build a new model for each new catchment. As all models are imperfect representations of reality, the knowledge of the imperfections is one of the prerequisites of a conscious model use. Hydrologists need generic model structures which can be used and calibrated over a wide range of catchments.

1.2. What Can Be Done About “Leaking” Catchments?

[3] One of the still unsolved problems of RR modeling is accounting for underground catchment losses. Even if college diagrams provide us with an idealized image of catchments (on these diagrams they always happen to be lying on an impervious horizon), we all know that surface catchments may either leak or gain water originating from outside the basin. These underground fluxes have been given different names by different authors: groundwater exchanges, leaks, artesian inflow, underground outflow, etc., and catchments have been called leaking, artesian, or nonconservative. For the sake of clarity we will adopt the following convention: Underground fluxes crossing topographic boundaries will be called intercatchment groundwater flow (IGF) [after *Eakin*, 1966], and associated leaking/gaining topographic catchments will be called non-conservative (NC) [*Goswami and O'Connor*, 2005].

[4] There are many well-known physical reasons for the existence of NC catchments. Karstification of limestone is perhaps the most widespread reason worldwide; however, similar transfer may exist in volcanic substrata as well as in chalk horizons. Many examples of large karstic systems have been described in the literature [*Margat*, 2006]. However, the real problem hydrologists face is not related to these large systems but rather to the great number of smaller, quasi-unnoticeable karstic or faulting systems which may be present over many natural catchments. Here the question becomes, Can RR models deal with such local groundwater drainage systems, possibly identify them, and perhaps even quantify the associated intercatchment groundwater flow?

[5] This is certainly a major challenge for hydrologic analysis: Catchments are systems with too many unknowns

¹Hydrosystems and Bioprocesses Research Unit, Cemagref, Antony, France.

for too few equations, where measurable inputs (precipitation) and measurable outputs (streamflow) are known with uncertainty, while two sources of output fluxes (an upward evaporation flux and a downward “leakage” flux) are mostly unobservable.

[6] In this paper we discuss the possible options to account for intercatchment groundwater flow. To test these options, we used a large set of natural, unregulated catchments located in France for which daily precipitation and streamflow data were available over the 1995–2005 time period. Our aim was to test different modeling options with no a priori knowledge (i.e., in the situation of a hydrologist with no knowledge of whether or not a significant intercatchment flow affects the system under investigation). In section 2 we present a literature review summarizing the methods which have been proposed to deal with nonconservative catchments. Then, in section 3 we provide a quick overview of the catchment set characteristics, and we present the RR models used in this study (namely, the GR4J and SMAR models). In section 4 we compare the efficiency of the different variants of the models, and in section 5 we discuss the physical likelihood of corrective parameter values. In section 6 we summarize our findings.

2. Dealing With Unobservable Flows at the Catchment Scale: A Commented Literature Review

2.1. Do Leaking Catchments Actually Exist?

[7] Many catchment modelers like to present surface (i.e., topographic) catchments as watertight systems. Under this assumption, catchment water balance can be written in instantaneous terms:

$$q = p - e - \frac{dS}{dt}, \quad (1)$$

where q is runoff rate, p is precipitation rate, e is evaporation rate, and S is catchment moisture storage. After integration over time step T this equation becomes

$$\frac{1}{T} \int_0^T q \cdot dt = \frac{1}{T} \int_0^T p \cdot dt - \frac{1}{T} \int_0^T e \cdot dt - \frac{S_{\text{fin}} - S_{\text{ini}}}{T}. \quad (2)$$

[8] Obviously, when T is large, the difference in catchment storage (last term of equation (2)) will become negligible, and mean annual runoff Q will be seen as the difference between mean annual precipitation (P) and mean actual evapotranspiration (AE), which can be written as a fraction of mean annual potential evapotranspiration (PE):

$$Q = P - \text{AE} = P - \alpha \text{PE} \quad \alpha < 1. \quad (3)$$

Figure 1 provides a nondimensional representation ($x = P/\text{PE}$, $y = Q/P$) where natural catchments should be limited to a narrow “physical” range, delimited by an upper limit where $Q = P$ (i.e., $y = Q/P = 1$) and a lower limit where $Q = P - \text{PE}$ ($Q/P = 1 - \text{PE}/P \iff y = 1 - 1/x$). Of course, the actual physical range is even narrower than that, and the catchments that are within the above limits but very close to one of them still have a quite unrealistic water balance. In

Figure 1, when we place the 1040 points corresponding to the catchment set used in this paper (see section 3), we see that not all the catchments are located within these physical limits (i.e., Figure 1a, middle strip). Two explanations are possible: (1) These catchments have measurement problems, and/or (2) the watertight substratum assumption does not hold (against outflows as well as inflows).

[9] On this topic, *Beven* [2001], p. 1655 argued that

We cannot currently close the water balance by measurement. Traditionally, there was no direct way of measuring actual evapotranspiration, so errors in the long term measured water balances tended to be assigned to the evapotranspiration term, despite the fact that we know that rainfall inputs, discharge outputs and changes of storage are not always accurately measured. There is still no way of checking whether the catchment is indeed watertight. The continuity equation is the most fundamental law in hydrology, but as a hypothesis it would appear that we cannot currently verify it at the catchment scale.

Indeed, theoretical developments and field studies tend to question the “watertight substratum” hypothesis in many physiographical settings. There is documented evidence for intercatchment groundwater flows in many environments, on the basis of water or solute balance discrepancies, which have already been extensively reviewed by *Genereux and Jordan* [2005].

[10] For the outliers in Figure 1, classical catchment moisture accounting tools using atmospheric forcings will surely fail to model runoff at the outlet. Some hydrologists consider these as mere “pathological cases” of minor importance, usually preferring to put the blame for the balance closure difficulty on uncertain precipitation and streamflow measurements and finally discarding them. However, we think that these cases are the rule rather than the exception: In many geological settings, hydraulic connections with lower geological horizons or neighboring catchments do exist.

[11] Hydraulic connection through continuous porous media can lead to the development of regional groundwater flows and diffuse leakage from surface catchments. Theoretical analysis by *Tóth* [1963] shows that even in homogeneous continuous permeable medium, appropriate topographic settings will allow regional flow systems to develop and cross out local topographic boundaries, depending on groundwater drainage patterns. *Konrad* [2006] shows that exchanges between shallow aquifers and streams may be an important feature of the water balance in sedimentary basins.

[12] Connection can also occur through discrete structures such as fractures in crystalline bedrock or chalk and solution sinkholes and channels in karstified limestone [*Margat*, 2006]. *Thyne et al.* [1999] show that similarly, igneous and metamorphic rocks cannot be considered as flow barriers where extended faulting and fracturation occurs. Tectonic discontinuities could allow ascending or descending drainage by connecting different confined bedrock aquifers [*Carillo-Rivera et al.*, 1996; *Carillo-Rivera*, 2000] as well as karstic aquifers [*Hudson and Mott*, 1997].

[13] The above review can be summarized in a few words that may sound trivial to hydrogeologists: In all cases where either continuous or discrete permeable pathways connect the catchment’s saturated domain with lower/higher hydraulic head regions, substantial groundwater loss/gain is likely to occur and affect the water balance of the topographic

W06428

LE MOINE ET AL.: DEALING WITH IGF IN RAINFALL-RUNOFF MODE

W06428

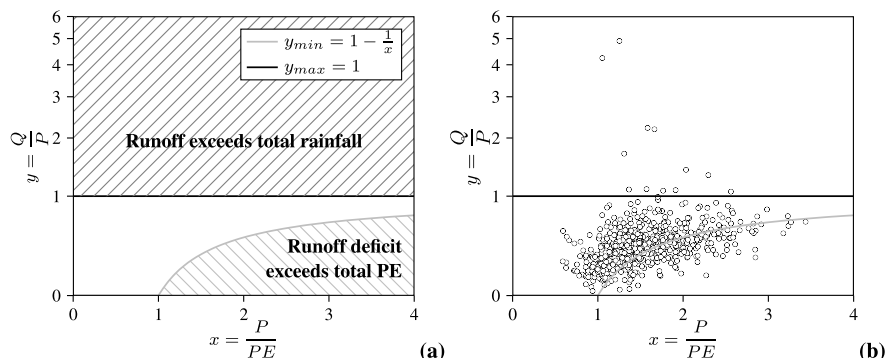


Figure 1. (a) Nondimensional graphical representation of catchments described by their mean streamflow Q , mean precipitation P , and mean potential evapotranspiration PE . (b) Presentation of the 1040 catchments of our set in the nondimensional space.

catchment. We insist on the fact that nonconservative behaviors can occur in very different places, even in the absence of large aquifers. Although the causes for the leaks are very different, karstic-type behavior can be observed both in limestone and in fractured, impermeable, and insoluble rocks, for example.

2.2. Attempts to Take Into Account Catchment Leakages Within Rainfall-Runoff Models

[14] Recognizing that surface catchments cannot be considered a priori as watertight systems, some model authors have attempted to take this fact into account in the structure of their rainfall-runoff models and to introduce functions meant to represent intercatchment groundwater flows.

[15] 1. The most detailed discussion on this topic is probably the one on GR2M monthly rainfall-runoff model by *Mouelhi et al.* [2006]. The authors discuss at length the necessity of the model’s underground exchange parameter and justify why such a parameter cannot be seen as a mere “fudge factor.” By comparing many different versions of GR2M and other alternative monthly RR models they show that such a parameter is, in fact, essential to represent the RR relationship.

[16] 2. The GR4J daily RR model presented by *Perrin et al.* [2003] also has an IGF formulation that is based on the exploratory work of *Nascimento* [1995], who worked specifically on nonperennial streams and found that an explicit formulation of IGF was both essential to model them and beneficial for perennial streams.

[17] 3. The present version of the *Pitman* [1973] model developed by *Hughes* [2004] uses a slightly more complex representation of the catchment, with local and regional head gradients and lumped Darcy formulations, allowing for leaks or gains.

[18] 4. *Ivkovic et al.* [2005] developed a version of the IHACRES model in order to take into account groundwater leakage and extractions.

[19] 5. *Goswami and O’Connor* [2005] worked on specific versions of the SMAR model, namely, SMAR-NC1 and SMAR-NC2, in order to model nonconservative behaviors. They showed that the addition of one parameter after the production function efficiently represents the rainfall-runoff relationship of otherwise anomalous catchments.

[20] 6. The MODHYDROLOG model [*Porter and McMahon, 1976*] has a deep seepage function that allows the computation of a groundwater recharge term [*Chiew and McMahon, 1990*].

3. Catchment Set, Model, and Method

3.1. Catchment Set

[21] We used a database of 1040 French catchments located throughout France (see map in Figure 2 and characteristics in Table 1) with daily rainfall, runoff, and Penman PE data over the 1995–2005 period. Figure 1 indicates that the P/PE ratio (long-term annual rainfall over annual PE) ranges from 0.4 to 3.6, which shows a rather wide climatic diversity. Physiographical settings are also quite diverse since the database contains small, Mediterranean catchments as well as larger, groundwater-dominated

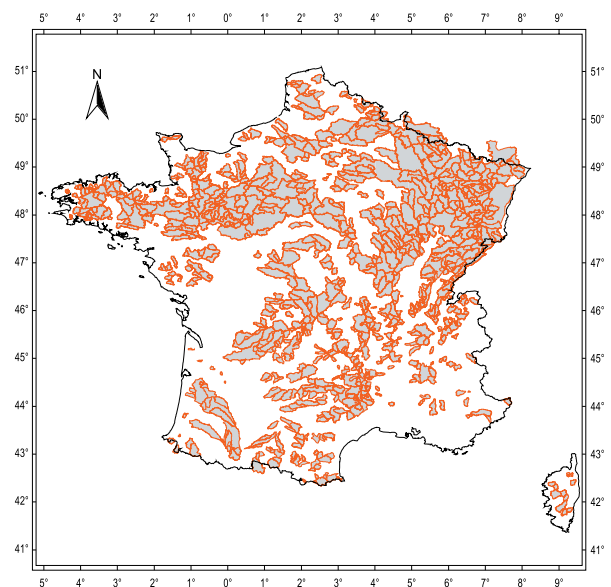


Figure 2. Location of the 1040 French catchments used in this study.

Table 1. Characteristics of the Catchment Set Used in the Study

Characteristic	Value
Number of catchments	1040
Average catchment area, km ²	480
Median catchment area, km ²	160
Maximum catchment area, km ²	9390
Minimum catchment area, km ²	10

basins or highland catchments. However, catchments where snow has a major hydrological role were excluded.

3.2. GR4J and SMAR Rainfall-Runoff Models and Their Variants

3.2.1. GR4J Model

[22] The four-parameter, daily time step GR4J model has been developed at Cemagref since the 1980s [Perrin et al., 2003] and tested over large sets of catchments. In the current standard version of the GR4J model one of the four calibrated parameters is meant specifically to account for intercatchment groundwater flows (Figure 3). The daily IGF is calculated from the storage ratio in the routing reservoir with the function

$$F = X_2(R/X_3)^{7/2}, \tag{4}$$

where R is the current level of the routing reservoir, X_3 is its capacity, and X_2 is a magnitude parameter for IGF (all in millimeters). X_2 can be either positive for water gains or negative for water losses. F is added (or subtracted) to both flow components so that the total IGF is

$$\text{IGF} = \begin{cases} 2F & \text{if } F > 0 \\ -\text{Min}(|F|; R) - \text{Min}(|F|; Q_1) & \text{if } F < 0 \end{cases} \tag{5}$$

The meaning of GR4J parameters is given in Table 2.

3.2.2. SMAR Model

[23] In addition to GR4J we used several models in order to assess the generality of our conclusions. For reasons of space we will only present SMAR. It is particularly relevant to this study as it is a well-known RR model [see, e.g., O’Connell et al., 1970; Gan and Biftu, 1996], significantly more complex than GR4J, and for which a new formulation was recently proposed in order to account specifically for IGF [Goswami and O’Connor, 2005; Goswami, 2006]. A sketch of SMAR is provided in Figure 4, and the reader is invited to refer to Goswami and O’Connor [2005] and Goswami [2006] for more details.

3.3. Can We Model Unobservable Fluxes and Still Avoid Introducing “Fudge Factors”?

[24] The main problem we face when trying to identify IGF on the sole basis of precipitation and streamflow observations is that we cannot do much more than suspect its existence when our model fails to provide an adequate calibration or when, in the nondimensional space in Figure 1, our catchment falls outside the physically realistic strip. However, even when the catchment is an obvious outlier, this does not help us to guess what type of physical fluxes the discharge deficit represents. Is evapotranspiration and/or

precipitation underestimated or overestimated? Has the topographic catchment been wrongly delimited? Is it leaking?

3.4. Methodology

[25] To investigate answers to the above questions, in this paper we propose to replace the existing one-parameter intercatchment groundwater flow function in the GR4J and SMAR models with several types of input data corrections: The conceptual IGF function is shut down, and the available free parameter is used for calibrating a correction or scaling factor of one of the input data, namely, catchment rainfall P , potential evapotranspiration PE, or discharge Q (the latter can also be seen as a correction of surface area A , scaling the discharge by α being the same as scaling the impluvium area by $1/\alpha$ from a numerical point of view). Table 3 summarizes the variants for the GR4J model that were also applied to the SMAR model.

[26] Our comparison will consist of the following: (1) checking how model efficiency can be affected when a parameter is dedicated to the scaling or the correction of input data, rather than to the explicit representation of IGF (section 4) and (2) checking the physical realism of each solution. For the physical realism we will look at the magnitude and distributions of corrective factors. We will also show that geological determinants exist and make an explicit accounting for IGF preferable from a physical point of view (section 5).

[27] The purpose of these tests is to answer the following two questions: (1) Does the explicit representation of intercatchment groundwater flows improve runoff simulation?

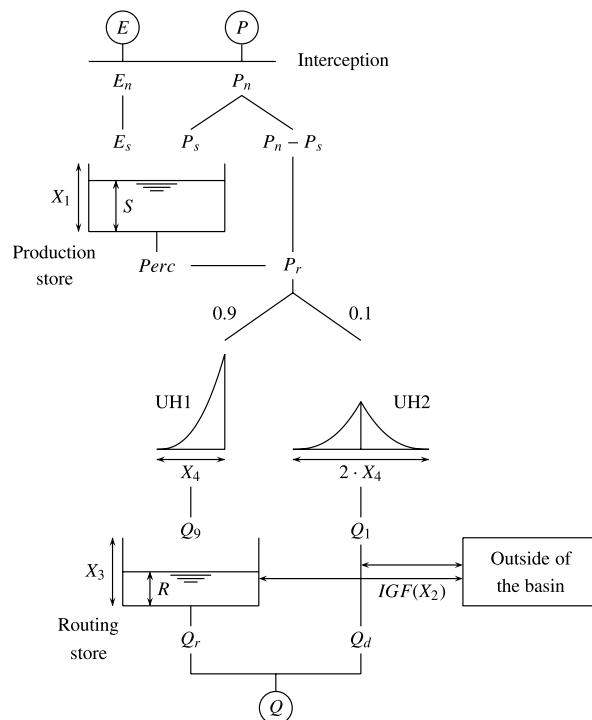


Figure 3. Structure of the GR4J daily time step model.

Table 2. List of the Parameters of the GR4J Rainfall-Runoff Model and Their Meaning

Parameter	Definition
X_1 , mm	capacity of the production reservoir
X_2 , mm	intercatchment groundwater flow magnitude
X_3 , mm	capacity of the nonlinear routing reservoir
X_4 , days	unit hydrograph time base

(2) How likely are the alternative solutions (mainly the distribution of scaling/correcting parameters)?

4. Comparison of Four Different Options in the GR4J Model

4.1. Distributions of Model Efficiency

[28] We applied the split-sample test scheme [Klemesš, 1986] and successively calibrated the model on the 1995–2000 and 2000–2005 periods with a test in validation mode on the other period each time. Some 2080 control tests were performed; items 1–3 below are based on these control results. For the sake of brevity we only present the results with the Nash-Sutcliffe criterion computed on \sqrt{Q} , which was shown by Oudin *et al.* [2006] to be a good compromise between several alternative criteria. Figure 5 shows the following.

[29] 1. Overall, on the basis of the large catchment set used here the best solution consists in devoting the fourth degree of freedom of the GR4J structure to an explicit representation of the IGF. At the median of the distribution it beats the next best four-parameter solution by 0.6 points (a small but not insignificant difference given the size of the

data set) and the three-parameter (watertight) solution by 8.0 points.

[30] 2. Moreover, the distribution of the model explicitly representing the IGF is close to the best solution curve (obtained by selecting, on each of the 1040 catchments, the best solution among the four). This shows that it is not only the best overall but also the best most of the time (out of 755 of the 2080 control periods).

[31] 3. However, even if we believe that the difference in efficiency between the explicit IGF-accounting solution and the competing variants is significant given the size of our data set, we must acknowledge that the variants involving a correcting/scaling parameter still manage to provide good results. This is the reason why we will investigate, in section 5, the hydrological likelihood of the correcting/scaling parameters.

4.2. Visualization of the Results in the Nondimensional Space

[32] In this section we present the a posteriori mapping of the interannual water balance in the nondimensional space, where we represent all catchments using, respectively, (1) $x = P_{\text{corr}}/PE$ and $y = Q/P_{\text{corr}}$ when we used a rainfall correction; (2) $x = P/PE_{\text{corr}}$ and $y = Q/P$ when we used a PE correction; (3) $x = P/PE$ and $y = Q_{\text{corr}}/P$ when we used a surface correction (or specific discharge correction); and (4) $x = P/PE$ and $y = (Q - IGF)/P$ when we computed an IGF term using uncorrected P , PE , and Q . We will compare the shape of the clusters obtained in these plots with a generalization of the well-known Turc-Pike relationship [Turc, 1954; Pike, 1964] that is frequently used to estimate interannual discharge:

$$Q = P \left(1 - \frac{1}{\left(1 + \left(\frac{P}{\alpha \cdot PE} \right)^\gamma \right)^{1/\gamma}} \right), \quad (6)$$

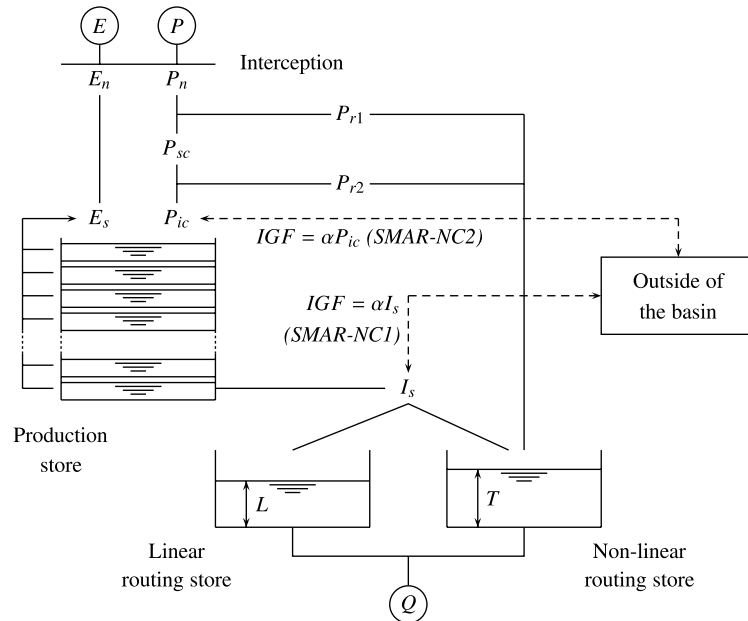


Figure 4. Structure of the SMAR daily time step model. The originality of the structure lies in the layered soil moisture accounting (production) store. The conservative version uses eight parameters. The NC1 and NC2 versions use a ninth parameter to account for intercatchment groundwater flow (IGF) at a different level.

Table 3. List of the Model Variants Tested in This Study^a

	Variants
Correction on rainfall input	$P' = X_2 \cdot P$
Correction on PE input	$PE' = X_2 \cdot PE$
Surface (or discharge) correction	$A' = X_2 \cdot A$ ($Q' = Q/X_2$)
Intercatchment groundwater flow	$IGF = f(X_2 \cdot (R/X_3)^{1/2})$

^a X_2 is the free parameter when the IGF function is shut down. Primed letters are corrected values. PE is mean potential evapotranspiration.

i.e.,

$$y = 1 - \frac{1}{\left(1 + \left(\frac{x}{\alpha}\right)^\gamma\right)^{1/\gamma}} \quad (7)$$

The original Turc-Pike relationship is given by $\alpha = 1$ and $\gamma = 2$, but we can obtain a whole family of curves which spread over the conservative domain in the nondimensional space (see Figure 6): The lower boundary $Q/P = 1 - PE/P$ is the limit curve when $\alpha = 1$ and $\gamma \rightarrow +\infty$, whereas the upper boundary $Q/P = 1$ is the limit curve when $\alpha \rightarrow 0$ and $\gamma \rightarrow 0$.

[33] Figure 7 shows that in terms of interannual water balance the alternative solutions clearly converge toward a generalized Turc-Pike relationship corresponding to $\alpha = 0.9$ and $\gamma = 2.5$ in equation (7) (it is an interesting result by itself, but discussing it is beyond the scope of this paper). We can see that the less scattered plot is obtained with the Q -IGF term. The reason why the solution accounting explicitly for IGF gives the best fit to a Turc-Pike relationship may be that once we have removed the “noise” of nonclimatic fluxes (here IGF, see equation (9)) from the catchment response (Q), it can take a smoother form depending only on climatic forcings P and PE:

$$\frac{Q}{P} = \frac{Q - IGF}{P} + \frac{IGF}{P} \quad (8)$$

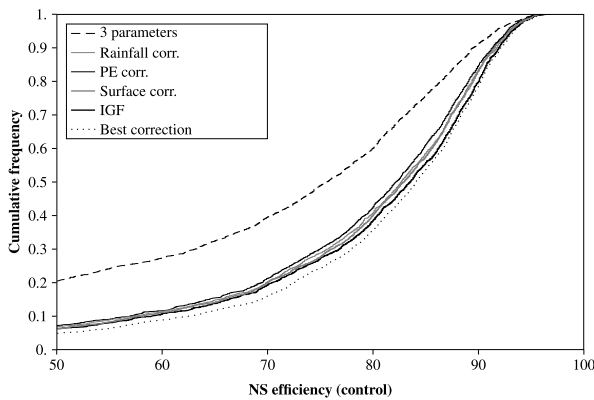


Figure 5. Distributions of Nash-Sutcliffe efficiencies in control mode for each corrective solution for the 1040 catchments (2080 control periods). The dotted line is the distribution of efficiencies when selecting the best solution among the four for each catchment (upper limit of what we can expect in selecting one solution).

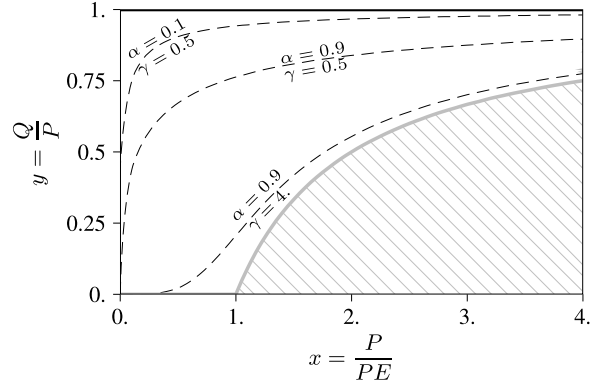


Figure 6. Family of curves given by the generalized Turc-Pike relationship presented in the nondimensional hydroclimatic coordinates. This family spreads over the conservative domain of Figure 1. Q , mean interannual streamflow; P , mean interannual catchment precipitation; PE, mean interannual Penman potential evapotranspiration; α and γ , parameters of the generalized Turc-Pike relationship (see equation (7)).

$$\frac{Q}{P} = f\left(\frac{P}{PE}\right) + \frac{IGF}{P} \quad (9)$$

Of course, it must be recognized that we do not have any formal guarantee that the IGF computed with this method are indeed correct since we rarely find a situation where we can measure them. However, at least the model will not twist or play with climatic inputs if there is obviously no way to explain the runoff with these inputs only, and our experience is that this is merely the rule rather than the exception. Section 5 will seek possible justifications for the likelihood or the unlikelihood of the four options discussed in this paper.

5. Realism of Calibrated Correcting and Scaling Parameters

[34] If there is no way to prove that the IGF computed by the RR model is true, we can at least point out the inconsistency of the alternative correcting and scaling solutions. The analysis of the distribution of correcting factors is additional evidence of the drawbacks of these methods.

5.1. Hydrological Realism

[35] Figure 8 shows the distributions of the calibrated correction factors applied to each input (using “box and whiskers” plots). We present here the distributions for the variants of both GR4J and SMAR in order to ensure that our interpretation is not structure-dependent.

[36] In the three cases and for the two base structures the distributions of the corrections are not centered on 1.0, which is not consistent with parameters intended to correct random measurement errors. There is no reason to believe that rainfall is almost systematically overestimated (most hydrometeorologists would rather advocate the opposite), that PE is underestimated, and that cartographs or digital elevation models processing software almost systematically

W06428

LE MOINE ET AL.: DEALING WITH IGF IN RAINFALL-RUNOFF MODE

W06428

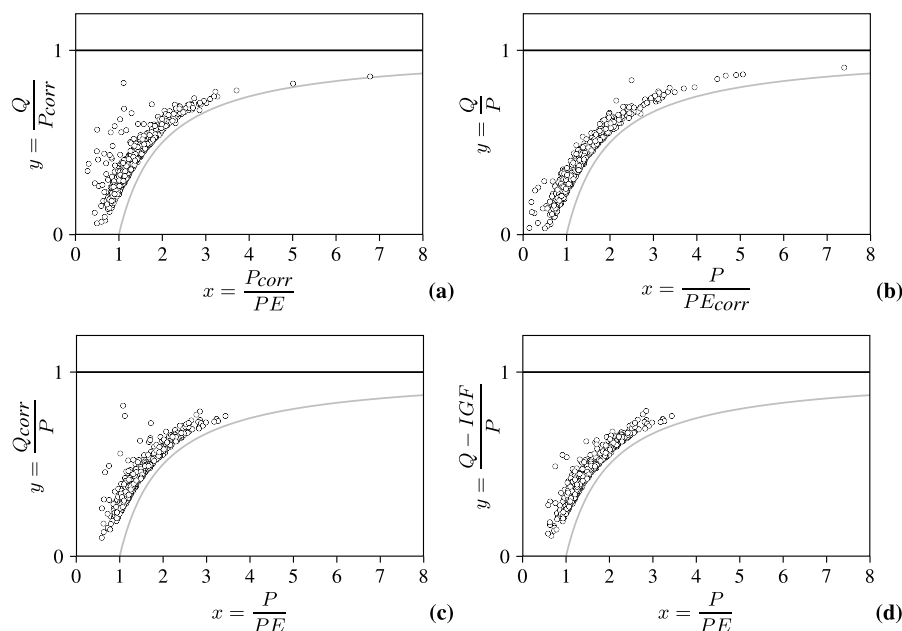


Figure 7. Simulated interannual water balance after applying corrective factors, mapped in nondimensional space for each solution: (a) rainfall correction, (b) PE correction, (c) surface (or discharge) correction, and (d) accounting explicitly for IGF. The Q -IGF/ P versus P/PE plot produces the least stretched and dispersed cluster.

overestimates catchment area. Our point of view here is that even if in some cases input data are known with a bias that could be corrected, there is no reason to accept any of these corrective solutions since a solution with the same number of parameters using raw observations can provide better simulations.

[37] IGF terms computed by the IGF-accounting versions of the GR4J and SMAR models are plotted in Figure 9. They are well correlated, although both structures are very different from a conceptual point of view.

[38] Of course, one can argue that corrections were applied one by one only, whereas errors can exist on rainfall, PE, and discharge simultaneously. Using two or three corrections at the same time would probably give more “realistic” values but would not change the fact that the distributions are not centered on 1 and that for most of the catchments, all corrections tend to discard some excess water (i.e., to reduce rainfall or surface or to increase PE). We have also seen in section 4.2 that all corrections produce the same kind of transformation in the nondimensional space, so that combining them would not change the result.

[39] Concerning the validity of PE data, the Penman model is only one of the possible inputs since there is no way to verify that it remains appropriate at the catchment scale. *Oudin et al.* [2005a, 2005b] show that a simplified PE model (based on extraterrestrial radiation and ground level temperature only) provides slightly better simulations for several well-known RR models, although its cumulative annual values are slightly lower than Penman values (around 80% on average). We will not discuss the respective benefits of choosing one PE formulation over another, but the nonuniqueness of acceptable PE input data is another reason to avoid PE scaling as a way to discard excess water.

Furthermore, Figure 10 shows that this operation is inefficient since maximum PE and maximum water availability rarely occur at the same time.

5.2. Geological Determinants

[40] The last evidence for the low likelihood of the three corrective solutions tested here is the existence of a correlation (even if not strong) between corrective factors and lithology. Figure 11 shows the results for the rainfall correction. For each class of lithology we selected the catchments underlain at least 50% by this lithology (homogeneous catchments). Figure 11 shows that the rainfall correction is correlated with lithology, which is proof by contradiction for the hydrological unlikelihood of this parameter.

[41] In contrast, we plotted the same distributions for the IGF/ P ratio in Figure 12, for which the geological correlation is sensible. Of course, great variability remains inside each class, which is consistent with the variability of the causes of IGF mentioned in the introduction. Crystalline bedrock, if not weathered, will act as an impermeable layer but will almost behave as a karst if it is widely faulted and weathered. What is more, on a regional basis we can find both gaining and leaking catchments for the same lithological unit, especially if it allows discrete pathways (e.g., a karstic resurgence will often be located next to a leaking catchment; this is less likely for regions where large, continuous aquifers exist). Figure 12 shows that a mere description of basin geology is not sufficient to narrow the uncertainty of the IGF parameter estimate: We would ideally need information concerning the extent and direction of “hydrologically active” faults, for instance. Only the rainfall-PE-runoff relationship at the catchment scale can demonstrate something about the conservativeness of the

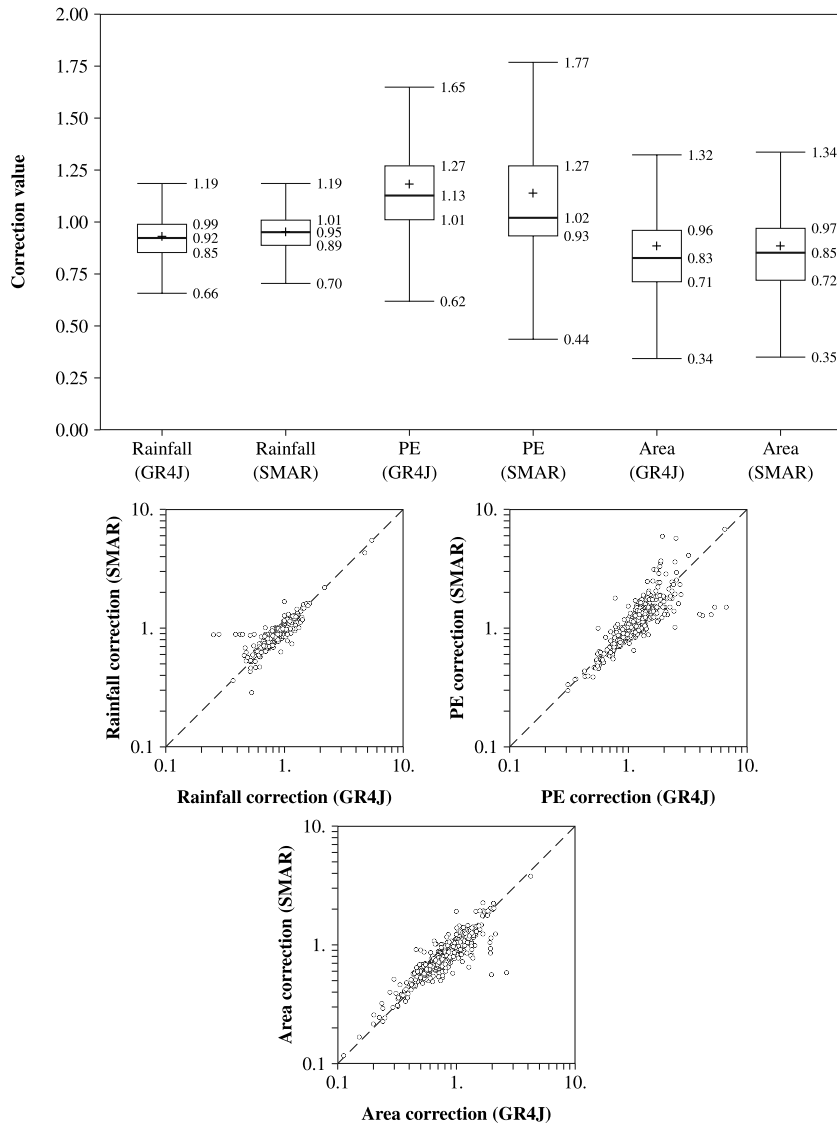


Figure 8. Distributions of parameter values for each corrective solution, represented using “box and whiskers” plots. The boxes are delimited by the first and third quartiles (Q_{25} and Q_{75}), the median is marked with a thick line, and the mean is marked with a cross. The length of the “whiskers” is 1.5 times the interquartile range. Noncentered distributions show that the parameters account for a systematic bias in the water balance and not only for randomly distributed measurement errors. The three bottom plots show that the correcting parameters that seem necessary to close the water balance are quite independent of the model structure.

system because we will never be able to describe and quantify all potential leaks, diffuse as well as localized ones.

6. Conclusion

[42] This article has focused on intercatchment groundwater flows, a problem which has been ignored for too long in rainfall-runoff modeling. Naturally, the exact hydrological nature of these water fluxes may vary. They might be small in many catchments, sometimes negligible. However, the hypothesis of IGF being equal to zero for every catchment at all times is indefensible. Proof by contradic-

tion has been given by *Mouelhi et al.* [2006], who argued that if IGF was to be zero for all catchments, moving the catchment outlet progressively upstream would lead to a situation where the catchment would have to be a slice of impermeable material, which is obviously absurd.

[43] In this paper we presented results obtained with two models and several modeling solutions, all aimed at solving the water balance issue, for a large set of catchments. All solutions were strictly equivalent from the numerical complexity point of view and only differed by the level at which they were applied. A solution devised to account explicitly for IGF could thus be compared to alternatives involving an

W06428

LE MOINE ET AL.: DEALING WITH IGF IN RAINFALL-RUNOFF MODE

W06428

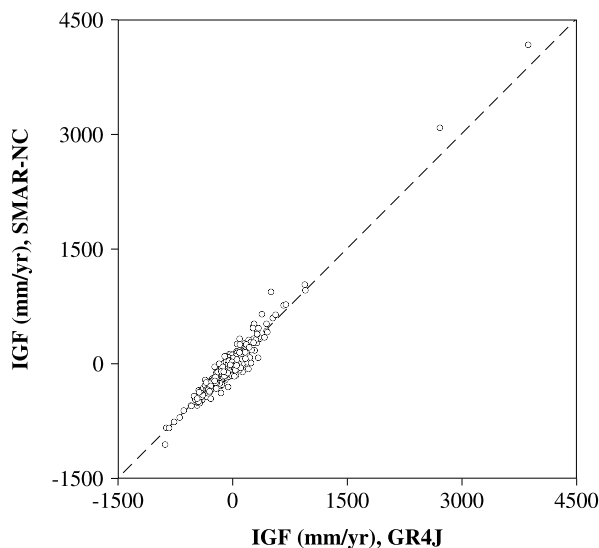


Figure 9. Comparison of interannual IGF terms computed with GR4J and SMAR-NC (the latter being the term given by the best of SMAR-NC1 and SMAR-NC2 on each catchment).

input correction or scaling parameter. What can be concluded from this experiment?

[44] First, we have shown that the explicit accounting of IGF does improve model efficiency (both in calibration and control). Given the size of our data set and its hydrological and geological diversity, we believe this result can be generalized.

[45] Second, we have also shown how dangerous alternative input correction parameters can be if applied blindly. Proof by contradiction was provided. For example, if

rainfall correction solutions were to be considered a valid alternative to the explicit representation of IGF, then the rainfall correction parameters would have some geologic determinant. This would obviously be disturbing. To test the generalizability of this statement even further, we conducted additional tests with other models (not reported herein for reasons of length) concerning the likelihood of input corrections. We applied the same methodology to versions of SMAR (see section 5.1 for partial results), HBV, and TOP-MODEL. The conclusions were basically the same as for the GR4J model.

[46] The only disappointing point here is that it was not possible to establish clear links between the lithology and IGFs. This can be understood, however, by considering that depending on bedrock age, weathering, faulting, connectivity of faults, the same lithology can have a very different hydrogeological behavior. What is more, the lumped, catchment-by-catchment approach used in this paper does not make any difference between transfers with adjacent catchments and transfers with downstream reaches. We could strengthen our conclusions with a semidistributed approach that would allow identifying leaking/gaining reaches more precisely (though we would still find groups of catchments for which the balance shows a deficit, e.g., those feeding aquifers flowing out to the sea). Hopefully, we will be able to report this work later.

[47] We all know that input data are known with uncertainty. Water balance computations are subject to errors which are inextricable combinations of model error, measurement error, and IGF. Thus, before starting the modeling phase, all efforts must be made to provide the best input estimates possible. However, once the RR modeling process has started, we can no longer act as if input data were wrong, and we should better work on large data sets where the sampling biases fade. Then our results indicate that the most realistic solution to identify the components of the

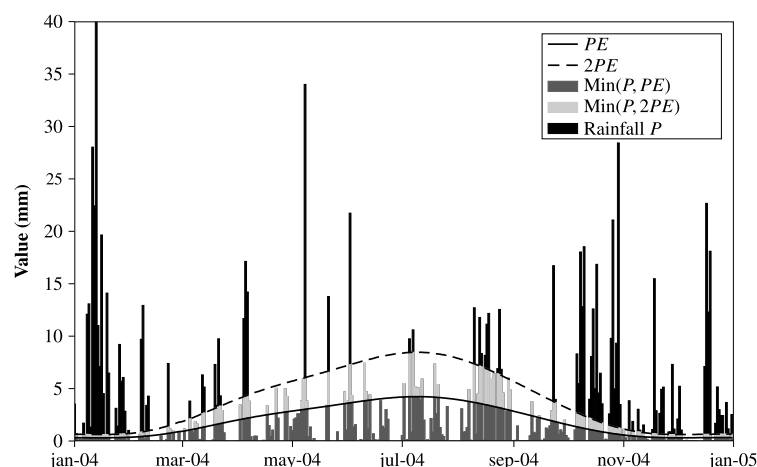


Figure 10. Illustration of the decreasing marginal effect of a potential evapotranspiration increase when aiming at increasing actual evapotranspiration. First, it can be seen that multiplying PE by 2 does not multiply interception by 2 (the rainfall intercepted by the area in light gray is less than twice the rainfall intercepted by the area in dark gray). Second, remember that soil moisture extraction has a decreasing yield. PE scaling is the clumsiest way to correct water balance in a hydrological model because maximum PE often occurs at times of minimum water availability.

W06428

LE MOINE ET AL.: DEALING WITH IGF IN RAINFALL-RUNOFF MODE

W06428

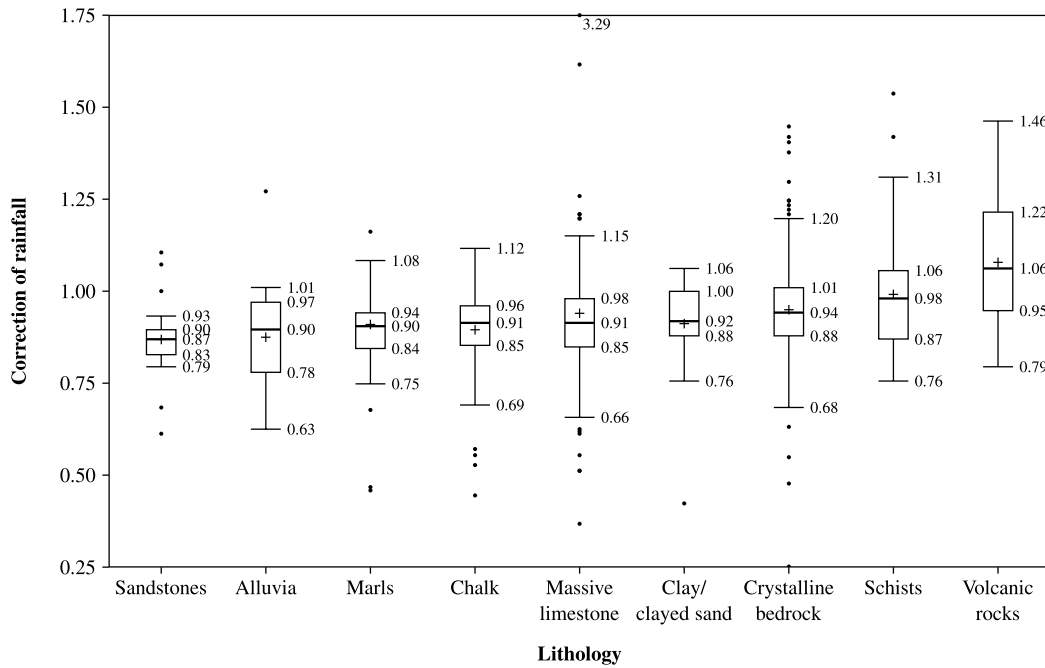


Figure 11. Distributions of the rainfall correction parameter for several lithology classes (outliers are plotted with dots). The link between lithology and the rainfall correcting parameter is clear proof by contradiction that such a correction cannot be trusted.

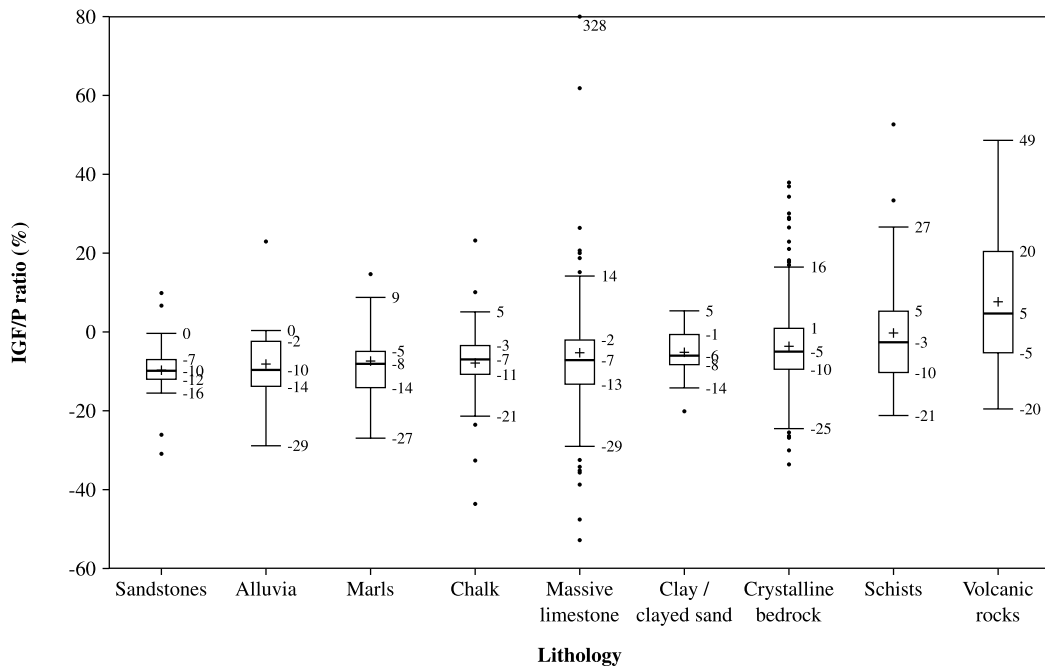


Figure 12. Distributions of the ratio between IGF and mean annual rainfall for several lithology classes, as computed by the GR4J model. The influence of lithology appears more sensible than for input corrections.

W06428

LE MOINE ET AL.: DEALING WITH IGF IN RAINFALL-RUNOFF MODE

W06428

surface water balance is to allow explicitly for a form of intercatchment groundwater flow in any RR model.

[48] **Acknowledgments.** This research is based on a very large hydrometeorological data set, and we would like to acknowledge the contribution of Météo France in providing the SAFRAN pluviographic archive (special thanks go to Joël Noilhan and Laurent Franchisteguy), the contribution of the SCHAPI in providing access to the HYDRO streamflow archive (special thanks go to Laurent Cobos), and the work of Jean-Louis Rosique who digitalized most of the 1040 catchment boundaries. The first author is grateful to K. O'Connor, who introduced him to the SMAR model during his stay in Galway.

References

- Beven, K. J. (2001), On hypothesis testing in hydrology, *Hydrol. Processes*, 15, 1655–1657.
- Carillo-Rivera, J. J. (2000), Application of the groundwater-balance equation to indicate interbasin and vertical flow in two semi-arid drainage basins, Mexico, *Hydrogeol. J.*, 8, 503–520.
- Carillo-Rivera, J. J., A. Cardona, and D. Moss (1996), Importance of the vertical component of groundwater flow: A hydrogeochemical approach in the valley of San Luis Potosi, Mexico, *J. Hydrol.*, 185, 23–44.
- Chiew, F. H. S., and T. A. McMahon (1990), Estimating groundwater recharge using a surface watershed modelling approach, *J. Hydrol.*, 114, 285–304.
- Eakin, T. E. (1966), A regional interbasin groundwater system in the White River area, southeastern Nevada, *Water Resour. Res.*, 2(2), 251–271.
- Gan, T. Y., and G. F. Biftu (1996), Automatic calibration of conceptual rainfall-runoff models: Optimization algorithms, catchment conditions, and model structure, *Water Resour. Res.*, 32(12), 3513–3524.
- Genereux, D. P., and M. T. Jordan (2005), Interbasin groundwater flow and groundwater interaction with surface water in a lowland rainforest, Costa Rica: A review, *J. Hydrol.*, 320, 385–399.
- Goswami, M. (2006), Rainfall-runoff simulation and forecasting using hydrological and data-driven models and techniques, Ph.D. thesis, Natl. Univ. of Ireland, Galway.
- Goswami, M., and K. M. O'Connor (2005), Application of a conceptual rainfall-runoff simulation model to three European catchments characterised by non-conservative system behaviour, in *Proceedings of the International Conference on Hydrological Perspectives for Sustainable Development (HYPESD-05)*, edited by M. Perumal et al., Indian Inst. of Technol. Roorkee, Roorkee.
- Hudson, M. R., and D. N. Mott (1997), Faulting and coincident interbasin groundwater flow in a karst aquifer, Buffalo National River region, northwestern Arkansas, *Geol. Soc. Am. Abstr. Programs*, 29(6), 181–182.
- Hughes, D. A. (2004), Incorporating groundwater recharge and discharge functions into an existing monthly rainfall-runoff model, *Hydrol. Sci. J.*, 49(2), 297–311.
- Ivkovic, K. M., B. F. W. Croke, R. Letcher, and W. R. Evans (2005), The development of a simple model to investigate the impact of groundwater extraction on river flows in the Namoi catchment, NSW Australia, paper presented at “Where Waters Meet” NZHS-IAH-NSSSS Conference, N. Z. Hydrol. Soc., Auckland, 28 Nov. to 2 Dec.
- Klemeš, V. (1986), Operational testing of hydrological simulation models, *Hydrol. Sci. J.*, 31(1), 13–24.
- Konrad, C. P. (2006), Location and timing of river-aquifer exchanges in six tributaries to the Columbia River in the Pacific Northwest of the United States, *J. Hydrol.*, 329, 444–470.
- Linsley, R. K. (1982), Rainfall-runoff models: An overview, in *Proceedings of the International Symposium on Rainfall-Runoff Modelling*, edited by V. P. Singh, pp. 3–22, Water Resour. Publ., Littleton, Colo.
- Margat, J. (2006), *Les eaux souterraines, une richesse mondiale*, U. N. Educ. Sci. and Cult. Organ., Paris.
- Mouelhi, S., C. Michel, C. Perrin, and V. Andréassian (2006), Stepwise development of a two-parameter monthly water balance model, *J. Hydrol.*, 318, 200–214.
- Nascimento, N. O. (1995), Appréciation à l'aide d'un modèle empirique des effets d'action anthropiques sur la relation pluie-débit à l'échelle du bassin versant, Ph.D. thesis, École Natl. des Ponts et Chaussées, Paris.
- O'Connell, P. E., J. E. Nash, and J. P. Farel (1970), River flow forecasting through conceptual models. Part II: The Brosna catchment at Ferbane, *J. Hydrol.*, 10, 317–329.
- Oudin, L., C. Michel, and F. Anctil (2005a), Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 1: Can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs?, *J. Hydrol.*, 303, 275–289.
- Oudin, L., F. Hervieu, C. Michel, C. Perrin, V. Andréassian, F. Anctil, and C. Loumagne (2005b), Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2: Toward a simple and efficient potential evapotranspiration model for rainfall-runoff modelling, *J. Hydrol.*, 303, 290–306.
- Oudin, L., V. Andréassian, T. Mathevet, C. Perrin, and C. Michel (2006), Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations, *Water Resour. Res.*, 42, W07410, doi:10.1029/2005WR004636.
- Perrin, C., C. Michel, and V. Andréassian (2003), Improvement of a parsimonious model for streamflow simulation, *J. Hydrol.*, 279, 275–289.
- Pike, J. G. (1964), The estimation of annual runoff from meteorological data in tropical climate, *J. Hydrol.*, 2, 116–123.
- Pitman, W. V. (1973), A mathematical model for generating river flows from meteorological data in South Africa, *Rep. 2/73*, Hydrol. Res. Unit, Univ. of the Witwatersrand, Johannesburg, South Africa.
- Porter, J. W., and T. A. McMahon (1976), The Monash model: User manual for daily program HYDROLOG, *Res. Rep. 2/76*, 41 pp., Dep. of Civ. Eng., Monash Univ., Clayton, Victoria, Australia.
- Thyne, G. D., J. M. Gillepsie, and J. R. Ostdick (1999), Evidence for interbasin flow through bedrock in the southeastern Sierra Nevada, *Geol. Soc. Am. Bull.*, 111(11), 1600–1616.
- Tóth, J. A. (1963), A theoretical analysis of groundwater flow in small drainage basins, *J. Geophys. Res.*, 68, 4795–4812.
- Turc, L. (1954), Le bilan d'eau des sols: Relation entre les précipitations, l'évapotranspiration et l'écoulement, *Ann. Agron., Ser.*, 491–595.

V. Andréassian, N. Le Moine, C. Michel, and C. Perrin, Hydrosystems and Bioprocesses Research Unit, Cemagref, Parc de Tourvoie, BP 44, F-92163 Antony Cedex, France. (nicolas.le-moine@cemagref.fr)

5

Confrontation du bilan de surface et souterrain sur le système karstique de la Rochefoucauld-Touvre (Charente)

Nicolas Le Moine, Vazken Andréassian, et Thibault Mathevet
Article publié dans *Water Resources Research*, vol. 44, 2008

“ La pataphysique sera surtout la science du particulier, quoiqu'on dise qu'il n'y a de science que du général. Elle étudiera les lois qui régissent les exceptions et expliquera l'univers supplémentaire à celui-ci; ou moins ambitieusement décrira un univers que l'on peut voir et que peut-être l'on doit voir à la place du traditionnel, les lois que l'on a cru découvrir de l'univers traditionnel étant des corrélations d'exceptions aussi, quoique plus fréquentes, en tous cas de faits accidentels qui, se réduisant à des exceptions peu exceptionnelles, n'ont même pas l'attrait de la singularité.”

Alfred JARRY – *Gestes et opinions du Docteur Faustroll, pataphysicien.*

Résumé en français

La résurgence de la Touvre est l'exutoire unique d'un grand système karstique, le karst de La Rochefoucauld, dans l'ouest de la France. Ce système est particulièrement intéressant du point de vue de la modélisation hydrologique, en ce qu'il permet d'étudier la façon dont les pertes et apports karstiques peuvent être pris en compte dans un modèle pluie-débit. En effet, cette résurgence est alimentée par les pertes de trois rivières, qui sont elles-mêmes jaugées avant et après avoir contribué à la recharge de l'aquifère karstique. Cette configuration est exceptionnelle puisqu'elle permet de comparer les pertes des bassins de surface (simulées par le modèle pluie-débit GR4J) avec les mesures de débit de la résurgence, à l'exutoire du système karstique. Elle offre une occasion de confronter les estimations de ces pertes — souvent suspectées mais presque jamais mesurables — à des observations. Nous montrons ici que la conceptualisation des pertes de surface dans le modèle GR4J aboutit à des volumes de pertes cohérents avec les volumes écoulés par la résurgence et que, moyennant l'ajout d'un réservoir modélisant le comportement retardant du réservoir karstique, un modèle intégré peut être proposé pour représenter de façon satisfaisante le comportement de l'ensemble du système. Ce cas d'étude sur la Touvre montre qu'une représentation parcimonieuse et empirique de la relation pluie-débit sur un système karstique est possible. Il fournit aussi une corroboration indépendante, quoiqu'unique, de la représentation conceptuelle des échanges surface-souterrain dans le modèle GR4J.



Confronting surface- and groundwater balances on the La Rochefoucauld-Touvre karstic system (Charente, France)

Nicolas Le Moine,¹ Vazken Andréassian,¹ and Thibault Mathevet²

Received 22 February 2007; revised 3 October 2007; accepted 20 November 2007; published 4 March 2008.

[1] The Touvre spring is the only outlet of a large karstic system, the La Rochefoucauld karst, located in the West of France. This system is particularly interesting from a modeler's point of view, for it allows to investigate how karstic losses and gains can be accounted for within a rainfall-runoff model. Indeed, this spring is fed by the losses of three rivers, which are themselves gaged after recharging the karstic aquifer: this setting is exceptional in that simulations of surface catchment losses (estimated by the GR4J rainfall-runoff model) can be compared with measurements of spring discharge at the outlet of the karstic system. It provides an opportunity to test the estimates of "catchment leakages", which are often suspected but seldom measurable, against observations. We show here that the conceptualization of surface losses in the GR4J rainfall-runoff model yields simulated leakage volumes that are consistent with the spring flow volumes and that, provided a routing store is added to GR4J to account for the specific retarding behavior of the karstic aquifer, an integrated model can be proposed to represent the behavior of the whole system adequately. The Touvre case study shows that a parsimonious, empirical rainfall-runoff representation of a karstic system is possible. It also provides an independent corroboration for the conceptualization of surface catchment leakages proposed in the GR4J rainfall-runoff model.

Citation: Le Moine, N., V. Andréassian, and T. Mathevet (2008), Confronting surface- and groundwater balances on the La Rochefoucauld-Touvre karstic system (Charente, France), *Water Resour. Res.*, 44, W03403, doi:10.1029/2007WR005984.

1. Introduction

1.1. Non-Conservative Catchments are More Frequent Than Commonly Expected

[2] One of the still unsolved problems of rainfall-runoff (RR) modeling is that of accounting for underground, point or non-point source gains or losses which affect the catchment-scale water balance. It must be recognized that surface catchments lying on a continuous impervious horizon are the exception rather than the rule: most surface catchments may either leak (by losing water to neighboring catchments or to the sea) or gain water (originating from outside the catchment). Several terms have been coined for these catchments. For clarity, we adopt the following terminology in this paper:

[3] ● leaking/gaining catchments are called non-conservative (NC);

[4] ● the associated underground flux is called intercatchment groundwater flow (IGF). This flux can be positive (inflowing water) or negative (outflowing water).

[5] Where fluxes transferred to or from NC catchments are small, ignoring them is not a matter of concern. However, when it comes to karstic catchments, large amounts of water may transit below the topographic water divide, and IGFs

cannot be ignored. These fluxes are made possible by the well-known karstification phenomena, widespread in limestone throughout the world, although similar phenomena can also exist in volcanic substrata, as well as in chalk horizons [Margat, 2008].

1.2. Karstic Flows are a Problem for Rainfall-Runoff Models

[6] Representing the rainfall-runoff transformation in catchments which are either partly or completely covered by karstic terrains is a challenge for hydrologic modeling: because neither evaporative fluxes nor underground fluxes are commonly observable, we find ourselves in the uncomfortable situation of having to model a system with too many unknowns for too few equations. Uncertainties on the measurable part of the fluxes (precipitation and streamflow) further complicate matters. In a recent paper [Le Moine *et al.*, 2007], we presented a review of solutions that may numerically allow to account for non-conservative behaviors. We tested alternative solutions using a large set of catchments and concluded that from both the distribution of model efficiency and the physical realism points of view, an explicit representation of IGF was statistically preferable to scaling or input correcting solutions. However, we could not directly test the physical realism of simulated underground fluxes, which are usually unmeasurable.

1.3. Scope of the Paper

[7] In this paper, we analyze a well-defined and widely investigated karstic system which geologists believe, for

¹CEMAGREF, Hydrosystems and Bioprocesses Research Unit, Antony Cedex, France.

²Électricité de France, Direction Technique Générale, Département Surveillance Service CADE, Grenoble Cedex 9, France.

structural reasons [see, e.g., *de Grissac et al.*, 2006], has a single outlet: the Touvre exsurgence. On this system, we can compare simulations of karstic losses by surface catchments with actual spring discharge measurements. As we will show, this confrontation is successful and leads us to propose a lumped representation for the overall system. In section 2, we present a rapid review of previous RR modeling experience with karstic catchments. Then section 3 presents the peculiarities of the Touvre exsurgence and shows why its natural setting is so advantageous for comparing simulations of surface catchment losses with actual exsurgence flow measurements. In section 4, we introduce the RR model used and its approach to leakage evaluation, and we show in section 5.1 that the modeled leakage amounts fit well the annual exsurgence flow. Last, section 5.2 discusses how the GR4J model can be completed with an additional store to obtain a complete representation for the rainfall-runoff behavior of the karstic spring, combining independently obtained recharge and storage terms (section 5.3). In conclusion (section 6), we discuss the limits of what we have interpreted as an independent check on the catchment leakage conceptualization of GR4J.

2. Rainfall-Runoff Modeling of Karstic Catchments: A Review

[8] Modeling the rainfall-runoff relationship within karstic catchments is a challenge because of the high complexity of such hydrological systems, which may present different levels of karstification, from less developed aquifers called “diffuse flow aquifers” to well-developed aquifers called “conduit flow aquifers” [*Atkinson*, 1977]. Streamflow at the outlet of karstic catchments is generated by a combination of diphasic flows in the non-saturated zone, Darcian flows in the saturated zone, and turbulent flows in open channels. Moreover, karstic catchments are usually characterized by great spatial and physical heterogeneity, which causes their particularly non-linear and non-stationary hydrological functioning. For these reasons, studies attempting to infer groundwater to surface water fluxes in karstic systems have been mostly based on geochemical evidence (solute budget studies, see e.g., *Katz et al.*, 1998) rather than on rainfall-runoff analyses. Because of the complexity of such hydrological systems, karstic catchments could be seen as ‘hybrid’ catchments, mixing groundwater aquifer and surface catchment behaviors. Therefore the attempts to model rainfall-runoff relationships within karstic catchments were made from different points of view: physical (hydrogeological) models, conceptual (reservoir) models, and black-box models [*Jukic and Denic-Jukic*, 2006]. Physical (or theoretical) models are based on hydraulic laws that represent different flow processes in porous media [*Eisenlohr et al.*, 1997a, 1997b]. Such models require extensive data collection to characterize the distributed aquifer properties (e.g., aquifer geometry, position and direction of fractures and open channels, hydraulic properties). Because of the shortage of such data, 2D or 3D complex numerical models are usually used in a degraded mode, with rather substantial simplifications of the physical and hydrogeological structures that these methods aim at representing. Conceptual models are based on a simplified but efficient representation of hydrological systems, using interconnected reservoirs and unit hydrographs [*Barrett and Charbeneau*, 1997; *Halihan et al.*,

1998; *Jukic and Denic-Jukic*, 2006; *Rimmer and Salinger*, 2006]. Such models require limited amounts of data (precipitation, potential evapotranspiration (PET) and streamflow): they are easier to implement and can be used for hypothesis testing to explore the complexity of such hydrological systems. Black-box models were introduced in karstic hydrology by *Mangin* [1981, 1984], who proposed to study karstic catchments as systems characterized by their input/output relationships. These models are based on the stochastic process theory developed by *Yevjevich* [1972] and *Box and Jenkins* [1976]. Rainfall-runoff relationships were first studied with a descriptive approach to characterize the properties of karstic catchments, using correlation and spectral analyses [*Mangin*, 1984; *Padilla and Pulido-Bosch*, 1995; *Laroque et al.*, 1998; *Labat et al.*, 2000a, 2000b], convolution, wavelet analyses, and multifractal analyses [*Labat et al.*, 2000a, 2000b] applied to various time series (e.g., rainfall, streamflow, water temperature, solid concentration). Studies that report quantitative modeling of rainfall-runoff relationships are much rarer: the tools used for that task include artificial neural networks (ANN) [*Johannet et al.*, 1994], linear and non-linear functions [*Labat et al.*, 1999; *Jukic and Denic-Jukic*, 2006].

3. The Touvre Spring and its Catchment

[9] The Touvre is a large karstic spring located in western France and is a tributary of the Charente River. The spring, with a mean annual flow around 13 m³/s, is the second largest in France and is the sole outlet of a large karstic network developed in Jurassic limestone (La Rochefoucauld karst): indeed the location of the spring is the only place where the karstic network intersects both the stratigraphic discontinuity (i.e., the limit between Jurassic limestone and the more impervious Cretaceous chalk) and a topographic minimum [see, e.g., *Queleuennec et al.*, 1971; *Laroque and Razack*, 1998; *Laroque et al.*, 1998; *de Grissac et al.*, 2006]. The Touvre spring flow is so important that it can be seen as the effective spring of the Charente River during summer. The Touvre spring is original in that its karstic network is fed by the losses of three rather large rivers running down from the western side of the Massif Central highlands: the Tardoire, the Bandiat, and the Bonnieure Rivers (see Figure 1 and Table 1). Before reaching the limestone plateau, the Tardoire and Bandiat rivers have a mean discharge of 6.1 and 4.1 m³/s respectively and a yield ranging between 0.35 and 0.43, typical for impermeable catchments in this crystalline bedrock area. Downstream from the limestone plateau the Tardoire River has a reduced streamflow and it fails to flow 2–8 months a year: the mean discharge at Coulgens is 5.0 m³/s which is only half of the sum of the discharges of the upstream subcatchments (in spite of a twofold increase of the area). The Bandiat fails to reach the Tardoire almost all the time because of the high density of sinkholes in its lower course.

[10] Data quality is generally good for the Bandiat and Tardoire streamflow measurements, whereas hydrometric problems occur at the Touvre’s gauging sites because of the growth of macrophytes that nearly entirely cover the river in summer. Table 1 presents elements of the water balance for the subcatchments, before and after they reach the karstic plateau. There is a clear, sudden decrease in

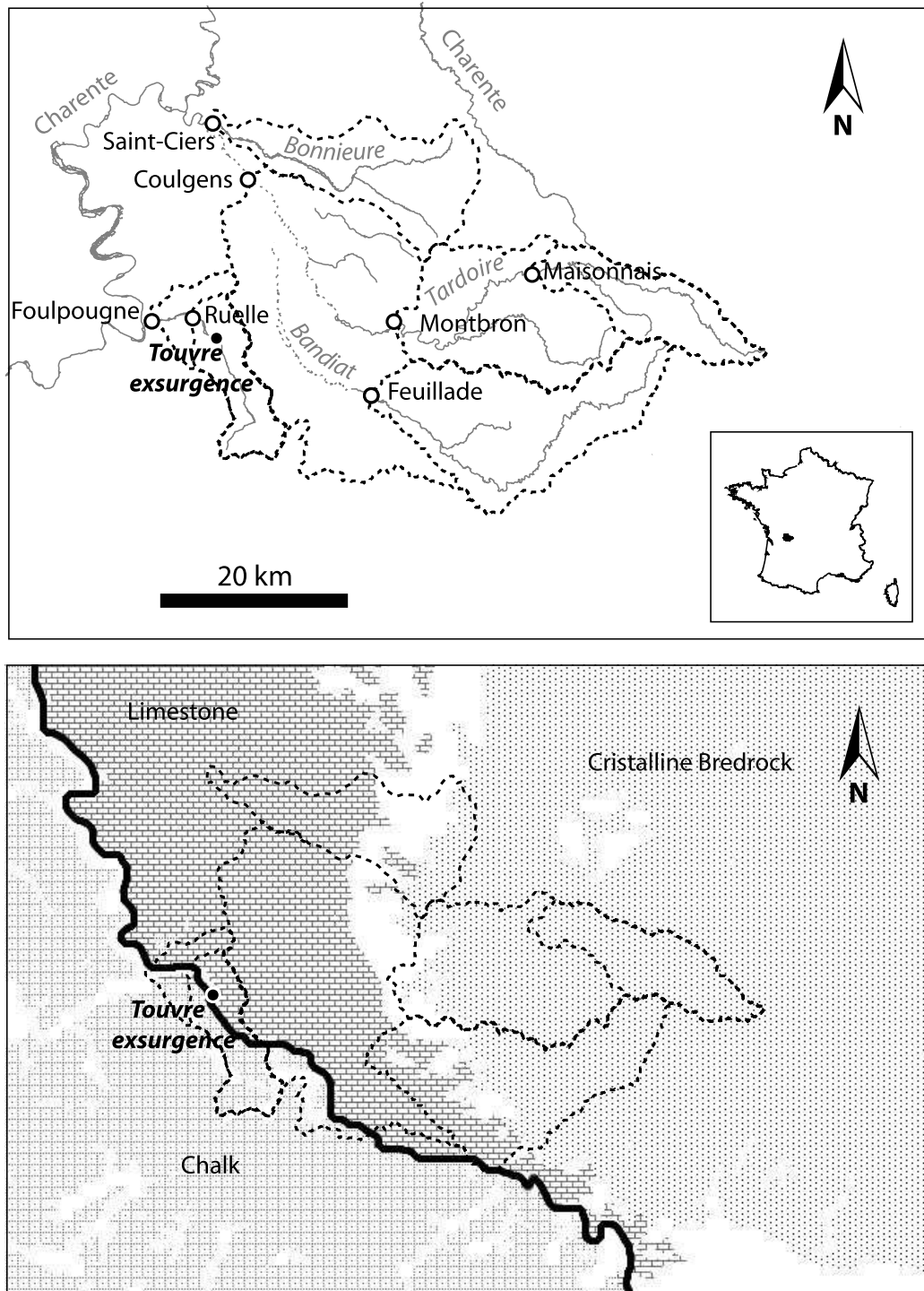


Figure 1. General maps of the *La Rochefoucauld* karstic system. Top: Topographical setting. Gauging sites are shown with white spots and topographic catchment boundaries with dashed lines. Bottom: Geological setting. The thick black line is the permeability contrast between karstified limestone and chalk, on which the spring is located.

Table 1. Gauged Catchments Used in This Study, and Their Mean Annual Water Balance Over the 1980–2004 Period (Except for the Touvre at Ruelle Where Streamflow was Available Only Until 1996)^a

Catchment	Gauging Site	Area, km ²	Mean, m ³ /s	Disch. (Q _m) mm/a	Mean Rain. (P _m) mm/a	Yield Q _m /P _m
Upper Tardoire (1)	Maisonnais	140	2.0	457	1061	0.43
Upper Tardoire (2)	Montbron	389	6.1	492	1132	0.43
Bandiat	Feuillade	333	4.1	390	1102	0.35
Lower Tardoire	Coulgens	1200	5.0	131	1055	0.12
Bonnieure	Saint-Ciers	203	1.6	250	955	0.26
Touvre (1)	Ruelle	91	13.4	4644	947	4.90
Touvre (2)	Foulpougne	126	13.4	3347	945	3.54

^aNote that the Touvre spring cannot be gauged directly at its exsurgence site: where streamflow measurements become possible, the own topographic catchment of the Touvre River must be taken into account. Thus we speak either of the Touvre spring or Touvre exsurgence when referring to the production of the karst, and of the Touvre River when referring to the available flow measurements.

annual water yield downstream from the karstic plateau, which results from two phenomena:

[11] • a deficit of surface runoff production over the karstic plateau, i.e., a decrease in specific discharge (m³/s/km²)

[12] • a decrease in absolute discharge (m³/s) due to leaking streambeds (channel transmission losses): a significant part of upstream production (crystalline bedrock catchments) is lost in the lower course.

[13] Moreover, the Touvre River itself has a yield that can obviously not be accounted for by its own small catchment.

4. GR4J: A Rainfall-Runoff Model With an Explicit Intercatchment Groundwater Flow Representation

4.1. GR4J

[14] In this study we decided to use a specific rainfall-runoff model, GR4J, which besides its classical structure (based on a standard soil moisture accounting approach like many of the models of its category) has the originality of having an explicit Intercatchment Groundwater Flow (IGF) function, for which one of its four parameters must be calibrated. Of course, since the possibilities to actually measure IGFs are extremely rare, this parameter is calibrated together with the three others, in order to minimize the Mean Square Error (MSE) on the target variable: surface streamflow. Note that this IGF formulation is an empirical one, which was based on results obtained with a large set of catchments, and that this kind of empirical function is not specific to GR4J: it does exist in several RR models such as SMAR [Goswami and O'Connor, 2005], MODHYDROLOG [Chiew and McMahon, 1990], and Pitman [Hughes, 2004].

[15] Initially, Nascimento [1995] introduced the IGF functionality in the model to improve the representation of non-perennial streams by GR4J, but subsequent work showed that explicit accounting of IGF was beneficial to all catchments. We showed it again in a recent paper [Le Moine et al., 2007], where we explored in detail the different options that may allow a RR model to simulate non-conservative behaviors. We showed that both from the perspective of overall model efficiency and physical realism, an explicit representation of intercatchment groundwater flows is preferable. This view is supported by the work of Mouelhi et al. [2006], who used a monthly rainfall-runoff model with a large set of catchments, over which they show

that this representation is in fact essential to a balanced, general RR model. The GR4J model [see Perrin et al., 2003 for a complete description], which will be used in this paper, has a specific function of this type, controlled by one of its calibrated parameters, to simulate IGF.

[16] How does GR4J usually function to model surface catchments? It is run in a lumped (basin-scale), continuous mode, at the daily time step. It is very parsimonious since its structure involves only four free parameters (see Figure 2)

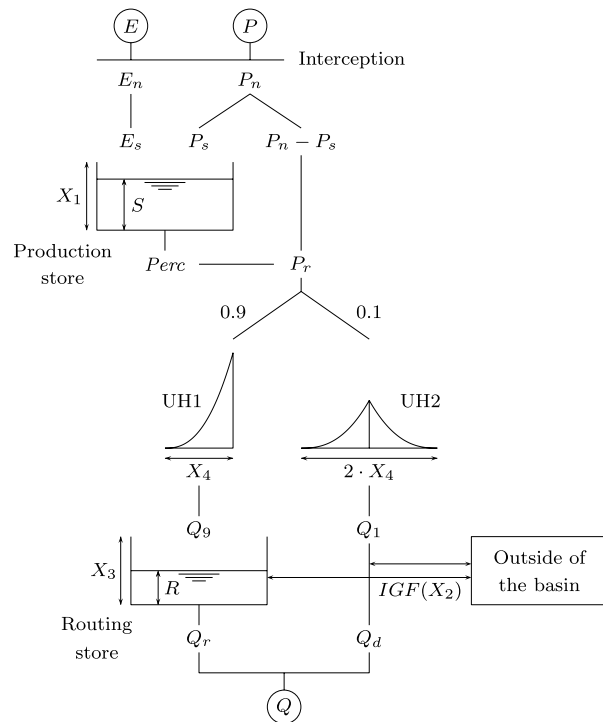


Figure 2. Structure of the GR4J daily time step model. Calibrated parameters are: X₁, capacity of the production reservoir (mm); X₂, intercatchment groundwater flow magnitude (mm/day); X₃, capacity of the non linear routing reservoir (mm); X₄, Unit Hydrograph time base (day). The amount of rainfall converted into net rainfall is a non linear function of the level in the production reservoir; this production is then routed through two Unit Hydrographs and a non linear routing module where the intercatchment groundwater flow (IGF) acts.

W03403

LE MOINE ET AL.: CLOSING THE CATCHMENT-SCALE WATER BALANCE-A CASE STUDY

W03403

Table 2. List of the Parameters of the GR4J Rainfall-Runoff Model and Their Meaning

X_1	capacity of the production reservoir, mm
X_2	intercatchment groundwater flow magnitude, mm
X_3	capacity of the non-linear routing reservoir, mm
X_4	unit hydrograph time base, day

requiring optimization (Table 2). The rest of the structure and its fixed coefficients have been determined through a top-down approach, aimed at identifying a generic structure able to represent the largest possible variety of catchments:

[17] • The production module of the model involves a soil moisture accounting store to determine effective rainfall and an intercatchment groundwater exchange function to simulate water exchanges to/from neighboring catchments or deep aquifers,

[18] • The routing module operates on effective rainfall, which is divided into two flow components. One (quick flow branch) is routed by a unit hydrograph and the other (slow flow branch) by a unit hydrograph and a non-linear routing store.

[19] The model has been extensively tested in several countries and has shown good results in comparison with other rainfall-runoff models [Perrin *et al.*, 2001]. It requires as only inputs rainfall and potential evapotranspiration. Streamflow time series are only needed during the calibration phase, and IGF ‘losses’ or ‘gains’ (which can usually never be measured) are determined as follows.

[20] The daily IGF is calculated as a function of the storage ratio in the routing reservoir of the model. First, a term F is computed as:

$$F = X_2 \cdot (R/X_3)^{7/2} \quad (1)$$

where R is the current level of the routing reservoir and X_3 is its capacity (in mm), and X_2 is a magnitude parameter for IGF (in mm/day). X_2 is either positive (water gains) or negative (water losses) for a given catchment. Then, F is added (algebraically) to both quick flow (Q_1) and slow flow (routing reservoir) components of the model so that the total IGF is:

$$IGF = \begin{cases} 2F & \text{if } F > 0 \text{ (i.e., } X_2 > 0) \\ -\{\text{Min}(|F|; R) + \text{Min}(|F|; Q_1)\} & \text{if } F < 0 \text{ (i.e., } X_2 < 0) \end{cases} \quad (2)$$

[21] Intuitively one might doubt about the possibility to constrain the calibration with two losses: the atmospheric loss (actual evapotranspiration) and the groundwater exchange. It is made possible by the very parsimony of the model which, for example, avoids the scaling of input variables (rainfall and potential evapotranspiration). Even though it is a purely empirical formulation, the validity of the IGF function lies in the extensive testing to which it was submitted. These tests were also supported by the work of *Le Moine et al.* [2007] who compared GR4J’s IGF values with that of another RR model (the SMAR model) which involves such a representation, and found it to be in good agreement for a set of more than 1000 catchments.

4.2. Application of GR4J to Five Leaking River Catchments

[22] The GR4J model was applied to the above mentioned surface catchments, using the standard split sample test scheme [Klemeš, 1986]. It was successively calibrated on the 1980–1992 and 1992–2004 periods with a test in verification mode on the other period each time. For the sake of brevity, we only present the results with the Nash-Sutcliffe (NS) efficiency criterion computed on \sqrt{Q} , which was shown by *Oudin et al.* [2006] to be a good compromise between several alternative criteria:

$$E_{NS} = 1 - \frac{\sum_t (\sqrt{\hat{Q}_t} - \sqrt{Q_t})^2}{\sum_t (E[\sqrt{Q}] - \sqrt{Q_t})^2} \quad (3)$$

where Q_t and \hat{Q}_t are the observed and simulated discharges at timestep t , and $E[\sqrt{Q}]$ is the mean of the square root of the observed discharge. What is more, this criterion gives more weight to no flow periods which are the most difficult part of the hydrograph to reproduce. The values in calibration and verification modes are given in Table 3. Figure 3 presents the simulated hydrograph for the downstream Tardoire catchment, and Table 4 gives the mean annual water balance for the rivers feeding the karstic aquifer. NS criteria as well as hydrographs show that the GR4J model is able to predict the streamflow of non-conservative catchments correctly but we still must look at the consistency of the water balance between adjacent catchments.

5. A Complete Rainfall-Runoff Model for the Touvre Exsurgence

[23] In this section, we show how GR4J’s function accounting for intercatchment groundwater flows can be used to calculate series of hypothetical karstic aquifer recharge, and we compare these annual recharge volumes with the exsurgence module. Then we propose a parsimonious parameterization for representing the karstic reservoir behavior, which finally results in a complete RR representation for the Touvre River.

5.1. Comparing Annual Volumes of Karstic Losses and Spring Flow

[24] Until now, we have used GR4J to analyze the losses of the Tardoire-Bandiat-Bonnieure river system over the La Rochefoucauld karstic plateau. By calibrating GR4J parameters using measured streamflow and precipitation records,

Table 3. Nash-Sutcliffe Performance Criteria^a

Station	Mean Calibration	Mean Verification
Upper Tardoire (1)	0.932	0.922
Upper Tardoire (2)	0.929	0.904
Bandiat	0.899	0.847
Lower Tardoire	0.913	0.911
Bonnieure	0.935	0.917

^aThe mean values refer to the symmetric split-sample test: we perform a calibration on the first period (1980–1992) and compute the verification criterion on the other period (1992–2004), then we switch the roles of the two periods. Thus there are two calibration criteria and two verification criteria and here we show their mean.

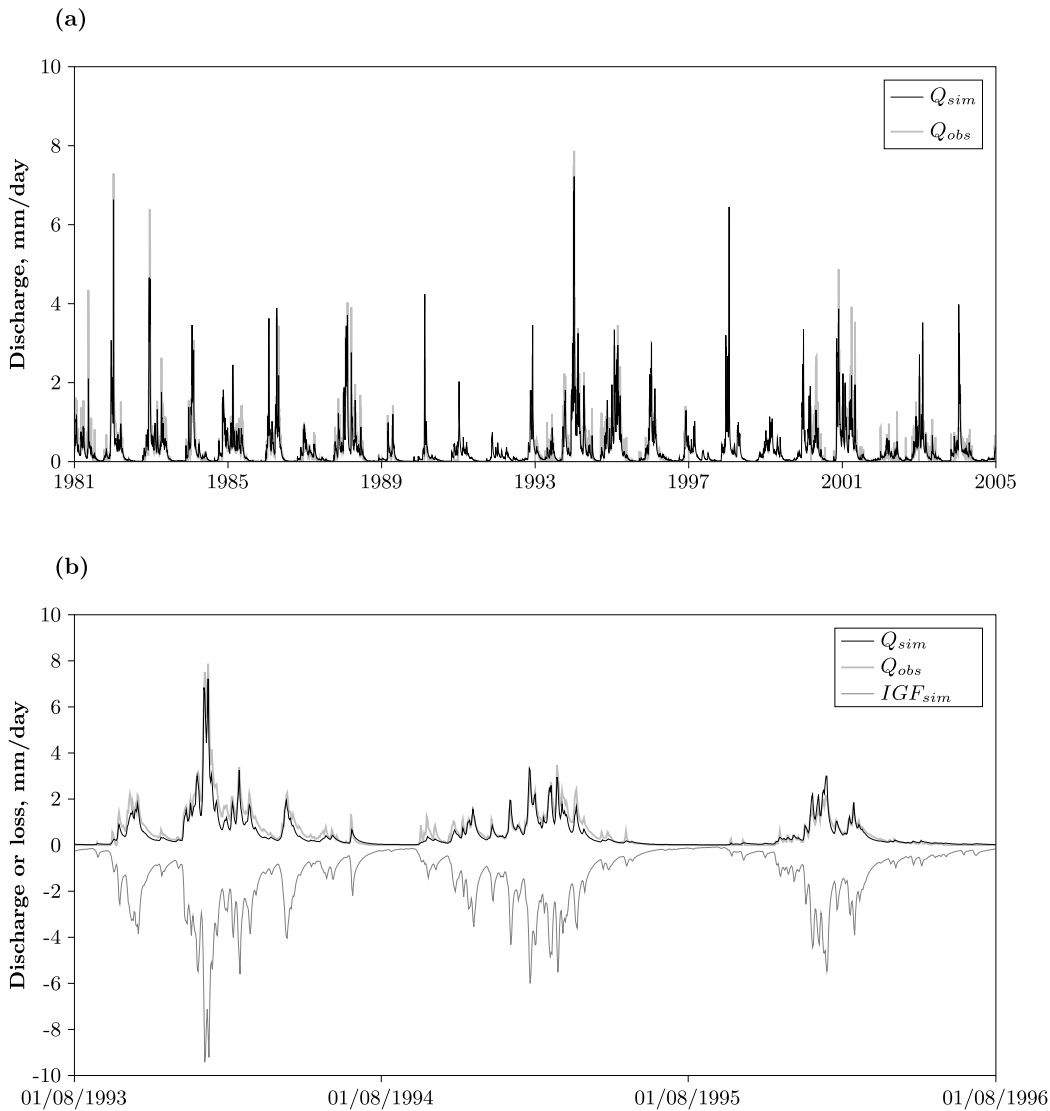


Figure 3. (a) Observed and simulated streamflow for the lower Tardoire (Coulgens gauging site) in verification mode (b) Detail on a three-year period with observed streamflow, simulated streamflow and simulated losses.

we are able to hypothesize estimates of recharge to the karstic reservoir. Let us here stress again that this recharge is computed on the sole basis of the surface stream, rainfall, and runoff time series, independently of any knowledge of exsurgence flows, which can thus be used to provide an independent corroboration of recharge volume estimates. Given mean rainfall $P_m = 945$ mm, on the basis of the average yield of conservative upper catchments of the Tardoire (which is approximately 0.4), we can have an estimation of the non-karstic part of the Touvre River flow (contribution of the Touvre’s own small topographic catchment which is 91 and 126 km² at Ruelle and Foulpougne respectively):

$$Q_{topo} = 0.4 \times P_m \approx 380 \text{ mm} \quad (4)$$

This estimation is quite rough but this term represents less than 10% of the Touvre streamflow: any error would have only a minor influence on the overall water balance. Using the two discharge series available downstream from the

Table 4. Mean Annual Water Balance (in mm) for the Three Rivers Feeding the Karstic Aquifer, Simulated by the GR4J Model^a

Station	P_{obs}	Q_{obs}	PE	Q_{sim}	AE_{sim}	IGF_{sim}
Upper Tardoire (1)	1061	457	756	423	610	-27 ^b
Upper Tardoire (2)	1132	492	756	487	626	-19 ^b
Bandiat	1102	390	756	383	627	-92
Lower Tardoire	1055	131	756	141	616	-298
Bonnieure	955	250	756	249	579	-123

^a P_{obs} : rainfall; Q_{obs} : measured streamflow; PE : potential evapotranspiration; Q_{sim} : simulated streamflow; AE_{sim} : simulated actual evapotranspiration; IGF_{sim} : simulated intercatchment groundwater flow.

^bNot significantly different from zero.

W03403

LE MOINE ET AL.: CLOSING THE CATCHMENT-SCALE WATER BALANCE-A CASE STUDY

W03403

Table 5. Comparison of Estimates of Karstic Aquifer Recharge Based on Different Information: Losses as Estimated by GR4J (RR Analysis, First Column) or Spring Discharge (As Measured Downstream the Exsurgence Site, Second and Third Columns)^a

Estimation Based on Simulated Surface Losses	Estimations Based on the Touvre Spring's Streamflow Measurements				
		At Ruelle		At Foulpougne	
	Water × Surface Depth (10 ⁶ m ²) (10 ⁻³ m)		Water × Depth Surface (10 ⁶ m ²) (10 ⁻³ m)		Water × Depth Surface (10 ⁶ m ²) (10 ⁻³ m)
Tardoire	298 × 1200	Runoff	4644 × 91	Runoff	3347 × 126
Bonnieure	123 × 203	(topo)	-380 × 91	(topo)	-380 × 126
	383 10⁶ m³/a		388 10⁶ m³/a		374 10⁶ m³/a

^aThe contribution of the Touvre's topographic catchment ($Q_{topo} \simeq 380$ mm) is estimated according to mean rainfall $P_m = 945$ mm and a runoff ratio set to 0.4.

spring, we can estimate the pure exsurgence flow (see calculations in Table 5). On the other hand, we can use the water balance computations summed up in Table 4 to estimate the actual volume lost from the surface system. Indeed, according to numerous studies [see, e.g., *Queleuennec et al.*, 1971; *Laroque and Razack*, 1998; *Laroque et al.*, 1998; *de Grissac et al.*, 2006], the recharge area of the spring is located within the limits of the lower Tardoire and Bonnieure topographic catchments. The results of the two methods match quite well:

[25] • The method using the Touvre River discharges gives a value of overall inflow between 374 and 388 × 10⁶ m³/a,

[26] • The GR4J model gives an overall loss of 383 × 10⁶ m³/a over the assumed recharge area (topographic catchments of the lower Tardoire and Bonnieure).

[27] These values are very close to the one given by *de Grissac et al.* [2006], which is 400 × 10⁶ m³/a. In this particular case, the GR4J model seems to have computed a realistic distribution of the water balance between atmospheric and groundwater losses. These good results allow us to try a more complex simulation to simulate the spring flow from rainfall over its actual catchment. The recharge (surface loss) series cannot give us the spring discharge at a daily time step directly because of the routing process through the karstic system. However, assuming a simple model for this reservoir, we will show that we can reproduce the daily dynamics of the spring discharge.

5.2. A Simple Mathematical Structure Describing the Storage and Depletion of the Karstic Aquifer

[28] Here we have chosen to test what we believe to be at the same time a very simple and efficient non-linear routing model, the lag-and-route model [*Linsley et al.*, 1949] using a single reservoir and three parameters:

[29] • the power of reservoir depletion p ,

[30] • the constant of reservoir depletion α , so that the cumulated discharge over a time step is given by the integration of:

$$Q = -\frac{dK}{dt} = \alpha K^p \quad (5)$$

where K is the current level in the reservoir,

[31] • a pure delay parameter D (in days).

5.3. A Complete RR Simulation Model of the Touvre Spring

[32] Let us now recapitulate the structure of the rainfall-runoff model representing the recharge and discharge pro-

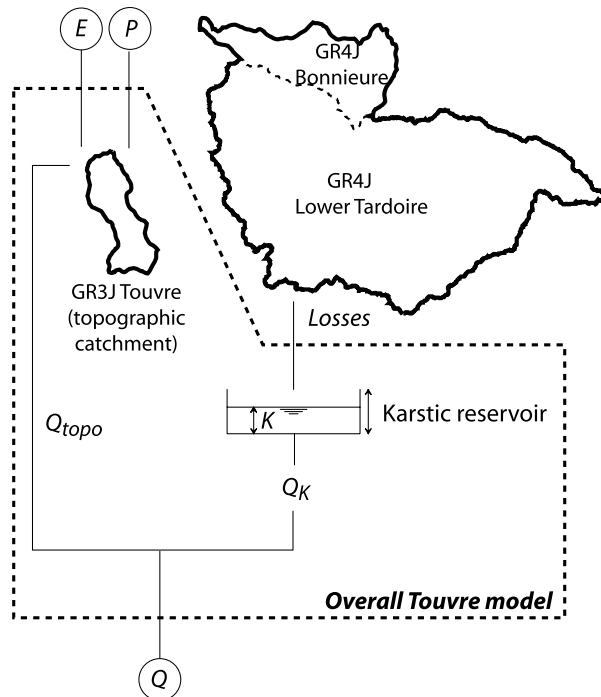


Figure 4. Structure of the six-parameter model derived from GR4J to model the Touvre streamflow. Calibrated (sensitive) parameters are: exponent p of karstic reservoir depletion law; constant of karstic reservoir depletion (time decay τ if $p = 1$); three parameters for the Touvre River's small topographic catchment (GR4J model degraded to conservative mode, i.e., imposing $X_2 = 0$).

Table 6. Nash-Sutcliffe Performance Criteria for the Karstic Model^a

Exponent	$p = 0.5$	$p = 1$	$p = 1.5$	$p = 2$	$p = 2.5$	$p = 3$
Calibration 1982–1989	0.876	0.876	0.840	0.868	0.866	0.865
Verification 1989–1996	0.724	0.737	0.760	0.733	0.731	0.729
Calibration 1989–1996	0.756	0.769	0.781	0.768	0.767	0.766
Verification 1982–1989	0.834	0.852	0.824	0.839	0.836	0.833
Average calibration	0.816	0.823	0.811	0.818	0.817	0.816
Average verification	0.779	0.795	0.792	0.786	0.784	0.781

^aThe best results are obtained for $p = 1$ which corresponds to a linear reservoir.

W03403

LE MOINE ET AL.: CLOSING THE CATCHMENT-SCALE WATER BALANCE-A CASE STUDY

W03403

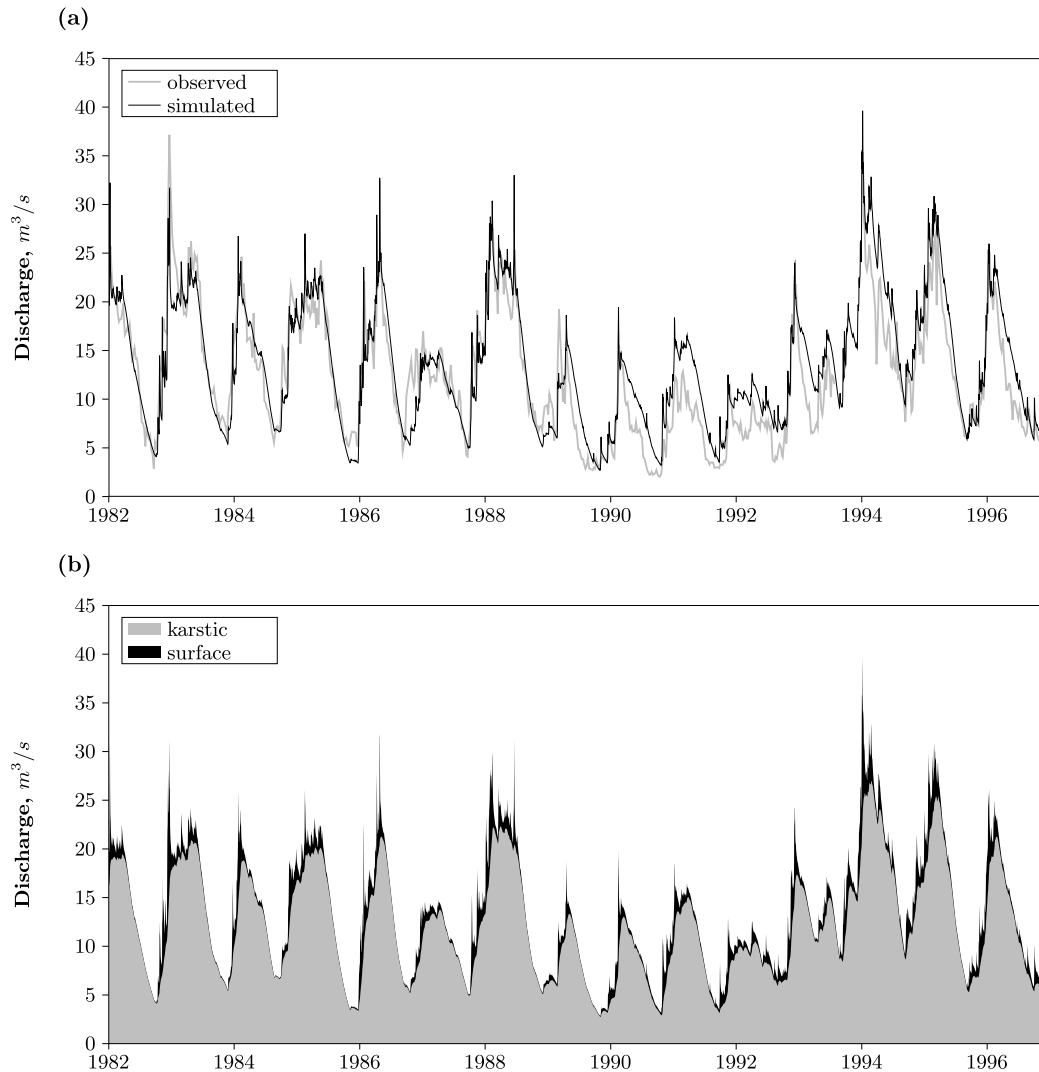


Figure 5. (a) Observed streamflow (thick gray) and simulated streamflow (thin black) for the Touvre discharge at Ruelle. The simulated hydrograph can be shown in verification mode for the whole period where data is available (1982–1996) because we use a symmetric split-sample test procedure over 1982–1989 and 1989–1996 (see also Table 6). (b) Flow components of the Touvre River discharge at Ruelle as computed by the integrated model. Surface flow refers to the contribution of the river’s own small topographic catchment: this contribution cannot be neglected during peak flows.

cess of the Touvre karstic system in Figure 4. Several values of exponent p were tested for the karstic reservoir and the best value was $p = 1$. Exponent $p = 0.5$, which would be a priori suggested by conduit flow hydraulics [Bonacci, 2001], gives good results for high flows but is less satisfactory for recessions. On the other hand, a linear reservoir seems to be compliant with the presumed dynamics of water release by the epikarst during summer [Laroque and Razack, 1998; Laroque et al., 1998]. The model has quite good predictive capabilities, as shown by the results in verification mode (Table 6). Hydrographs are plotted in Figure 5 and observed versus simulated streamflow is presented in Figure 6. Thus we have been able to propose a parsimonious rainfall-flow model for the complete karstic system, including:

[33] • A recharge module calculating continuous time series of karstic reservoir input, based on estimating the intercatchment groundwater flows (IGF). It was calibrated on the discharge values of the Tardoire and Bonnieure rivers after they passed the karstic plateau where they contribute to aquifer recharge, and it requires eight parameters (four for each river).

[34] • A storage model, which transforms these inputs into a time series of karstic flow components for the spring. The behavior of the aquifer is mainly characterized by a linear reservoir ($p = 1$ in equation (5)) with a characteristic time $\tau = \frac{1}{\alpha} = 46$ days. The pure delay D in the lag-and-route procedure was found to be very small (less than one day) and not sensitive. It was set to $D = 0$.

W03403

LE MOINE ET AL.: CLOSING THE CATCHMENT-SCALE WATER BALANCE-A CASE STUDY

W03403

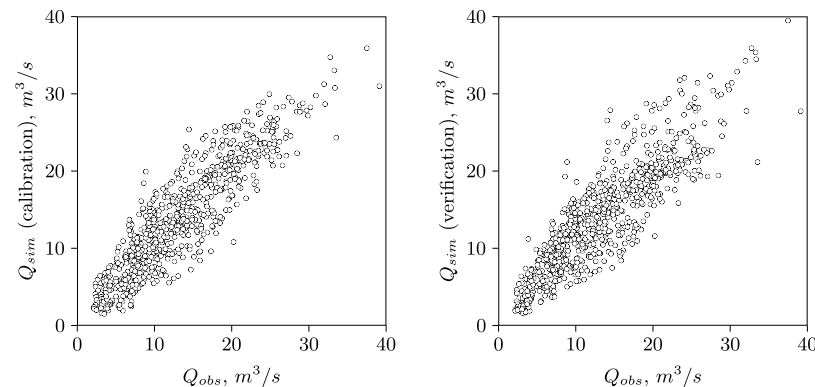


Figure 6. Observed versus simulated discharge for the Touvre at Ruelle.

[35] • A surface module accounting for the surface flow component of the Touvre discharge, using three parameters (GR4J model degraded to conservative mode).

6. Conclusion

[36] The aim of this article was to present a unique case of karstic catchment, where we were able to estimate the recharge and the flow at the exsurgence independently. Comparing daily time series of recharge input and exsurgence flow made it possible to look for an adequate representation of the behavior of the karstic reservoir. On the basis of several tests, we found that a lag-and-route model with a linear store was the best option. This study was also able to set up a full RR model for the Touvre exsurgence. From a RR modeler's point of view, the good match between calculated losses (based on our vision of the RR behavior of the surface catchments, using climatic forcings) and measured spring flow is encouraging. Even if the empirical function regulating the intercachment groundwater flow in GR4J has been tested and discussed extensively [Le Moine et al., 2007], a true verification requires a comparison with measurements. Let us acknowledge, however, that a single case study is obviously not sufficient to validate GR4J's structure: many more similar cases would be needed. For now, it is satisfactory to show that in this precise case, GR4J leakage estimates have been realistic, and that a parsimonious complete RR model of the Touvre exsurgence was set up with only six free parameters to calibrate, providing a realistic and efficient simulation tool to test the ability of the Touvre system to withstand a variety of drought scenarios.

[37] From a more general point of view, this work also emphasizes the importance of surface water measurements as signals to estimate groundwater recharge, not only its fraction in the annual water balance, but also its dynamics: this flux between surface- and groundwater may often look like a succession of episodes, more or less correlated to rainfall and surface streams, rather than like a smooth seasonal curve like those used by hydrogeologists to define the boundary conditions of groundwater models.

[38] **Acknowledgments.** We would like to acknowledge the help of Jean Margat (BRGM) and Stéphane Rénié (Hydro Invest engineering company) for all hydrogeological information, Marie-José Garcia (DIREN Poitou-Charente/SEMA) and Pierre Delavallade (Président de l'Association

du Grand Karst de La Rochefoucauld) for the streamflow data of the Touvre at Foulpouagne and Ruelle-sur-Touvre, and Météo France for the pluviographic archive. We would also like to thank the reviewers for their constructive critics and comments on this paper.

References

- Ababou, R., and A. Mangin (1999), Linear and nonlinear input/output models for karstic springflow and flood prediction at different time scales, *Stochastic Environ. Res. Risk Assessment*, 13(5), 337–364.
- Atkinson, T. C. (1977), Diffuse flow and conduit flow in limestone terrain in Mendip-Hills, Somerset (Great Britain), *J. Hydrol.*, 35, 93–110.
- Barrett, M. E., and R. J. Charbeneau (1997), A parsimonious model for simulating flow in karst aquifer, *J. Hydrol.*, 196, 47–65.
- Bonacci, O. (2001), Analysis of the maximum discharge of karst springs, *Hydrogeology J.*, 9, 328–338.
- Box, G. E. P., and G. M. Jenkins (1976), *Time Series Analysis: Forecasting and Control*, rev. ed., Holden-Day, San Francisco, CA.
- Chiew, F. H. S., and T. A. McMahon (1990), Estimating groundwater recharge using a surface watershed modelling approach, *J. Hydrol.*, 114, 285–304.
- de Grissac, B., S. Rénié, and J.-C. Roux (2006), Sources de la Touvre et karst de la Rochefoucauld, in *Aquifères et eaux souterraines en France*, pp. 458–460, BRGM-AIH.
- Eisenlohr, L., M. Bouzelboudjen, L. Kiraly, and Y. Rossier (1997a), Numerical versus statistical modeling of natural response of a karst hydrogeological system, *J. Hydrol.*, 202, 244–262.
- Eisenlohr, L., L. Kiraly, M. Bouzelboudjen, and Y. Rossier (1997b), A numerical simulation as a tool for checking the interpretation of karst springs hydrographs, *J. Hydrol.*, 193, 306–315.
- Goswami, M., and K. M. O'Connor (2005), Application of a conceptual rainfall-runoff simulation model to three European catchments characterised by non-conservative system behaviour, *International Conference on Hydrological Perspectives for Sustainable Development (HYPESD-'05)*, Indian Institute of Technology Roorkee, India.
- Halihan, T., C. M. Wicks, and J. F. Engeln (1998), Physical response of a karst drainage basin to flood pulses: Example of the Devil's Icebox cave system (Missouri, USA), *J. Hydrol.*, 204, 24–36.
- Hughes, D. A. (2004), Incorporating groundwater recharge and discharge functions into an existing monthly rainfall-runoff model, *Hydrol. Sci. J.*, 49(2), 297–311.
- Johannet, A., A. Mangin, and D. D'Hulst (1994), Subterranean water infiltration modelling by neural networks: use of water source flow, in *ICANN '94 - Proceedings of the International Conference on artificial neural networks, Sorrento, Italy, 26–29 May 1994*, pp. 1033–1036, Springer-Verlag, London.
- Jukic, D., and V. Denic-Jukic (2006), Nonlinear kernel functions for karst aquifers, *J. Hydrol.*, 323(1–2), 360–374.
- Katz, B. G., J. S. Catches, T. D. Bullen, and R. L. Michel (1998), Changes in isotopic and chemical composition of ground water resulting from a recharge pulse from a sinking stream, *J. Hydrol.*, 211, 178–207.
- Klemeš, V. (1986), Operational testing of hydrological simulation models, *Hydrol. Sci. J.*, 31(1), 13–24.

W03403

LE MOINE ET AL.: CLOSING THE CATCHMENT-SCALE WATER BALANCE-A CASE STUDY

W03403

- Labat, D., R. Ababou, and A. Mangin (1999), Linear and nonlinear input/output models for karstic spring flow and flood prediction at different time scales, *Stoch. Env. Res. Risk A.*, 13(5), 337–364.
- Labat, D., R. Ababou, and A. Mangin (2000a), Rainfall-runoff relations for karstic springs. Part I: Convolution and spectral analyses, *J. Hydrol.*, 238(3–4), 123–148.
- Labat, D., R. Ababou, and A. Mangin (2000b), Rainfall-runoff relations for karstic springs. Part II: Continuous wavelet and discrete orthogonal multiresolution analyses, *J. Hydrol.*, 238(3–4), 149–178.
- Laroque, M., and M. Razack (1998), Hydrogéologie du karst de La Rochefoucauld (Charente): Bilan des nouvelles connaissances, *Hydrogéologie*, 3, 35–45.
- Laroque, M., A. Mangin, M. Razack, and O. Banton (1998), Contribution of correlation and spectral analyses to the regional study of a large karst aquifer (Charente, France), *J. Hydrol.*, 205, 217–231.
- Le Moine, N., V. Andréassian, C. Perrin, and C. Michel (2007), How can rainfall-runoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments, *Water Resour. Res.*, 43(6), W06428, doi:10.1029/2006WR005608.
- Linsley, R. K., M. A. Kohler, and J. L. H. Paulus (1949), *Applied Hydrology*; McGraw-Hill, Inc., New York, NY.
- Mangin, A. (1981), Apport des analyses corrélatoire et spectrale dans l'approche des systèmes hydrologiques, *C. R. Acad. Sci. Paris, Série II*, 293, 401–404.
- Mangin, A. (1984), Pour une meilleure connaissance des systèmes hydrologiques à partir des analyses corrélatoires et spectrales, *J. Hydrol.*, 67, 25–43.
- Margat, J. (2008), *Les eaux souterraines, une richesse mondiale*, UNESCO Press, Paris.
- Mouelhi, S., C. Michel, C. Perrin, and V. Andréassian (2006), Stepwise development of a two-parameter monthly water balance model, *J. Hydrol.*, 318(1–4), 200–214.
- Nascimento, N. O. (1995), Appréciation à l'aide d'un modèle empirique des effets d'action anthropiques sur la relation pluie-débit l'échelle du bassin versant, Ph.D. thesis, ENPC, Paris.
- Oudin, L., V. Andréassian, T. Mathevet, C. Perrin, and C. Michel (2006), Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations, *Water Resour. Res.*, 42(7), W07410, doi:10.1029/2005WR004636.
- Padilla, A., and A. Pulido-Bosch (1995), Study of hydrographs of karstic aquifers by means of correlation and cross-spectral analysis, *J. Hydrol.*, 168(1–4), 73–89.
- Perrin, C., C. Michel, and V. Andréassian (2001), Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments, *J. Hydrol.*, 242(3–4), 275–301.
- Perrin, C., C. Michel, and V. Andréassian (2003), Improvement of a parsimonious model for streamflow simulation, *J. Hydrol.*, 279, 275–289.
- Quellenec, R. E., J. C. Sauret, M. Seguin, and J. Voue (1971), Les résurgences de la Touvre – Étude préliminaire, in *Actes du Colloque d'Hydrologie en Pays Calcaire, Besançon*, Annales Scientifiques de l'Université de Besançon, fascicule 15, 3ème série, Géologie, 1971.
- Rimmer, A., and Y. Salingar (2006), Modelling precipitation-streamflow processes in karst basin: The case of the Jordan River sources, Israel, *J. Hydrol.*, 331, 524–542.
- Yevjevich, V. (1972), *Stochastic Processes in Hydrology*, Water Research Publications, Fort Collins, CO.

V. Andréassian and N. Le Moine, CEMAGREF, Hydrosystems and Bioprocesses Research Unit, Parc de Tourvoie, BP 44, 92163 Antony Cedex, France. (nicolas.le-moine@cemagref.fr)

T. Mathevet, Électricité de France, Direction Technique Générale, Département Surveillance Service CADE, BP 41, 38040 Grenoble Cedex 9, France.

6

L'approche semi-distribuée comme test de la cohérence interne des modèles

6.1 Principe de l'étude

Nous avons expliqué au chapitre 3 que l'évaluation des modèles testés dans cette thèse se concentrerait sur leurs performances en simulation des débits de surface, aussi bien pour des raisons pratiques (la vocation de ces modèles est la prédétermination, simulation ou prévision des débits) que théoriques, puisqu'il n'y a pas d'autres flux mesurés — ni mesurables — à l'échelle du système retenu, le bassin versant. Malgré tout, dans la « quête » probablement sans fin d'une fermeture réaliste du bilan en eau, nous pouvons suivre les recommandations de *Beven* (2006) et exiger des solutions retenues qu'elles satisfassent, autant que possible, certains principes physiques. Le premier d'entre eux étant l'équation de continuité sur les volumes de contrôle choisis, nous pouvons contruire un test pour évaluer la cohérence des modèles sur ce point. Ce test consiste à faire varier les volumes de contrôle en choisissant différents découpages topographiques pour une zone donnée : il consiste à faire de la semi-distribution.

Une façon simple de présenter l'approche est de considérer une confluence entre deux rivières (Figure 6.1). Si nous disposons d'une station hydrométrique sur chacune des branches à l'amont de la confluence, ainsi qu'une autre juste à l'aval, nous pouvons définir trois systèmes (*Baudez*, 1997) : le bassin versant aval, noté B_0 , et les deux bassins versants amont notés B_1 et B_2 .

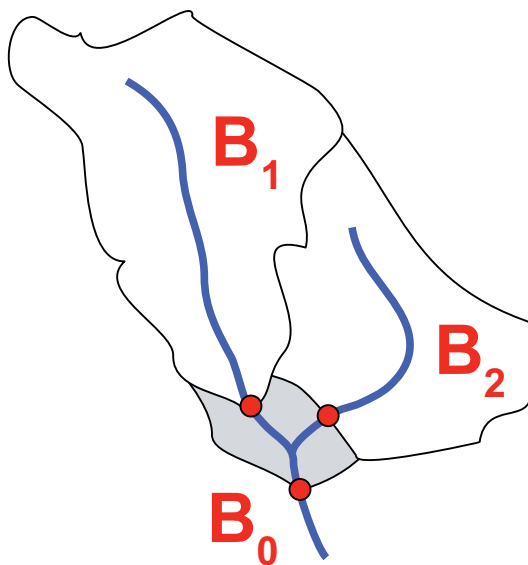


FIG. 6.1: Exemple de triplet de stations hydrométriques associé à une confluence.

En notant respectivement A_0 , A_1 et A_2 les surfaces de ces bassins versants ($A_0 \simeq A_1 + A_2$), nous pouvons donc écrire trois équations de continuité en volume au cours d'un pas de temps (les termes de pluie P , d'évapotranspiration réelle E , de débit Q , d'échange souterrain

G et de variation de stock ΔS étant exprimés en hauteur d'eau) :

$$A_1 \cdot (P_1 + G_1) = A_1 \cdot (Q_1 + E_1 + \Delta S_1)$$

$$A_2 \cdot (P_2 + G_2) = A_2 \cdot (Q_2 + E_2 + \Delta S_2)$$

$$A_0 \cdot (P_0 + G_0) = A_0 \cdot (Q_0 + E_0 + \Delta S_0)$$

Or, étant donné que $B_0 = B_1 \cup B_2$ (en négligeant la zone intermédiaire), la somme des deux premières équations doit être égale à la troisième. On a, par définition de la pluie de bassin, $A_0 \cdot P_0 = A_1 \cdot P_1 + A_2 \cdot P_2$ et à la confluence $A_0 \cdot Q_0 = A_1 \cdot Q_1 + A_2 \cdot Q_2$. En considérant un pas de temps suffisamment long pour permettre la propagation des variations de stock à l'intérieur du système B_0 (i.e., écrire $A_0 \cdot \Delta S_0 \simeq A_1 \cdot \Delta S_1 + A_2 \cdot \Delta S_2$), nous obtenons la condition :

$$A_1 \cdot E_1 + A_2 \cdot E_2 - A_1 \cdot G_1 - A_2 \cdot G_2 = A_0 \cdot E_0 - A_0 \cdot G_0$$

En toute rigueur, nous ne pouvons pas aller plus loin et nous retombons à nouveau sur le problème qui nous occupe : on ne peut contraindre que le résultat net de l'évapotranspiration et des flux souterrains. En effet, rien n'interdit qu'une partie de l'infiltration profonde sur le bassin B_1 s'écoule vers le bassin B_2 et y soit en partie perdue par évapotranspiration. Laisant ces cas extrêmes de côté, nous écrirons raisonnablement que les conditions à vérifier sont :

$$\begin{cases} A_1 \cdot E_1 + A_2 \cdot E_2 = A_0 \cdot E_0 \\ A_1 \cdot G_1 + A_2 \cdot G_2 = A_0 \cdot G_0 \end{cases} \quad (6.1)$$

Ce système fournit une possibilité pour d'évaluer le réalisme physique des flux calculés : il s'agit ici d'un critère de cohérence interne. Ce que nous souhaitons vérifier, c'est qu'un modèle parvenant à simuler correctement le débit aux trois stations d'une telle configuration, ne le fait pas grâce à des artifices¹ : par exemple en prédisant une évaporation nulle sur chacun des bassins B_1 et B_2 (tout le déficit d'écoulement étant simulé par des pertes souterraines), et une infiltration profonde nulle sur le système global B_0 . Cela serait la preuve flagrante que ces deux termes (infiltration profonde et évapotranspiration réelle) ne correspondent en rien à ce que nous désignons par les mêmes noms dans le modèle².

Insistons sur un point : le but ici n'est pas de tester l'intérêt de la semi-distribution pour simuler le débit Q_0 de la station aval (ce résultat est annexe). Cette étude a été faite par *Baudez* (1997) au pas de temps journalier. Ici, il s'agit de poursuivre à un pas de temps plus fin la logique de test de cohérence interne initiée au pas de temps mensuel (*Le Moine*, 2005).

¹Ou du moins, pas grâce à des artifices trop grossiers...

²Rappelons une nouvelle fois l'asymétrie fondamentale de ce type de démonstration : nous pouvons démontrer que le modèle est faux, en aucun cas le valider.

6.2 Le bassin intermédiaire : un système à plusieurs entrées

Dans la situation que nous venons de présenter, chaque maille (B_1 et B_2 d'un côté, B_0 de l'autre) reste délimitée par une ligne de crête, conformément à l'approche que nous avons choisie et expliquée au chapitre 1. Cependant, le fait de se limiter à l'amont et à l'aval immédiats de points de confluence est assez contraignant pour deux raisons :

- les cas où l'on a ce genre de triplets de stations hydrométriques sont rarissimes (un quinzaine dans *Baudez, 1997* et *Le Moine, 2005*),
- d'autre part, nous pourrions rarement définir plus de deux mailles par bassin aval : il est rare d'avoir une triple confluence, ou deux confluences de même ordre très rapprochées...

Nous proposons donc de généraliser le découpage précédent en acceptant, d'une part, que la zone intermédiaire ait une surface importante, et d'autre part en la traitant comme un sous-bassin versant classique dans le modèle. La seule différence sera l'existence d'entrées supplémentaires, constituées par les débits amont entrant dans cette zone intermédiaire. Cette hypothèse un peu hardie se justifie malgré tout par le fait que la nature de ce *bassin intermédiaire* n'est pas radicalement différente de celle d'un bassin versant au sens strict (*Lerat et al., soumis*). C'est aussi un ensemble de points de la surface topographique convergeant vers une même section : simplement, ce n'est pas *l'ensemble* de ces points. La difficulté sera alors de savoir comment le débit amont interagit avec la production du bassin intermédiaire.

6.3 Choix des combinaisons de stations

Nous avons sélectionné, dans l'échantillon de 1040 stations hydrométriques, toutes celles qui ont à leur tour au moins une station à l'amont. La Figure 6.2 illustre la façon dont on construit le maillage topographique pour la semi-distribution : pour une station donnée, on utilise un certain nombre de stations amont comme points d'entrée, ce qui définit un bassin versant intermédiaire. Considérons une station aval S_0 , comportant n stations amont (S_1, \dots, S_n) sur des branches distinctes du linéaire (i.e., aucune de ces n stations n'est à l'amont de l'une des $n - 1$ autres, elles ont toutes pour plus proche station aval S_0). Si nous considérons que ces n stations sont des têtes de bassin (« non-découpables », pas de stations plus loin à l'amont), on obtient facilement le nombre $B(S_0)$ de bassins intermédiaires (i.e., possédant au moins un débit amont injecté) dont l'exutoire est S_0 :

$$B(S_0) = 2^n - 1$$

2^n étant le nombre de combinaisons des n stations amont, auquel il faut enlever le cas « aucune station amont », qui correspond à l'approche globale en S_0 .

Si cette station S_0 comporte en outre à l'amont une station S_{n+1} sur laquelle on peut définir $B(S_{n+1})$ bassins intermédiaires dont elle est l'exutoire, le nombre $B(S_0)$ devient :

$$B(S_0) = [2^n \times (B(S_{n+1}) + 2)] - 1$$

Il faut en effet ajouter à $B(S_{n+1})$ les cas où le bassin versant global défini en S_{n+1} est inclus ou exclu en entier. Cet exemple nous permet de définir par une relation de récurrence le nombre de bassins versants intermédiaires dont l'exutoire est S_0 :

$$B(S_0) = \left[\prod_{k=1}^n (B(S_k) + 2) \right] - 1$$

où les S_k sont les stations immédiatement à l'amont de S_0 en remontant les différentes branches du linéaire (avec bien sûr $B(S_k) = 0$ si S_k est une station « non-découpable », sans station amont jaugée).

Pour le cas de la Figure 6.2, la station n°3 possède deux stations amont « non-découpables », on peut donc lui associer $B(S_3) = 2^2 - 1 = 3$ bassins intermédiaires. La station n°6 possède deux stations amont non-découpables (numérotées 4 et 5) et une découpable, la station n°3. On peut donc définir $B(S_6) = 2^2 \times (B(S_3) + 2) - 1 = 19$ bassins versants intermédiaires dont elle est l'exutoire. Vu l'explosion du nombre de combinaisons, on se limitera à une règle simple : dans les cas où des stations amont sont disponibles pour une station donnée, on prend en compte toutes celles immédiatement à l'amont en remontant les différentes branches du linéaire, sans en éliminer aucune ; on définit ainsi les bassins versants intermédiaires les plus petits possibles associés à chaque station (pour un réseau donné). Nous avons malgré tout tenu à présenter la façon de définir tous les maillages topographiques possibles associés à chaque station, pour bien mettre en évidence l'ampleur potentielle de ces tests : bien qu'ils ne soient pas tous indépendants, le nombre de maillages définissables croît exponentiellement avec la taille du réseau de stations hydrométriques, offrant un nombre quasi-illimité de configurations dans lesquelles comparer des bilans globaux et semi-distribués.

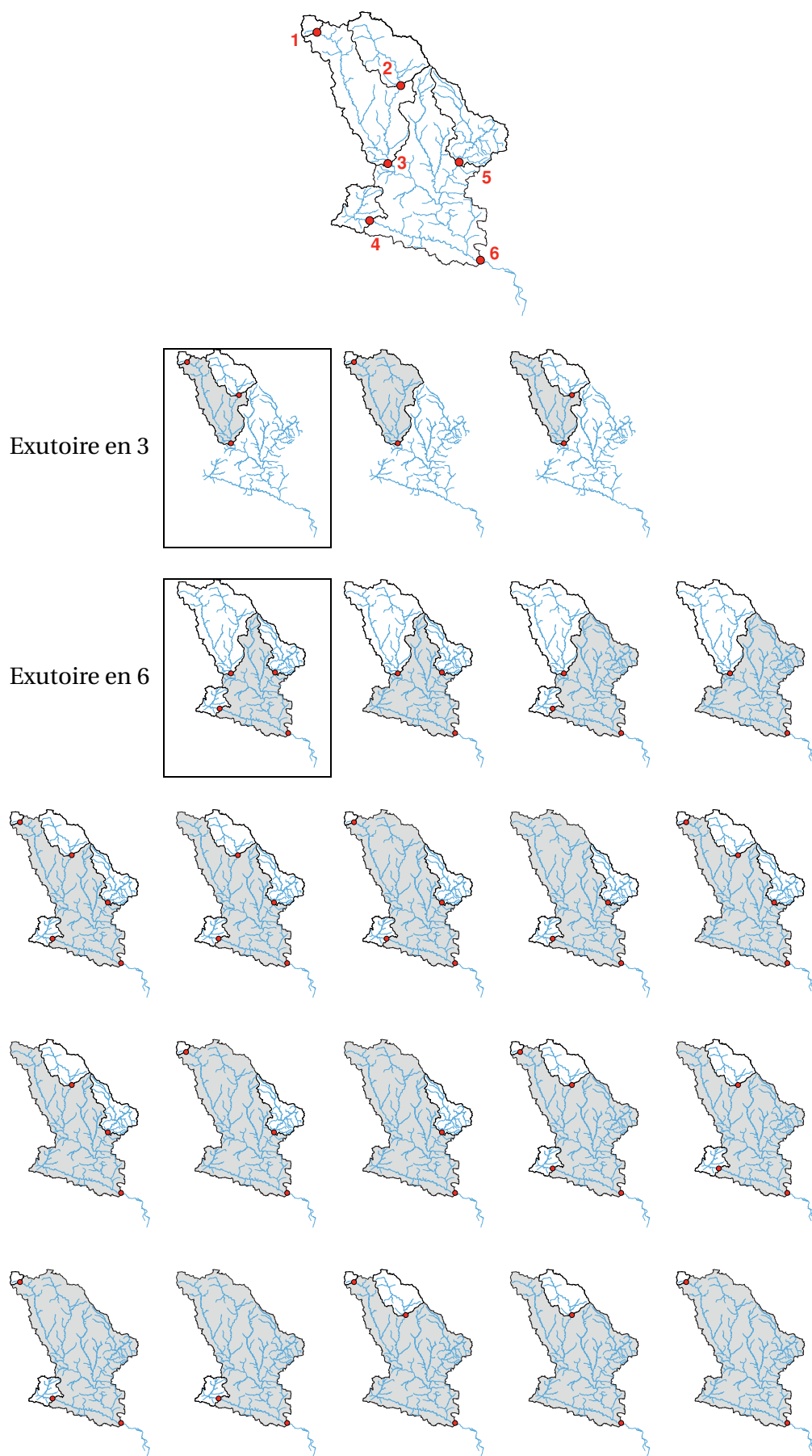


FIG. 6.2: Exemple de construction du maillage hydrologique utilisé en semi-distribution : ensemble de 6 stations (haut) et bassins versants intermédiaires possibles (bas). Les encadrés correspondent aux seuls cas retenus.

Sur l'échantillon complet, on compte 284 stations possédant au moins une station à l'amont, auxquelles on associera donc 284 bassins versants intermédiaires « minimaux » (cf. Figure 6.3). Le but de ce découpage est d'évaluer la cohérence du bilan en eau calculé globalement sur un bassin donné (approche classique), avec le bilan agrégé sur le bassin versant intermédiaire et les bassins amont jaugés qui le composent. Cette vérification est importante si l'on a pour ambition de développer un modèle capable de représenter la relation pluie-débit, avec une composante souterraine, à des échelles variées (à partir de quelques km² jusqu'à l'échelle de très grands bassins). En effet, l'équation de continuité impose que les bilans calculés soient proches dans les deux cas.

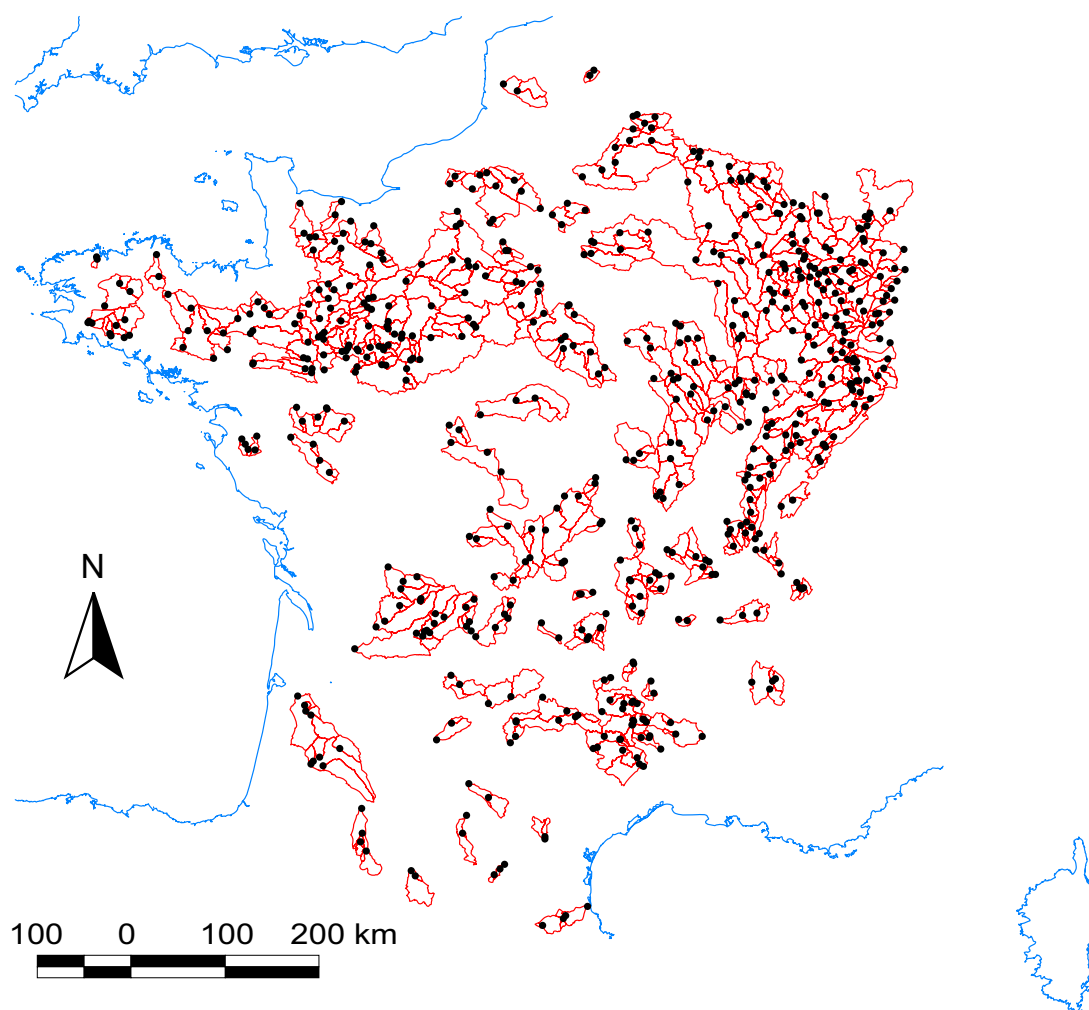


FIG. 6.3: Localisation des 630 stations « emboîtées » utilisées pour le test de semi-distribution. Ces 630 stations incluent (a) 284 stations possédant au moins une station à leur amont, auxquelles on définit 284 bassins intermédiaires, et (b) 346 stations « têtes de bassins ».

6.4 Construction des pluies et ETP sur les bassins intermédiaires

Les grandeurs intensives comme la pluie ou l'ETP de bassin sont calculées simplement à partir de ces mêmes grandeurs sur les bassins globaux (méthode présentée au chapitre 2). Par construction des pluies de bassin SAFRAN, la pluie sur un bassin intermédiaire comportant N débits amont vérifie :

$$A_0 \cdot P_0 = A_{\text{int}} \cdot P_{\text{int}} + \sum_{k=1}^N A_k \cdot P_k$$

$$P_{\text{int}} = \frac{1}{A_{\text{int}}} \left(A_0 \cdot P_0 - \sum_{k=1}^N A_k \cdot P_k \right)$$

L'ETP de bassin intermédiaire vérifie la même équation. Quant aux débits entrant et sortant d'un bassin intermédiaire, on les norme tous par la surface de ce bassin intermédiaire.

6.5 Adaptation du modèle, méthodes de calage et de validation

Une fois le maillage défini, on cale le modèle normalement sur chaque bassin versant intermédiaire, à la différence près que l'on a des entrées supplémentaires qui sont les débits amont (en fonction des cas nous utiliserons soit les débits amont observés, soit les débits simulés). On utilise donc un modèle de propagation simple, conceptuel, à deux paramètres, quel que soit le nombre de stations amont : une vitesse de transfert V uniforme et une forme d'hydrogramme unitaire pour le routage de ces débits amont (plus ou moins pointue en fonction de l'exposant p , calé). Pour chaque station amont k , on calcule le temps de base T_k de l'hydrogramme unitaire à partir de sa distance hydraulique D_k à l'exutoire : $T_k = \frac{D_k}{V}$. La distance D_k est calculée d'après un modèle numérique de terrain (MNT) au pas de 75 m et est donc dépendante de cette échelle d'analyse, mais cela est sans importance puisque de toute façon nous calons la vitesse V . Chaque bassin intermédiaire possède donc six paramètres : les quatre du modèle GR4J, auxquels on ajoute deux paramètres de routage des débits amont. On ajoute les débits amont sur la branche de routage direct, avant que ne s'applique le terme d'échange F : cela permet à la fonction d'échange du bassin intermédiaire d'agir également, le cas échéant, sur les débits amont (cas de pertes par le lit par exemple, voir *Costelloe et al.*, 2003). Le routage des débits amont par des hydrogrammes unitaires mériterait incontestablement une fondation plus solide, et l'on gagnerait à tester

différentes hypothèses telles que celles proposées par *Lerat et al.* (soumis) : *lag-and-route* (*Linsley et al.*, 1949), modèle de *Hayami* (1951) etc. Cependant, une solution de routage par hydrogramme unitaire, plutôt que par un réservoir conceptuel, présente l'avantage pratique de pouvoir gérer assez simplement les lacunes sur les débits amont : en effet, chaque journée de lacune sur un débit amont QAM_k n'a d'effet que sur les $2T_k$ pas de temps suivants (durée de l'hydrogramme unitaire associé). Pendant cette durée, on ne calcule pas de Nash sur les débits simulés à l'exutoire du bassin intermédiaire puisqu'il manque des apports, mais les états du modèle de bassin intermédiaire continuent à être calculés normalement.

La structure du modèle de bassin intermédiaire est schématisée sur la Figure 6.4.

Plusieurs possibilités sont envisageables pour caler puis contrôler le modèle sur chaque bassin intermédiaire. Une seule méthode est retenue pour le calage : elle consiste à utiliser les débits amont observés, convertis en mm sur le bassin intermédiaire. Pendant les périodes de contrôle, deux solutions s'offrent à nous :

- soit continuer à utiliser les débits observés, dans ce cas on valide simplement à l'échelle du bassin intermédiaire,
- soit utiliser les débits amont simulés en contrôle, c'est-à-dire qu'on se place dans le cas d'une validation globale « en cascade » (aucun débit intermédiaire n'est alors supposé connu).

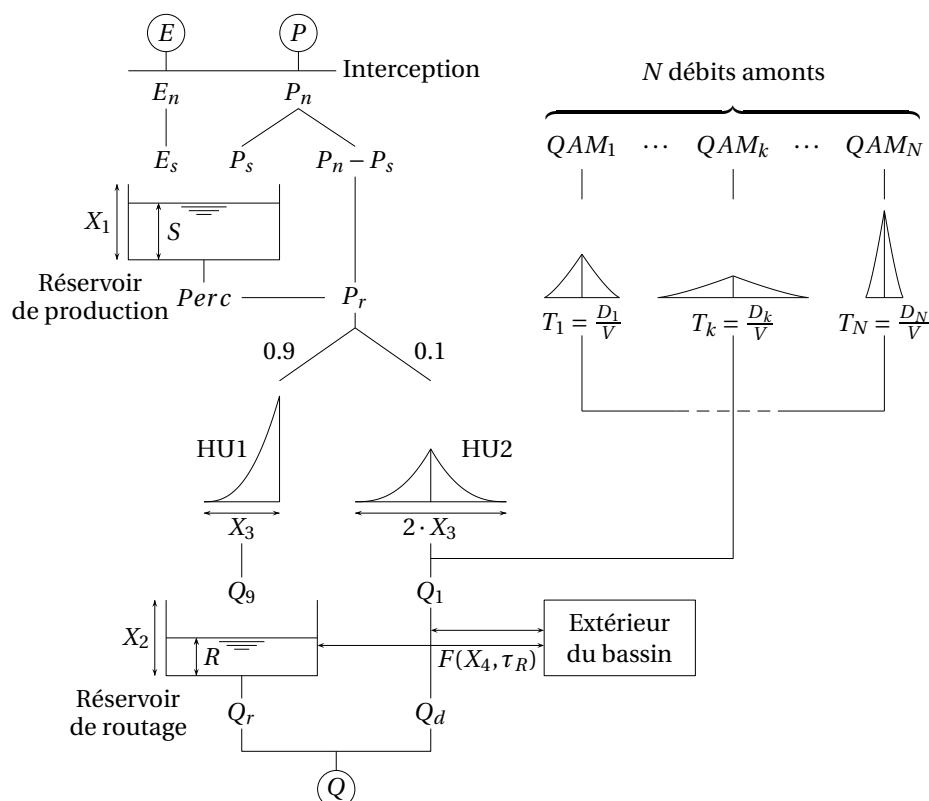


FIG. 6.4: Insertion de débits amont dans le modèle GR4J en semi-distribution utilisant deux paramètres calés supplémentaires : une vitesse V et un exposant p donnant la forme des N hydrogrammes unitaires (les deux mêmes paramètres pour tous les débits amont).

6.6 Résultats

Dans ce chapitre, nous présentons les résultats des tests pour deux structures proches, différant simplement par leur fonction d'échange : l'une des structures repose sur un modèle pluie-débit à quatre paramètres (auxquels on ajoute deux paramètres de routage) et l'autre sur un modèle à cinq paramètres (soit sept avec le routage). Nous donnerons les détails de ces deux structures au Chapitre 8 : ce chapitre se borne à montrer la sensibilité du critère de cohérence à la formulation de la fonction d'échange. Il s'agit plutôt d'une « preuve de concept » que d'une étude exhaustive sur toutes les structures testées dans la partie 3, car les simulations sont relativement lourdes à mettre en oeuvre.

6.6.1 Performance en restitution des débits

Les distributions de performance en restitution des débits (Figure 6.5) montre que l'approche semi-distribuée n'est pas une manière directe d'améliorer les simulations à un exutoire donné, ce qui confirme certains résultats de la littérature (e.g. *Diermanse*, 2001 ; *Booij*, 2002). Lorsque l'on utilise les débits observés en contrôle, la simulation du débit à l'exutoire des bassins intermédiaires est naturellement améliorée par rapport au cas global (le débit amont apporte beaucoup d'information). En revanche, dans le cas où l'on utilise les

simulations amont en contrôle (i.e., contrôle « en cascade »), les performances chutent en raison du cumul des erreurs amont, et sont à peine supérieures à celle du modèle global au même point.

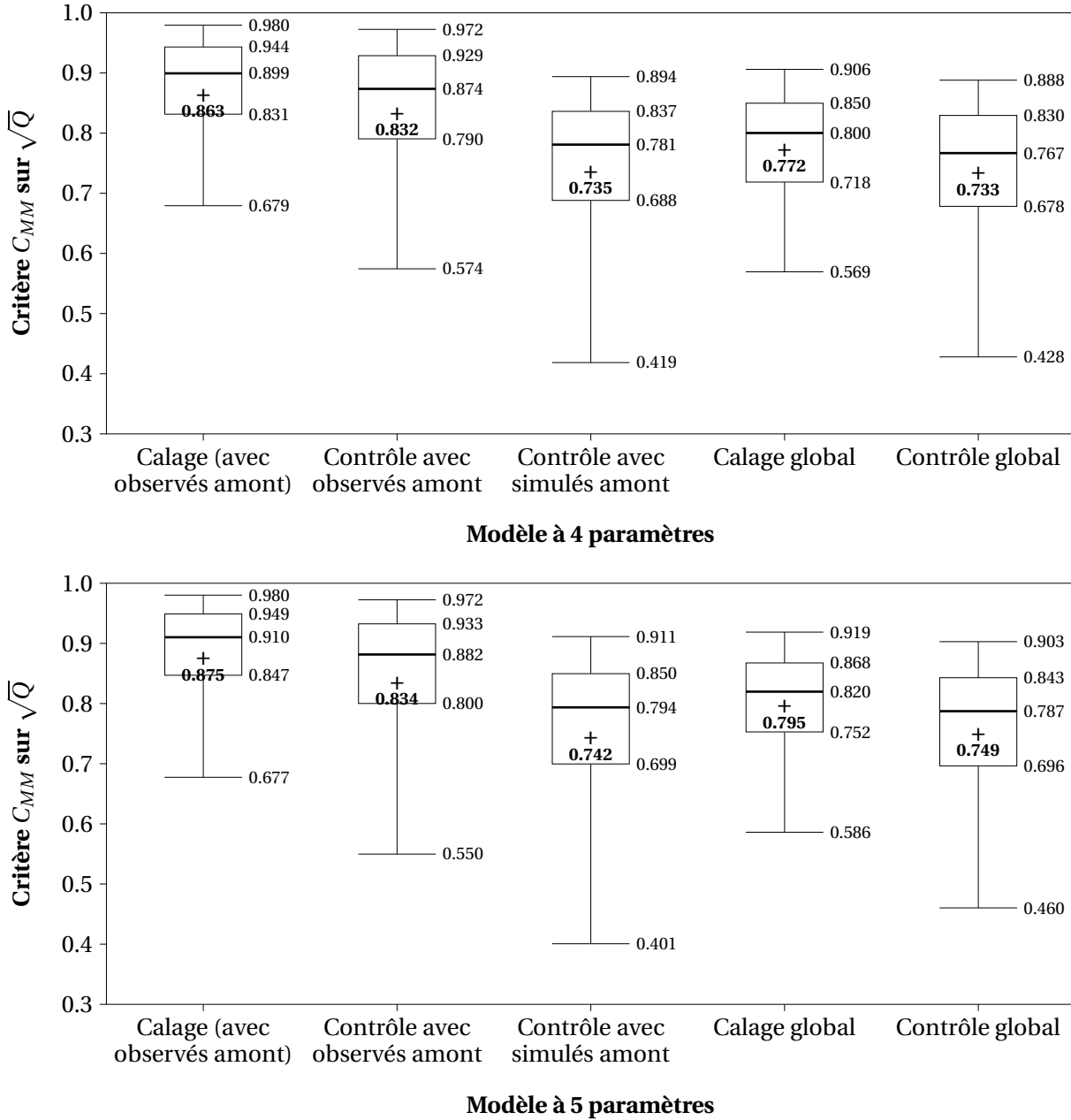


FIG. 6.5: Résultats en calage et en contrôle pour les structures à quatre et cinq paramètres. Les performances en contrôle « en cascade » (utilisant les simulations amont successives, en contrôle elles aussi) sont équivalentes à celles en contrôle global sur l'échantillon de 284 stations, voire moins bonnes pour la moyenne (croix) : les mauvais résultats d'une station amont se répercutent à l'aval lorsque l'on passe des observés aux simulés.

6.6.2 Distribution des paramètres sur les bassins intermédiaires

On a tenu à vérifier (Figure 6.6) que les distributions des paramètres sur les bassins intermédiaires ne sont pas trop différentes de celles obtenues en calage global, pour voir si le modèle ne se « tord » pas trop en semi-distribution (un bassin intermédiaire et un bassin versant classique n'étant pas le même objet).

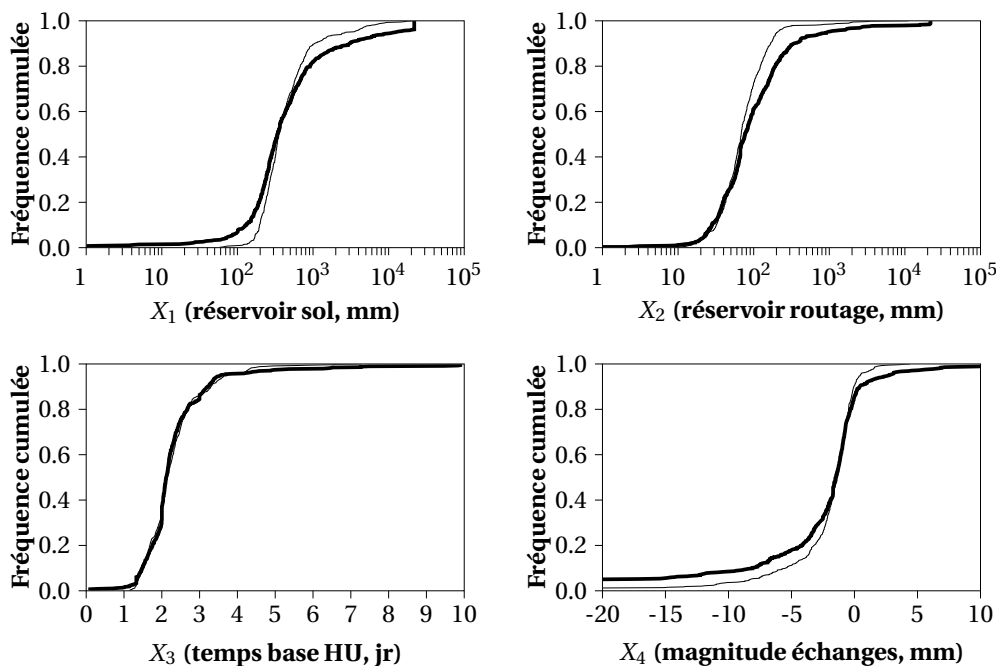


FIG. 6.6: Distributions des quatre paramètres de GR4J sur les 284 bassins intermédiaires (lignes épaisses) et sur les 284 stations en global (lignes fines).

On constate que les médianes sont assez proches mais que les paramètres X_1 , X_2 et X_4 , la variance est plus grande en semi-distribution qu'en global. Cela paraît logique puisqu'en isolant les bassins intermédiaires (i.e. en choisissant des entités plus petites), on accroît la variété des situations. D'autre part, la variance plus grande du paramètre d'échange en semi-distribution semble confirmer l'idée que ce paramètre ne corrige pas une entrée telle que la pluie (cf. Chapitre 4), faute de quoi il devrait diminuer quand on élimine les têtes de bassins (zones d'altitudes les plus élevées).

6.6.3 Critère de cohérence spatiale du bilan

La vérification de la cohérence spatiale du bilan est le véritable objectif de ce test de semi-distribution. Comparons le bilan calculé globalement sur un bassin versant de surface A pendant un intervalle Δt , avec la somme des bilans sur N bassins intermédiaires et têtes de bassin, de surfaces $(a_1, \dots, a_k, \dots, a_N)$:

$$\begin{aligned}
S_k(t) + P_k(\Delta t) + G_k(\Delta t) &= Q_k(\Delta t) + E_k(\Delta t) + S_k(t + \Delta t) \\
\sum_{k=1}^N a_k S_k(t) + \sum_{k=1}^N a_k P_k(\Delta t) + \sum_{k=1}^N a_k G_k(\Delta t) &= \sum_{k=1}^N a_k Q_k(\Delta t) + \sum_{k=1}^N a_k E_k(\Delta t) + \sum_{k=1}^N a_k S_k(t + \Delta t) \\
A \cdot S(t) + A \cdot P(\Delta t) + A \cdot G(\Delta t) &= A \cdot Q(\Delta t) + A \cdot E(\Delta t) + A \cdot S(t + \Delta t)
\end{aligned}$$

(tous les termes du bilan : P = pluie, Q = débit écoulé, S = stock, E = évapotranspiration et G = échange étant exprimés en hauteur d'eau)

On doit donc vérifier que les $G_k(\Delta t)$ calculés sur tous les sous-ensembles k d'un bassin versant sont cohérents avec la relation

$$\sum_{k=1}^N a_k \cdot G_k(\Delta t) = A \cdot G(\Delta t)$$

De façon similaire, on doit vérifier pour l'évapotranspiration réelle (ETR) :

$$\sum_{k=1}^N a_k \cdot E_k(\Delta t) = A \cdot E(\Delta t)$$

Sur la Figure 6.7, on constate une assez bonne concordance des échanges en valeur interannuelle ($\Delta t = 10$ ans).

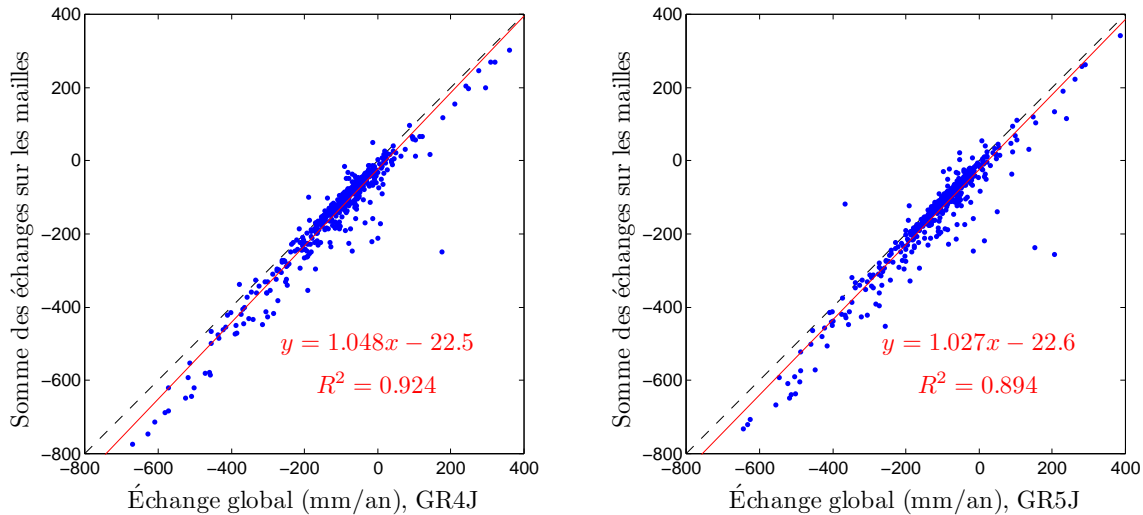


FIG. 6.7: Comparaison des échanges interannuels sur les 284 stations aval, calculés globalement (abscisses) ou en semi-distribution (ordonnées). À gauche : structure à 4 paramètres ; À droite : structure à 5 paramètres.

La Figure 6.8 présente les distributions d'un critère d'adéquation entre les échanges mensuels agrégés sur toutes les mailles d'un bassin, et les échanges mensuels calculés par le modèle global ($\Delta t = 1$ mois, on utilise les volumes mensuels de façon à négliger les effets

de routage à travers le bassin global). Ce critère est une légère variante du critère C_{MM} (et donc, du critère de Nash) : plutôt que de normer l'erreur quadratique moyenne par la variance d'une des deux séries prise comme référence (échanges globaux ou distribués), nous normons par le produit des écarts-type. En effet, il n'y a pas de raison de considérer l'une ou l'autre de ces séries comme étant plus « vraie » que l'autre. Le critère est donc :

$$C_{\text{éch}} = \frac{1 - \frac{F}{F_0}}{1 + \frac{F}{F_0}}$$

$$\frac{F}{F_0} = \frac{\sum_t (G_{\text{glo},t} - G_{\text{dis},t})^2}{N \cdot \sigma_{G_{\text{glo}}} \sigma_{G_{\text{dis}}}}$$

Comme le critère C_{MM} , ce critère est compris entre -1 et 1 , la valeur 1 signifiant que les deux séries coïncident exactement. De la même façon, nous pouvons calculer un critère d'adéquation pour les flux atmosphériques (« évapotranspiration réelle ») simulés par le modèle en global et en semi-distribution :

$$C_{\text{ETR}} = \frac{1 - \frac{F}{F_0}}{1 + \frac{F}{F_0}}$$

$$\frac{F}{F_0} = \frac{\sum_t (E_{\text{glo},t} - E_{\text{dis},t})^2}{N \cdot \sigma_{E_{\text{glo}}} \sigma_{E_{\text{dis}}}}$$

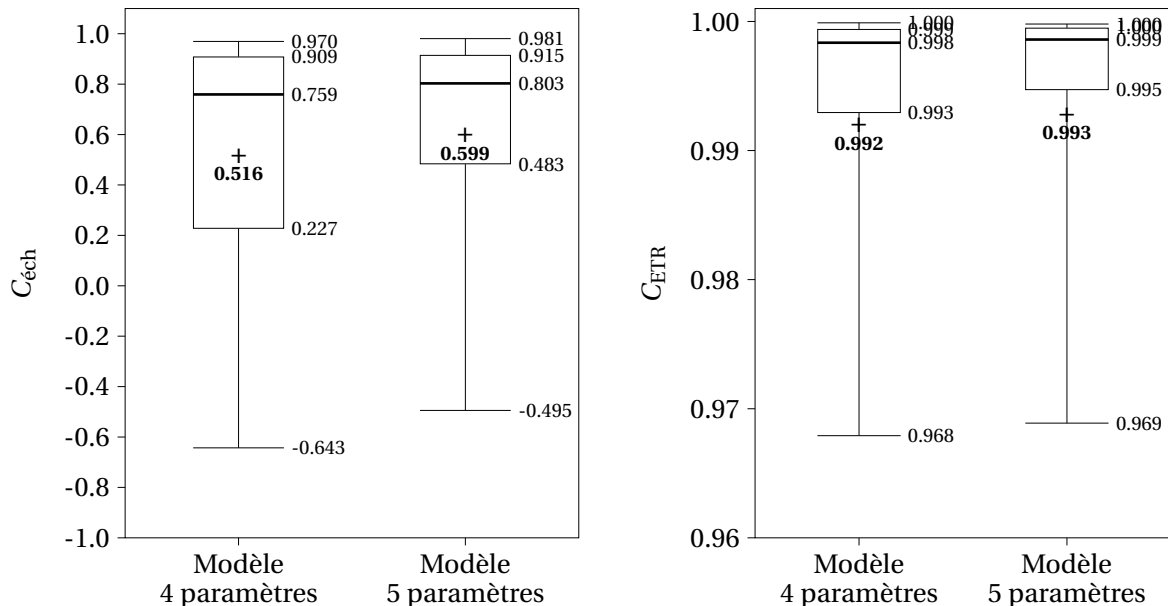


FIG. 6.8: Distribution des critères d'adéquation entre échanges globaux et échanges distribués (gauche), et entre ETR globale et ETR distribuée (droite).

Les distributions de ces critères mettent en évidence un assez bon accord entre les dynamiques des échanges simulés en global et en semi-distribué, ainsi qu'une légère améliora-

tion lorsque l'on passe de la structure à quatre paramètres à la structure à cinq paramètres (avec échanges linéaires et seuil). Cela ne veut pas dire que les termes calculés sont plus exacts, mais que cette structure semble un peu plus stable quand on change d'échelle. On constate malgré tout un certain nombre de fortes divergences entre global et semi-distribué ($C_{éch}$ faibles voire négatifs), qu'il faut cependant nuancer par le fait qu'une erreur sur de faibles échanges entraîne quand même un mauvais critère (les implications de ce point seront discutées en conclusions).

D'autre part, les critères d'adéquation de l'ETR sont toujours très proches de 1. Outre la faible sensibilité des modèles GR à la valeur de l'ETP, déjà observée par Oudin (2004), ce résultat tient peut-être aussi au fait que l'ETP est assez peu variable spatialement entre les différents bassins amont, et donc que les paramètres du réservoir de production sont eux-aussi peu différents. La Figure 6.9 illustre la façon de le critère $C_{éch}$ complète l'information donnée par le critère de performance sur les débits. Pour chaque modèle, le graphique de droite montre la répartition des résultats selon les axes « performance » et « cohérence » : la performance est représentée par le critère $C_{MM}(\sqrt{Q})$ en contrôle (axe des abscisses) tandis que la cohérence du bilan est mesurée par le critère $C_{éch}$. Le nombre de points situés dans le coin haut gauche est quasi nul (i.e., on ne risque pas d'avoir la cohérence du bilan si on n'a pas d'abord une bonne simulation des débits aux exutoires), en revanche un certain nombre de points sont situés dans le coin bas-droite, c'est-à-dire que le modèle a été capable de simuler assez bien les débits mais en ayant recours à un « tour de passe-passe » au niveau du bilan. La fréquence de ces incohérences doit toutefois être relativisée par chacun de graphiques de gauche, sur lesquels on a tracé la valeur du critère $C_{éch}$ en fonction de l'importance des échanges (moyennée entre mode global et semi-distribué) : on constate que bon nombre de points où le critère $C_{éch}$ est très mauvais correspondent à des échanges faibles. Dans ces cas, le critère continue à pénaliser les biais relatifs alors que les biais absolus sont très faibles (quelques dizaines de millimètres annuellement).

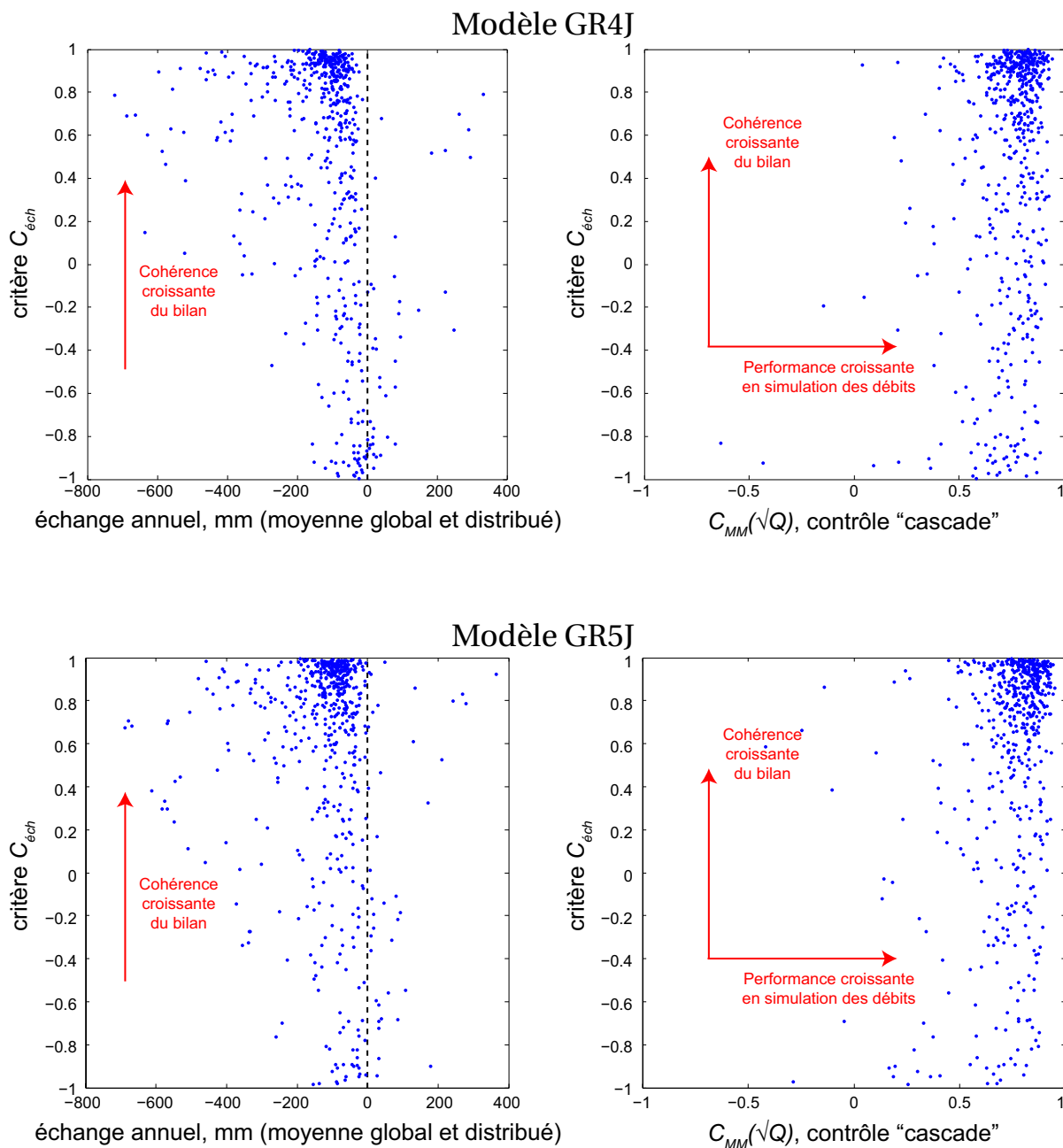


FIG. 6.9: Dépendance et complémentarité entre le critère de performance en simulation des débits, et le critère de cohérence du bilan

6.7 Un cas d'étude en semi-distribution

Dans ce paragraphe nous proposons une illustration de la méthode utilisée. La carte de la Figure 6.10 et le Tableau 6.1 présentent un groupe de six stations hydrométriques dans le bassin de la Seine amont : Ce groupe de bassins versants est intéressant car il s'agit

Rivière	Code	Surface (km ²)
Laigne	H0203030	674
Ource 1	H0301010	173
Ource 2	H0321030	548
Seine 1	H0100010	371
Seine 2	H0100020	704
Seine 3	H0400010	2340

TAB. 6.1: Sous-bassins jaugés du bassin de la Seine à Bar-sur-Seine (H0400010, 2340 km²).

d'une zone karstique où les échanges entre bassins versants topographiques sont fréquents, particulièrement dans la vallée de la Laigne (cf. flèches noires sur la carte). Dans cette configuration, nous avons trois possibilités de confronter un bilan distribué avec un bilan global : sur la station Ource 2, la station Seine 2, et la station Seine 3. Les résultats sont présentés dans le Tableau 6.2 et sur la Figure 6.11, pour le modèle GR5J.

Station	S_{glo} (km ²)	S_{int} (km ²)	$C_{MM}(\sqrt{Q})$ glo.	$C_{MM}(\sqrt{Q})$ dis.	G_{glo} (mm/an)	G_{int} (mm/an)	G_{dis} (mm/an)	V_{glo} (m ³ /s)	V_{int} (m ³ /s)	V_{dist} (m ³ /s)
Laigne	674	<i>n.a.</i>	0.755	<i>n.a.</i>	-217	<i>n.a.</i>	<i>n.a.</i>	-4.63	<i>n.a.</i>	<i>n.a.</i>
Ource 1	173	<i>n.a.</i>	0.822	<i>n.a.</i>	-72	<i>n.a.</i>	<i>n.a.</i>	-0.39	<i>n.a.</i>	<i>n.a.</i>
Ource 2	548	375	0.832	0.839	-74	-93	-87	-1.28	-1.11	-1.50
Seine 1	371	<i>n.a.</i>	0.758	<i>n.a.</i>	-50	<i>n.a.</i>	<i>n.a.</i>	-0.58	<i>n.a.</i>	<i>n.a.</i>
Seine 2	704	333	0.796	0.796	50	109	26	+1.11	+1.15	+0.57
Seine 3	2340	414	0.822	0.825	-57	+11	-73	-4.21	+0.14	-5.42

TAB. 6.2: Termes d'échange simulés globalement et en semi-distribution par le modèle à cinq paramètres. S_{glo} : surface du bassin global associé à la station; S_{int} : surface du bassin intermédiaire associé à la station; G_{glo} : échange global; G_{int} : échange sur le bassin intermédiaire; G_{dis} : résultante des échanges sur les mailles constituant le bassin global; Les termes V_{glo} , V_{int} , V_{dist} correspondent aux termes G_{glo} , G_{int} , G_{dist} mais exprimés cette fois en m³/s. Les couples de valeurs tests sont indiqués en gras.

On constate tout d'abord sur le tableau 6.2 que le modèle parvient à simuler correctement les débits aux exutoires, que ce soit en mode global (quatrième colonne, $C_{MM}(\sqrt{Q})$ glo., pour tous les bassins) ou en mode semi-distribué (cinquième colonne, $C_{MM}(\sqrt{Q})$ dis., pour les trois stations où cette solution est applicable). L'accord entre les termes d'échange en mode global et en mode semi-distribué est relativement satisfaisant, à l'exception de la station Seine 2 : l'écart est important en valeurs relatives (facteur deux), écart qui doit toutefois être relativisé par la faiblesse des termes en valeur absolue (cf. conclusion).

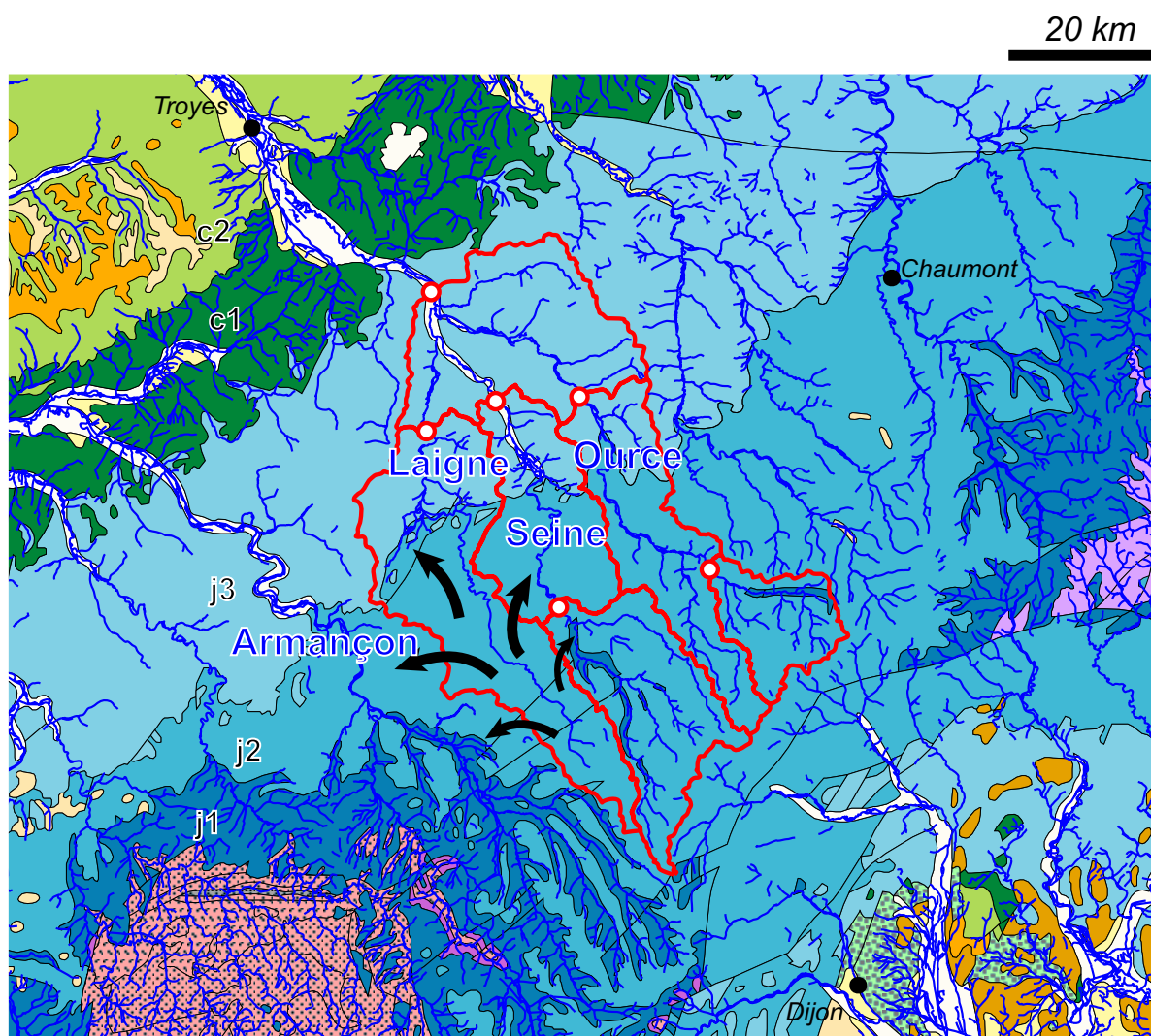


FIG. 6.10: Bassin et sous-bassins versants de la Seine à Bar-sur-Seine (2340 km²). Les deux affluents principaux sont l'Ource à l'est, et la Laigne à l'ouest. L'existence d'écoulements souterrains divergents (flèches noires) est avérée dans le bassin de la Laigne, ces flux alimentant les bassins de la Seine et de l'Armançon (une partie réapparaît dans le bassin même de la Laigne, sous forme de sources karstiques à la limite entre les calcaires du Jurassique moyen (j2) et du Jurassique supérieur (j3)). Voir par exemple *Cornet et Lelong* (2006) ou *Amiot* (1989) pour une description des circulations souterraines.

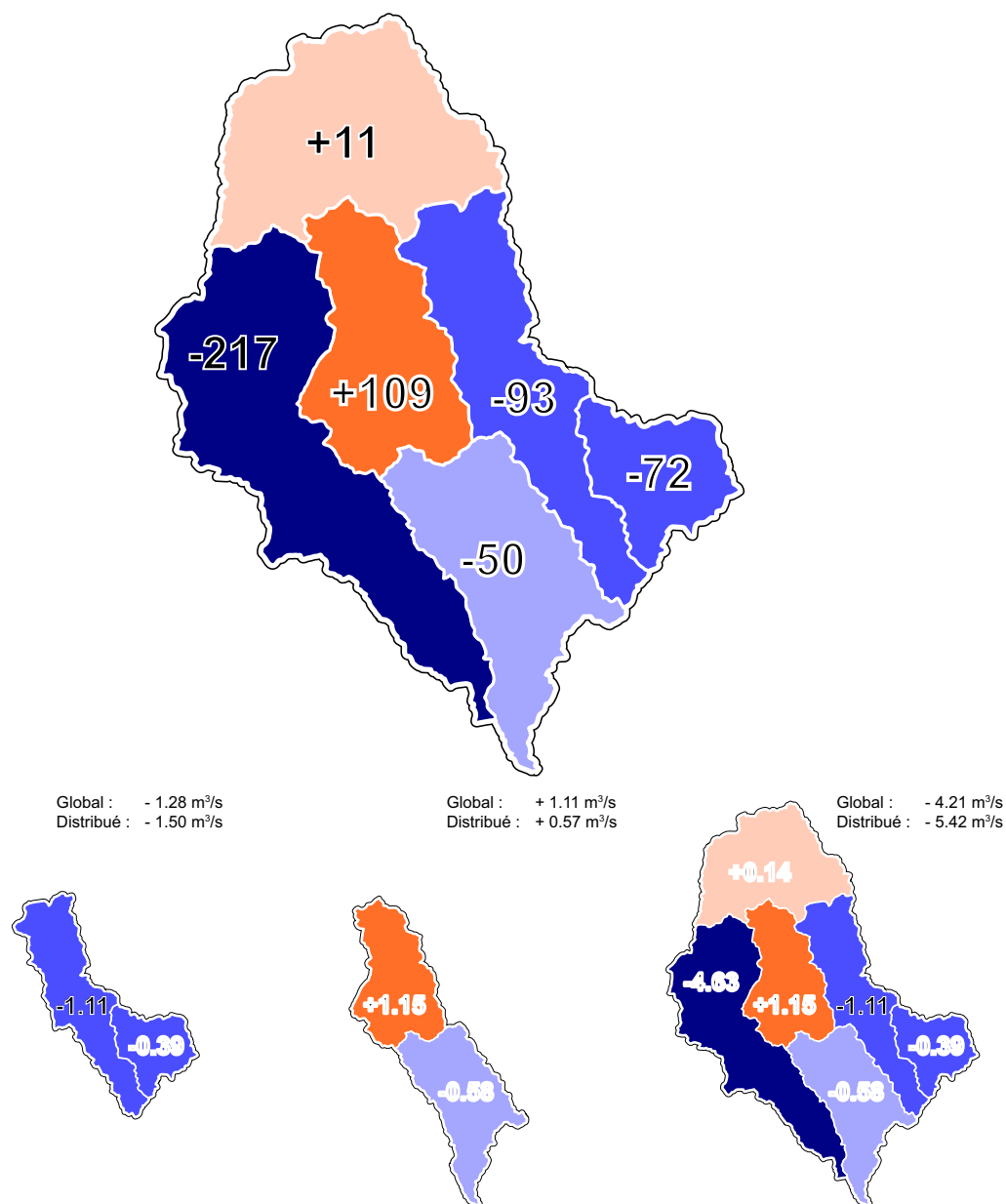


FIG. 6.11: Comparaison des bilans globaux et distribués sur le bassin de la Seine à Bar-sur-Seine. **Haut** : Lames d'eau échangées (mm/an) telles que simulées par le modèle à cinq paramètres sur les bassins / sous-bassins élémentaires. **Bas** : Comparaison aux stations Ource 2 (gauche), Seine 2 (centre), et Seine 3 (droite). Les flux échangés, exprimés cette fois en m³/s pour pouvoir faire les sommer, sont toujours des moyennes interannuelles.

Sur la Figure 6.12, nous avons tracé les chroniques d'échanges simulés pour les trois stations-tests (Ource 2, Seine 2 et Seine 3). Les valeurs sont agrégées mensuellement de façon à tenir compte des temps de transfert entre les compartiments.

6.8 Synthèse

Dans ce chapitre, nous avons tenté de construire un test permettant d'évaluer la cohérence interne du modèle, lorsque l'on passe du mode global à un mode semi-distribué sur un ensemble de sous-bassins versants définissant une partition du bassin initial. Ici, c'est la capacité du modèle à fournir des estimations de flux interbassins (pertes / gains) qui nous sert à juger la cohérence. Ces tests, assez lourds à mettre en oeuvre, ont porté sur deux structures différant simplement par la formulation de la fonction d'échange. Ils nous ont permis d'établir que :

- en moyenne interannuelle, on observe une bonne cohérence entre les échanges calculés globalement et en semi-distribution,
- en valeurs mensuelles, l'accord reste assez bon,
- enfin, il nous a semblé particulièrement satisfaisant que le critère de cohérence spatiale que nous avons défini s'avère sensible à la formulation des échanges. L'objectif de ce chapitre était de montrer que la construction d'un tel critère était possible : en effet, plus que la qualité du modèle à un moment donné, ce qui compte vraiment est de disposer d'outils pour le faire progresser !

Cependant, il ne faut voir ce critère que comme un complément au critère d'efficacité classique sur les débits, pour discriminer plus avant les modèles. En effet, si le modèle ne marche pas, il est inutile de recourir à de tels tests ! Il y a d'ailleurs fort à parier que le critère de cohérence ne veut plus rien dire en cas d'échec du modèle à simuler les débits.

En outre, ce test permet d'avoir une estimation de la précision à laquelle on peut éventuellement détecter un signal d'influence souterraine dans la relation pluie-débit. En effet, nous avons écrit plus haut qu'il ne fallait pas attacher trop de signification aux échanges de faible amplitude (et donc aux écarts associés en semi-distribution). Nous pouvons donc, inversement, supposer que l'écart-type des erreurs entre les bilans calculés en global et semi-distribution sont un bon ordre de grandeur de cette incertitude. Cet écart-type valant environ 40 mm, nous pourrions nous risquer à l'affirmation suivante : **lorsque le modèle parvient à simuler correctement les débits, et que le terme d'échange simulé dépasse quelques dizaines de millimètres annuellement (en valeur absolue), nous pouvons raisonnablement suspecter des écoulements souterrains interbassin.**

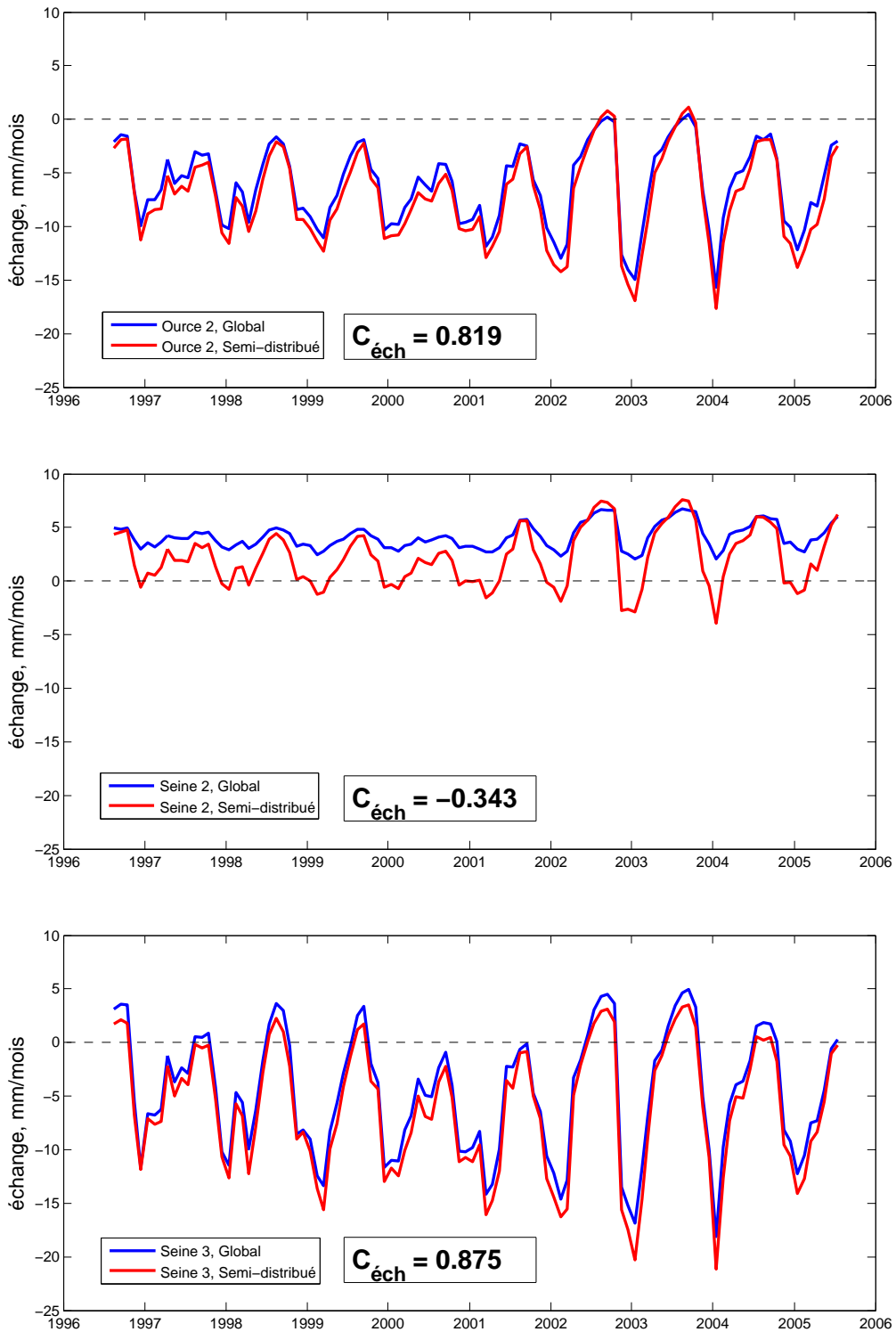


FIG. 6.12: Chroniques d'échanges mensuels calculés globalement (bleu) et en semi-distribution (rouge) pour les trois stations-tests : Ource 2 (H0321030), Seine 2 (H0100020) et Seine 3 (H0400010). Nous constatons un très bon accord pour les stations Ource 2 et Seine 3, mais des écarts importants en crue (hiver) pour la station Seine 2 (toutefois, dans ce dernier cas, il s'agit d'écarts absolus de l'ordre de quelques millimètres par mois). Dans les trois cas on constate une légère différence d'amplitude entre les deux moitiés de chaque chronique (1996–2000 et 2001–2005) : ce défaut est normal puisque chaque chronique est composée de deux hydrogrammes en contrôle mis bout-à-bout, et donc obtenus avec deux jeux de paramètres pas tout à fait identiques (*split-sample test* symétrique, cf. page 16).

7

Conclusion à la deuxième partie

7.1 Les limites de l'approche vérificationniste

Dès l'introduction, nous avons mentionné la principale difficulté du sujet auquel nous nous intéressons : les flux échangés entre les hydrosystèmes de surface et les hydrosystèmes souterrains rentrent généralement dans la catégorie de ce que *De Marsily* (1994) appelle « *phénomènes inobservables* ». Dans ces conditions, les possibilités de tester les flux prédits par les modèles se limitent à deux possibilités :

1. des tests de *pouvoir prédictif* (i.e., des confrontations à des observations), mais limités à des cas particuliers, comme celui du karst de la Rochefoucauld, où les flux sont mesurables,
2. des tests qui, tout en étant généralisables (semi-distribution), se limitent à la *profondeur explicative* (i.e., à l'évaluation de la cohérence interne) du modèle.

Cette situation est très insatisfaisante, et il semble finalement assez hasardeux de se prononcer sur le caractère réaliste des simulations de ces flux souterrains par les modèles GR. En l'absence d'observations du phénomène, ces simulations ne semblent tout simplement pas testables.

7.2 Peut-on utiliser la donnée piézométrique pour contraindre les simulations ? Un premier calcul de coin de table

Il existe des données, dont la collecte est largement aussi répandue que celle des données de pluie et de débit, et que les hydrogéologues pourraient nous reprocher d'avoir jusqu'à présent écartées : les observations piézométriques. En effet, les épisodes de recharge ou de vidange d'une nappe se traduisent en principe par des variations de niveau observables dans des puits. À partir de ces niveaux, pouvons-nous déduire les stockages ou déstockages correspondants ? À première vue, la réponse semble évidemment positive, puisque c'est une pratique courante en modélisation hydrogéologique. Pourtant, il est nécessaire d'examiner les hypothèses couramment faites dans ce type de modélisation.

Considérons à nouveau le cas unidimensionnel de la surface piézométrique sous un interfluve de largeur L , présenté au Chapitre 3, et faisons l'hypothèse que la « vraie » situation est la situation 2 de la Figure 3.5 (p. 57) : une perméabilité uniforme K et une recharge spatialement variable (situation pas très éloignée de celle présentée dans *Jackson et Rushton*, 1987), recharge que nous prendrons schématiquement de forme $q(x) = (q_B - q_A) \frac{x}{L} + q_A$. Sur

la Figure 7.1, schéma du haut, nous avons représenté une telle situation pour

$$\left\{ \begin{array}{l} L = 10000 \text{ m} \\ K = 10^{-4} \text{ m} \cdot \text{s}^{-1} \\ h_A = h_B = 5 \text{ m} \\ q_A = 0 \text{ mm} \cdot \text{an}^{-1} \\ q_B = 300 \text{ mm} \cdot \text{an}^{-1} \end{array} \right.$$

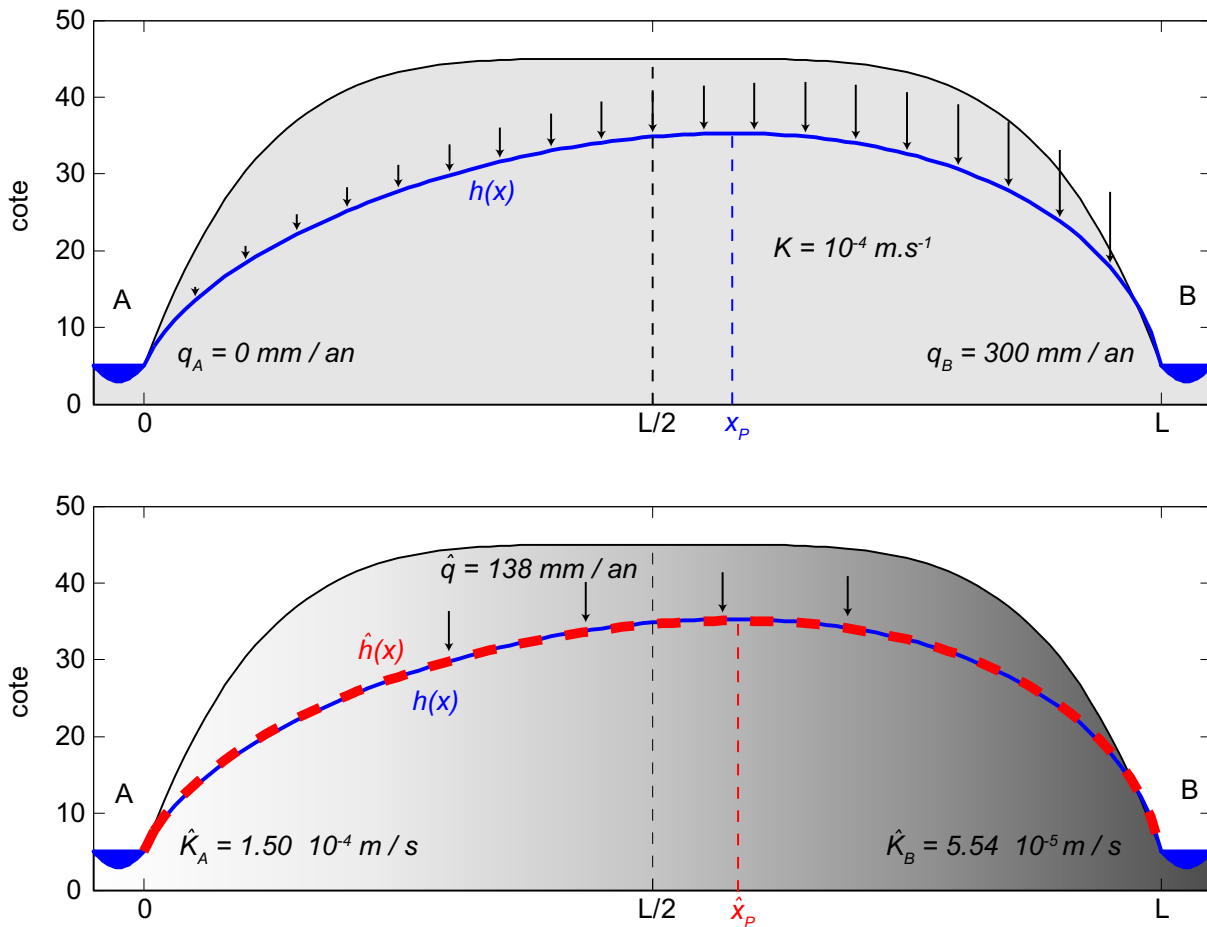


FIG. 7.1: Exemple d'équifinalité dans une modélisation hydrogéologique inverse. En haut, la situation « réelle » (prise comme telle) avec une perméabilité uniforme et une recharge croissante entre A et B ; En bas, l'une des solutions possible au problème inverse en l'absence d'informations précises sur la structure spatiale de la recharge et de la perméabilité : cette solution à recharge uniforme et une perméabilité décroissante entre A et B (gradient constant sur le logarithme) permet de reproduire correctement la piézométrie, mais pas les flux totaux en A et B.

Nous obtenons l'abscisse x_p de la ligne de partage des eaux en annulant la dérivée $\frac{dh^2}{dx}$ (polynôme d'ordre 2 en x) : le résultat est $x_p = 5774 \text{ m}$, les débits de base $Q_{b,A}$ et $Q_{b,B}$ valent

donc respectivement

$$\begin{cases} Q_{b,A} = \int_{x=0}^{x=x_p} q(x) dx = x_p \cdot \frac{(q_A + q(x_p))}{2} = 1.58 \text{ m}^2 \cdot \text{s}^{-1} \\ Q_{b,B} = \int_{x=x_p}^{x=L} q(x) dx = (L - x_p) \cdot \frac{(q(x_p) + q_B)}{2} = 3.17 \text{ m}^2 \cdot \text{s}^{-1} \end{cases}$$

Nous remarquons que la ligne de partage des eaux est déplacée vers la droite, mais comme la recharge est plus forte à droite qu'à gauche, le débit de base en B reste plus fort que celui en A .

Revenons à notre problème inverse, et supposons que nous ne connaissions précisément ni K , ni la recharge $q(x)$. Nous pourrions alors (1) faire l'hypothèse, malencontreuse mais pourtant raisonnable faute d'informations, que la recharge est à peu près uniforme sur l'interfluve, de valeur \hat{q} ; et (2) penser que si une grandeur varie spatialement, c'est la perméabilité K . Supposons donc que nous cherchions le champ de perméabilité sous la même forme qu'à la situation 3 de la Figure 3.5, avec un gradient constant sur le logarithme :

$$\ln \hat{K}(x) = (\ln \hat{K}_B - \ln \hat{K}_A) \frac{x}{L} + \ln \hat{K}_A$$

La question est alors : quelle erreur commet-on dans l'estimation des débits de base lorsque l'on essaye de reproduire la piézométrie de la Figure 7.1 (haut), résultant d'une recharge non-uniforme, avec le modèle ainsi spécifié, i.e. en calant les paramètres « physiques » (\hat{q} , \hat{K}_A , \hat{K}_B) (ou (\hat{q} , \hat{K}_A , $\hat{\beta}$), avec $\beta = \frac{1}{L} \ln \left(\frac{\hat{K}_B}{\hat{K}_A} \right)$) ? Le schéma du bas sur la Figure 7.1 montre la piézométrie simulée $\hat{h}(x)$: l'accord est très bon (Figure 7.1, bas), et les valeurs obtenues ne sont pas « irréalistes » :

$$\begin{aligned} \hat{K}_A &= 1.50 \times 10^{-4} \text{ m} \cdot \text{s}^{-1} \\ \hat{K}_B &= 5.54 \times 10^{-5} \text{ m} \cdot \text{s}^{-1} \\ \hat{q} &= 138 \text{ mm} \cdot \text{an}^{-1} \end{aligned}$$

En particulier, la valeur \hat{q} estimée est proche de la moyenne « réelle » sur le domaine (138 contre 150 mm/an). Comme nous reproduisons correctement la piézométrie, nous estimons en particulier correctement la position de la ligne de partage des eaux (ici $\hat{x}_p = \frac{1}{\hat{\beta}} \left[1 - \frac{\beta^2 \hat{K}_A C}{2 \hat{q}} \right] = 5816 \text{ m}$). Mais comme nous avons fait l'hypothèse d'une recharge uniforme, les débits de base que nous estimons sont :

$$\begin{cases} \hat{Q}_{b,A} = \int_{x=0}^{x=\hat{x}_p} \hat{q}(x) dx = \hat{x}_p \hat{q} = 2.54 \text{ m}^2 \cdot \text{s}^{-1} \\ \hat{Q}_{b,B} = \int_{x=\hat{x}_p}^{x=L} \hat{q}(x) dx = (L - \hat{x}_p) \hat{q} = 1.83 \text{ m}^2 \cdot \text{s}^{-1} \end{cases}$$

Cela correspond respectivement à une surestimation de 61% en A et une sous-estimation de 42% en B ! Bien sûr, il ne faut pas accorder trop d'importance à ces valeurs, mais plutôt au résultat qualitatif : si, pour reproduire la piézométrie, nous avons besoin de faire des hypothèses sur la recharge, alors le réalisme des flux souterrains calculés n'est pas assuré, même si les équations utilisées ont un fondement physique. Par cette étude théorique, nous avons simplement voulu monter concrètement l'équifinalité de ce type de modélisation inverse lorsque le problème est sous-contraint (ce qui est *toujours* le cas). Le fait de prendre comme « référence » la situation de recharge variable et de perméabilité uniforme ne signifie pas que nous considérons la recharge comme plus variable spatialement que la perméabilité (ni le contraire, d'ailleurs). Simplement, force est de constater que les hydrogéologues sont plus enclins à tourner le bouton « perméabilités » dans leurs modèles et à considérer la condition à la limite supérieure (le flux de recharge) comme plus ou moins acquise, ou moins sensible.

Bien sûr, en pratique, on ne cherche pas une solution analytique mais on utilise un maillage, ce qui a tendance à multiplier le nombre de degrés de liberté et l'équifinalité associée. Toutes ces conclusions sont généralisables en deux dimensions et en régime transitoire, l'équation de continuité en nappe libre s'écrivant :

$$\vec{\nabla} \cdot \left(-\frac{1}{2} K \vec{\nabla} h^2 \right) = q + \omega_d \frac{\partial h}{\partial t}$$

avec $q = q(x, y, t)$, $h = h(x, y, t)$, $K = K(x, y)$ et $\omega_d = \omega_d(x, y)$ la porosité de drainage. Dans ce cas, l'ajout d'un champ supplémentaire à identifier (ω_d) complique encore un peu plus le problème inverse. La situation que nous avons choisie était bien évidemment un peu caricaturale, et les intervalles possibles pour q et K ne sont heureusement pas infinis en pratique (on disposera d'estimations locales du champ de perméabilités, par des essais de pompages par exemple). Par contre, la surface piézométrique ne sera jamais complètement connue, l'interpolation des charges ajoutant beaucoup d'incertitudes par rapport à notre cas schématique. Ainsi, l'absence de solution unique aux problèmes inverses en hydrogéologie reste bien une réalité (*de Marsily et al.*, 1999 ; *de Marsily et al.*, 2005 ; *Carrera*, 2005), ce qui les rend difficiles à valider (*Konikow et Bredehoeft*, 1992). Les hydrogéologues ne peuvent *in fine* pas se passer d'une estimation de la recharge et, inversement, l'observation du seul niveau piézométrique n'est pas d'un grand secours à l'hydrologue pour boucler son bilan en eau.

En fait, l'une des façons d'ajouter des contraintes à ce problème est de forcer le modèle à identifier correctement le(s) débit(s) de base du/des cours d'eau drainant l'aquifère étudié (*Sanford*, 2002). Mais cela implique d'en obtenir une estimation indépendamment de toute hypothèse sur la structure de l'aquifère (ou à défaut, de réaliser une modélisation totalement

couplée, point que nous aborderons plus loin) : c'est précisément le sujet qui nous occupe ! Nous pouvons essayer de procéder de la sorte, parce que le terme de recharge q est lui-même une conséquence de la relation pluie-débit et que, sans être négligeable, il est en général plutôt un « ordre 2 ». Ce point fera l'objet du paragraphe 7.4.

7.3 Quelles sont les alternatives basées sur la mesure ?

La mesure de certaines grandeurs intensives associées aux écoulements d'eau, que ce soit la concentration de certains isotopes (voir par exemple *Kendall et McDonnell*, 1998) ou d'autres grandeurs comme la température (e.g., *Tabbagh et al.*, 1999), est une information complémentaire intéressante pour contraindre le type de problèmes inverses mentionnés. Cependant, l'emploi de ces méthodes de traçage pour estimer la recharge et le débit de base reste intimement lié à la spécification d'hypothèses sur le transport (e.g., l'écoulement est-il plutôt piston ou mélange parfait?) ainsi que sur les termes sources / puits (où et à quel rythme les espèces tracées sont-elles produites ou consommées?). *Sandford* note la persistante de cette difficulté (p. 117) : « *les modélisations inverses utilisant les traceurs chimiques révèlent que de telles observations sont importantes pour estimer les flux du système mais d'autres paramètres, comme la porosité, associés au calcul de vitesses d'écoulement, sont toujours indéterminés avec les seules données de traceurs et de niveaux piézométriques* (Medina et Carrera, 1996; Portniaguine et Solomon, 1998). *La situation idéale pour estimer le maximum de paramètres est d'être capable d'utiliser à la fois la mesure du débit de base et des temps de transfert, car cela permet une estimation indépendante de la porosité* ». En d'autres termes, on retombe toujours sur la nécessité d'estimer le débit de base des cours d'eau drainant l'aquifère : cette estimation nécessite donc *à la fois une mesure du débit total, et une méthode exclusivement hydrologique de séparation des écoulements*.

Comparées à ces méthodes géochimiques, les approches géodésiques offrent une alternative intéressante dans la mesure où elles s'intéressent à la redistribution des masses d'eau à l'échelle globale, indistinctement des compartiments du bassin où ont lieu les transferts entre surface et souterrain. En effet, nous avons vu que la *recharge* et *l'écoulement souterrain ne réapparaissant pas à l'exutoire* sont deux notions distinctes. L'estimation de la première est utile pour les modèles hydrogéologiques, mais elle n'aide pas à contraindre les simulations à l'échelle du bassin versant. L'utilisation de méthodes géodésiques satellitales, telles que GRACE (e.g., *Andersen et al.*, 2005; *Schmidt et Hinderer*, 2006), ou au sol (e.g. *Longuevergne*, 2008) est une piste très intéressante, même si leur précision reste pour l'instant limitée¹.

¹en fait, elle ne l'est pas plus que celle des modèles hydrologiques à contraindre : de l'ordre de la centaine de millimètre annuellement.

7.4 N'est-il pas absurde de chercher à modéliser des influences extérieures au bassin topographique, sans modéliser cet extérieur ? Un second calcul de coin de table

Reprenons maintenant la situation plus simple où la recharge et la perméabilité sont uniformes, la non-superposition des bassins topographiques et hydrogéologiques provenant de la différence entre les conditions h_A et h_B (Figure 7.2) :

$$h(x) = \sqrt{-\frac{q}{K}(x-x_p)^2 + h_A^2 - x_p^2} ,$$

$$\text{où } x_p = \frac{L}{2} \left(1 + \frac{K(h_B^2 - h_A^2)}{qL^2} \right)$$

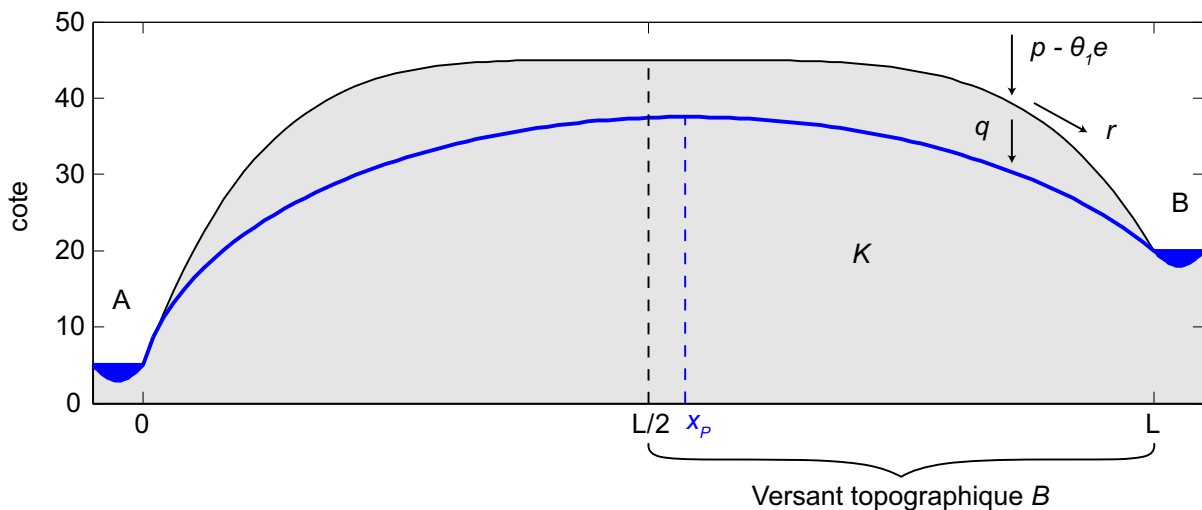


FIG. 7.2: Relation pluie-débit schématique sur le versant B.

Imaginons maintenant que la relation pluie-débit se résume à la fonction de production (pluie nette) suivante : $p_n = p - \theta_1 e$, où p est la pluie, e l'évapotranspiration potentielle et θ_1 un paramètre dépendant du bassin versant ; et que l'infiltration puisse se résumer à un coefficient de partage α de cette pluie nette :

$$\begin{cases} r = \alpha (p - \theta_1 e) & \text{pour le ruissellement,} \\ q = (1 - \alpha) (p - \theta_1 e) & \text{pour la recharge} \end{cases}$$

Le débit spécifique du versant B , de longueur topographique $\frac{L}{2}$, est alors :

$$q_{s,B} = \frac{\alpha(p - \theta_1 e) \frac{L}{2} + Q_{b,B}}{L/2}$$

Comme d'autre part, $Q_{b,B} = q(L - x_p) = q \frac{L}{2} - \frac{K}{L^2} (h_B^2 - h_A^2) \frac{L}{2}$,

$$q_{s,B} = \alpha(p - \theta_1 e) + (1 - \alpha)(p - \theta_1 e) - \frac{K}{L^2} (h_B^2 - h_A^2)$$

$$q_{s,B} = (p - \theta_1 e) - \frac{K}{L^2} (h_B^2 - h_A^2)$$

En notant $\theta_2 = \frac{K}{L^2}$ et $\theta_3 = h_A^2$, nous obtenons l'équation de bilan (spécifique) :

$$q_{s,B} = \underbrace{(p - \theta_1 e)}_{\text{fonction de production}} - \underbrace{\theta_2 (h_B^2 - \theta_3)}_{\text{fonction d'échange}}$$

Nous arrivons alors à la justification de ce développement un peu long. En effet, nous avons volontairement utilisé deux types de notation pour les paramètres : les lettres K , L , h , etc. pour les paramètres décrivant des « propriétés » du système, et les lettres θ_i pour désigner des paramètres « conceptuels »¹. En écrivant une (improbable) équation de bilan pour le bassin B , ne dépendant plus que des **forçages climatiques** (p, e), de **paramètres « conceptuels »** ($\theta_1, \theta_2, \theta_3$) et **d'états du système de surface** (symbolisés par h_B), notre but était de mettre en évidence la logique qui a motivé les développements présentés dans la troisième partie :

- localement, la géométrie de l'aquifère est une conséquence de la relation pluie-débit sur le bassin considéré et ses voisins,
- le taux de retour (rapport recharge / débit de base) est une fonction de la géométrie de l'aquifère,
- si l'on isole un bassin topographique particulier en faisant l'hypothèse, en première approximation, que son environnement immédiat est dans un état « d'équilibre », il n'est donc pas absurde de chercher à représenter le terme d'apport / perte comme une fonction des états d'un modèle simulant la relation pluie-débit sur ce bassin versant topographique, et d'un petit nombre de paramètres conceptuels, sans décrire explicitement le système souterrain. Dans l'exemple précédent, θ_3 traduit l'hypothèse de faible variabilité des conditions avoisinante, et θ_2 est un paramètre « boîte noire » qui résume l'importance des flux souterrains.

Insistons sur le fait que ces calculs de coin de table ne sont, ni de près ni de loin, une tentative de justification « physique » de la forme des fonctions d'échange testées : les essais

¹bien sûr, dans ces calculs, tout est conceptuel étant donné le caractère caricatural de la situation choisie : régime permanent, relation linéaire entre pluie et écoulement, etc., mais les conclusions qualitatives restent les mêmes pour le « vrai » cycle de l'eau.

de paramétrisation dont les résultats sont présentés dans la partie suivante sont purement empiriques. Cette approche, qui exclut toute observation de l'hydrosystème souterrain, peut légitimement être qualifiée de « *modélisation couplée surface-souterrain du pauvre* » : on se référera à nouveau au Chapitre 1 pour en trouver les justifications tant pratiques que philosophiques ! D'ailleurs, la plupart des codes de calcul couplés (e.g., HydroSphere, Therrien et Sudicky, 1996 ; SIM, Rousset et al., 2004) comportent un module de surface et un module souterrain qui opèrent, non de façon simultanée (option numériquement intenable, étant donné la différence des échelles de discrétisation spatiale et temporelle nécessaires pour les écoulements dans la zone saturée et dans la zone non saturée, voir e.g. Jorgensen et al., 1989a, 1989b), mais de façon itérative. Par la fonction d'échange que nous cherchons, nous n'avons pas pour ambition d'aller plus loin que la *première itération* d'un schéma de surface : tenter d'identifier l'ordre 1 de la composante souterraine (dynamique et amplitude), qui est elle-même souvent un ordre 2 dans la relation pluie-débit.

Troisième partie

Tests systématiques de structures à différents pas de temps

8

Développements au pas de temps journalier

8.1 Point de départ

Comme nous l’avons mentionné au Chapitre 3, notre étude ne partait pas de zéro puisque *Nascimento* (1995) et *Perrin* (2000) ont successivement introduit et amélioré la prise en compte des échanges dans le modèle journalier du CEMAGREF. Cette version « point de départ », que nous avons déjà utilisée pour les études des Chapitres 4 et 5, est schématisée en Figure 8.1 (modèle GR4J).

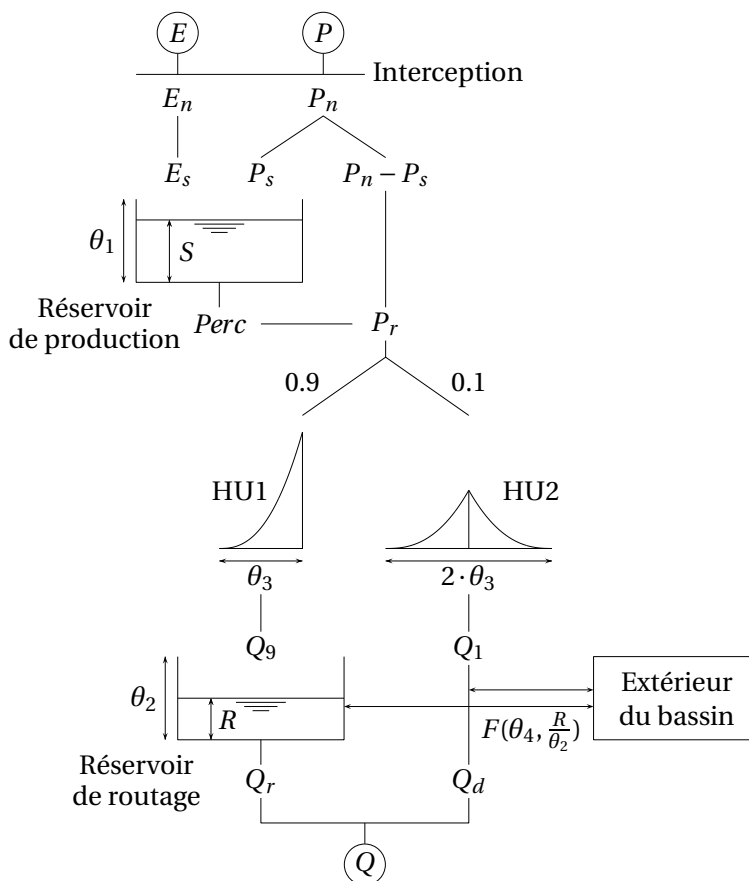


FIG. 8.1: Structure du modèle journalier GR4J

Rappelons que dans cette structure, la fonction d’échange a comme formulation empirique :

$$F_t = \theta_4 \left(\frac{R_t}{\theta_2} \right)^{\frac{7}{2}} \tag{8.1}$$

Dans cette formule, R_t désigne le contenu (en mm) à la date t du réservoir de routage, dont la capacité est paramétrée par θ_2 (le rapport $\frac{R}{\theta_2}$ étant donc le taux de remplissage). θ_4 (mm) est un paramètre fixant le signe et l’amplitude maximale de l’échange journalier. Le terme F_t est ajouté (algébriquement) aux deux branches du routage, dans la limite des stocks disponibles,

de sorte que l'échange à la date t vaut :

$$G_t = \begin{cases} 2F_t & \text{si } F_t > 0 \text{ (i.e., } \theta_4 > 0) \\ -\{\text{Min}(|F_t|; R_t) + \text{Min}(|F_t|; Q_{1,t})\} & \text{si } F_t < 0 \text{ (i.e., } \theta_4 < 0) \end{cases} \quad (8.2)$$

Nous voyons dans cette formulation que l'échange est, pour un bassin donné, toujours du même signe au cours du temps, ce signe étant celui du paramètre θ_4 .

8.2 Canevas des modifications testées

8.2.1 Principe

Nous sommes partis de la structure existante en cherchant des modifications empiriques à cette structure. Comme le rappelle *Mathevet* (2005), l'éventail des modifications possible est très vaste (pour ainsi dire infini) d'autant que chaque modification peut être testée seule ou en combinaison avec d'autres. Les modifications que nous avons testées portent donc principalement :

- sur la redéfinition de la forme fonction $F_t(\cdot)$, qui pourra dépendre de différents états du modèle (taux de remplissage d'un réservoir, etc.) et/ou de variables exogènes (fonction périodique du jour de l'année, par exemple), et d'un plus ou moins grand nombre de paramètres ;
- sur des changements de point d'insertion dans le modèle, i.e. des stocks / flux auxquels on va ajouter ou retrancher le terme F_t . Dans la suite, lorsque nous parlerons d'ajout, il s'agira d'ajout algébrique (ajout de $|F_t|$ si $F_t > 0$, retrait de $|F_t|$ si $F_t < 0$).

Le canevas général des modifications est illustré sur la Figure 8.2. La pertinence des modifications testées est évaluée selon la procédure décrite au Chapitre 1 :

- pour chacun des N bassins versants de l'échantillon, le jeu de données est découpé en n segments ; le modèle est calé sur un segment puis contrôlé sur les $n - 1$ autres segments, le segment utilisé pour le calage étant permuté circulairement (*split-sample test* symétrique). La fonction objectif utilisée pour identifier les paramètres est **l'erreur quadratique moyenne sur \sqrt{Q}** , que l'on transforme en critère :

$$C_{MM}(\sqrt{Q}) = \frac{1 - \frac{F}{F_0}}{1 + \frac{F}{F_0}}$$

$$\text{où } \frac{F}{F_0} = \frac{\sum_t \left(\sqrt{\hat{Q}_t} - \sqrt{Q_t} \right)^2}{\sum_t \left(\mathbb{E}[\sqrt{Q_t}] - \sqrt{Q_t} \right)^2}$$

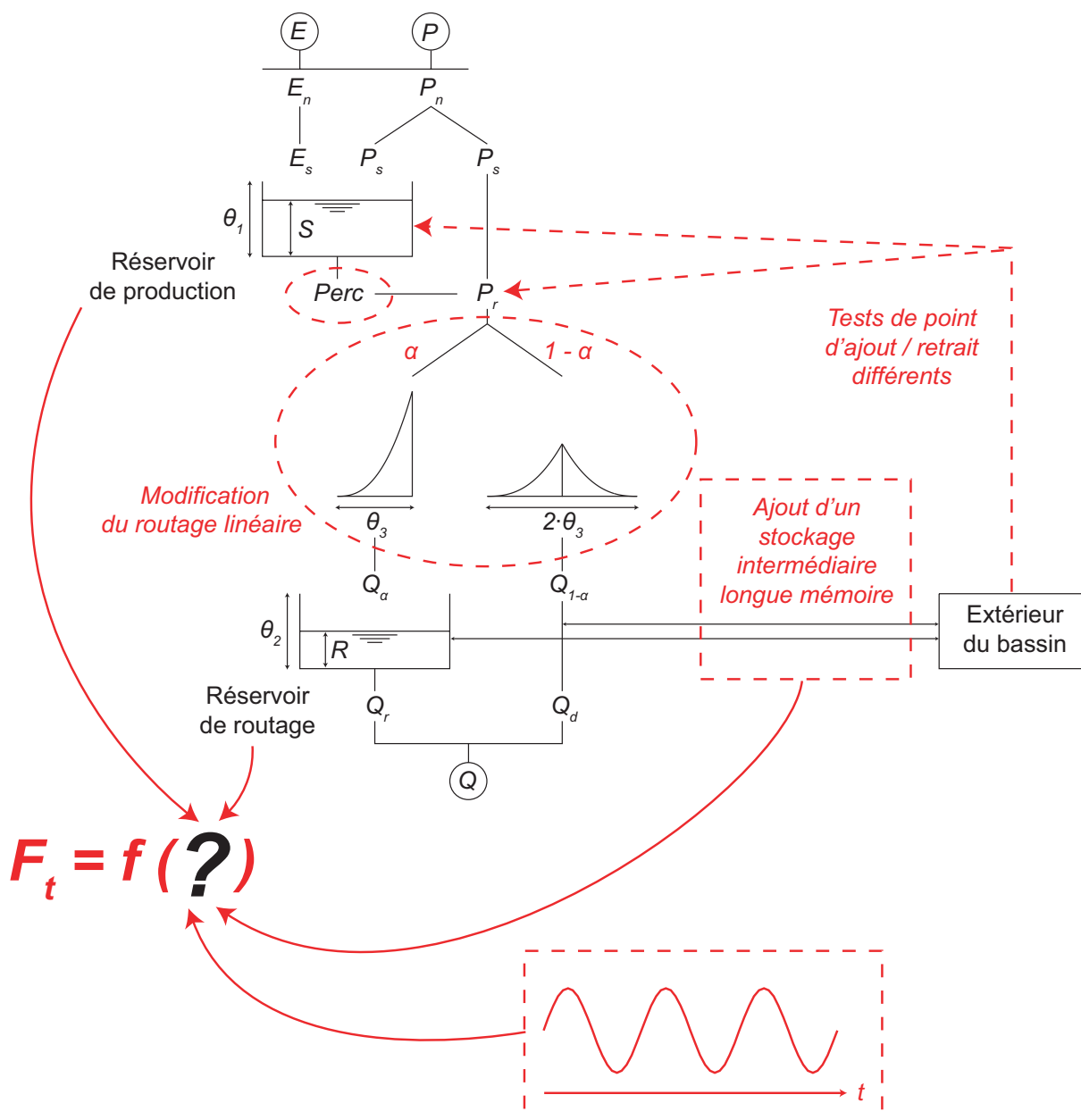


FIG. 8.2: Canevas des modifications apportées à la structure du modèle journalier GR4J.

- chaque période commence par une période d'initialisation d'un an,
- la méthode de calage utilisée pour identifier les paramètres est la méthode « Pas-à-Pas » avec préfiltrage des points de départ (Mathevet, 2005). Le préfiltrage est une méthode globale permettant de choisir le point de départ de l'optimisation locale, et ainsi minimiser le risque d'atteindre un optimum local. Nous choisissons k_i points de départ pour chaque paramètre θ_i , par exemple les quantiles $\left(\frac{1}{k_i+1}, \frac{2}{k_i+1}, \dots, \frac{k_i}{k_i+1}\right)$ de sa distribution (*a priori* s'il s'agit d'un nouveau paramètre jamais testé, *a posteriori* sinon), puis nous évaluons la fonction objectif aux $\prod_i k_i$ points de départ ainsi définis. Le nombre k_i pourra dépendre de la sensibilité du paramètre, les paramètres les plus sensibles étant préfiltrés plus finement. Nous choisissons les k_i de telle sorte $\prod_i k_i < 300$: à titre indicatif, nous utilisons typiquement quatre valeurs par paramètre pour un modèle à quatre paramètres ($4^4 = 256$ points de départ), ou trois par paramètre pour un modèle à cinq paramètres ($3^5 = 243$ points de départ). Le meilleur de ces points de départ est choisi pour initier la méthode « Pas-à-Pas »¹.
- **les $Nn(n-1)$ critères de contrôle obtenus sont résumés par leur moyenne.**

8.2.2 Quels progrès attendre d'une complexification ?

Avant de présenter les résultats des tests, il est nécessaire de s'accorder sur ce que l'on appelle une amélioration... Cette question n'est pas simple.

Nous nous la sommes déjà posée au Chapitre 3 lorsque nous avons comparé les performances de différentes solutions corrigeant les forçages climatiques ou la surface des bassins versants (cf. Figure 5, page 78, où les résultats sont encore exprimés en critère de Nash-Sutcliffe). À partir de quelle différence moyenne sur le critère C_{MM} ou Nash-Sutcliffe peut-on raisonnablement penser que nous avons progressé ? Cette question fera l'objet d'un chapitre entier (le onzième). Pour l'heure, nous avons déjà plusieurs méthodes à notre disposition pour estimer la significativité des différences. Une première solution (Mathevet, 2005) est de considérer les séries de critères de performance (en contrôle) des modèles pris deux-à-deux (M_i et M_j), et de choisir entre les termes de l'alternative suivante :

$$\left\{ \begin{array}{l} H_0 : \text{ Les deux séries sont issues de la même distribution, i.e. } M_i \text{ et } M_j \text{ sont } \textit{ex-aequo}, \\ H_1 : \text{ Les deux séries ne sont pas issues de la même distribution,} \\ \quad \text{i.e. l'un des modèles est supérieur à l'autre.} \end{array} \right.$$

¹Cette méthode suit le gradient de la fonction objectif et s'arrête lorsque le pas de recherche devient inférieur à un certain seuil.

Dans l'annexe B, nous proposons une méthode basée sur le test non-paramétrique de Wilcoxon « somme des rangs » (*Wilcoxon*, 1945 ; *Mann et Whitney*, 1947), pour tester cette alternative. Nous construisons les courbes empiriques de taux de rejet de l'hypothèse H_0 avec le risque de première espèce α et nous sélectionnons, en fonction de la taille de l'échantillon, la différence moyenne des performances qui donne un taux de rejet empirique supérieur à un seuil choisi. Les ordres de grandeurs auxquels nous arrivons sont tout à fait comparables à ceux avancés par *Mathevet* : de l'ordre d'un point (0.01) de critère C_{MM} pour un échantillon de plusieurs centaines de bassins.

Une seconde possibilité pour estimer la marge de progression que l'on peut atteindre est de chercher les complémentarités entre structures, au moyen d'une approche multi-modèle. Si l'on dispose de N modèles, on peut définir un « super-modèle » en sélectionnant, pour chaque bassin versant, celui qui donne les meilleurs résultats. Cependant, pour estimer correctement la marge de progression, cette sélection doit se faire en calage, pour chaque bassin versant (ou bassin-période) :

- On cale les N modèles sur le bassin considéré,
- On choisit celui qui donne le meilleur critère de calage,
- Le critère de contrôle du multi-modèle sur ce bassin est celui du modèle sélectionné, en contrôle.

Cette procédure (multi-modèle *a priori*) n'est donc pas tout à fait équivalente à celle qui prendrait le meilleur critère de contrôle pour chaque bassin-période : cette dernière façon de faire donne une autre estimation, un peu supérieure, de la marge de progression (multi-modèle *a posteriori*).

8.2.3 Catégories de modifications testées

Nous ne détaillerons pas toutes les structures testées mais nous donnons ici, dans les grandes lignes, les paramétrisations testées. Pour un maximum de clarté, précisons les notations utilisées :

1. les **paramètres optimisés** sont notés par les θ_i ,
2. les **états du modèle** sont les grandeurs indicées par le temps (x_t),
3. les autres lettres (A, a, b, α, ϕ , etc.) désignent des **constantes de structure**, dont nous cherchons éventuellement à optimiser la valeur de façon à maximiser les performances moyennes du modèle sur tous les bassins de l'échantillon, mais qui ne sont pas optimisées bassin par bassin (i.e., on ne les modifie pas dans le calage Pas-à-Pas).

En pratique, la recherche du meilleur compromis performance / parcimonie consiste à faire passer certaines valeurs de la catégorie 1 (paramètres optimisés) à la catégorie 3 (paramètres

fixes) ou inversement, en réduisant ou au contraire en augmentant le nombre de degrés de liberté d'une structure.

Formulation de la fonction d'échange

Les fonctions d'échange testées en premier lieu se divisent en cinq groupes :

- **f1** : des fonctions utilisant un seul paramètre, de forme $F_t = \theta_4 x_t^\alpha$ où x_t est le taux de remplissage d'un des réservoirs ($x_t = \frac{S_t}{\theta_1}$ ou $x_t = \frac{R_t}{\theta_2}$). Cette forme inclut l'échange constant, pour $\alpha = 0$,
- **f2** : des fonctions utilisant deux paramètres (θ_4 et θ_5) et dépendant du niveau d'un des réservoirs (S_t ou R_t), de la forme décrite à la Figure 8.3 gauche. L'exposant α de l'échange est le même quel que soit le signe de $x_t - \theta_5$.
- **f3** : des fonctions utilisant deux paramètres et dépendant du niveau d'un des réservoirs (S_t ou R_t), de la forme décrite à la Figure 8.3 droite. L'exposant de l'échange dépendant du signe de $x_t - \theta_5$: α si $(x_t - \theta_5) > 0$, β sinon,
- **f4** : des fonctions à deux paramètres, périodiques de période $T = 1$ an ($F_t = a \cos(\frac{2\pi}{T}t + \phi) + b$, les valeurs optimisées étant les couples $(\theta_4, \theta_5) = (a, b)$ ou $(\theta_4, \theta_5) = (a, \phi)$ mais pas les trois valeurs (a, b, ϕ) à la fois).

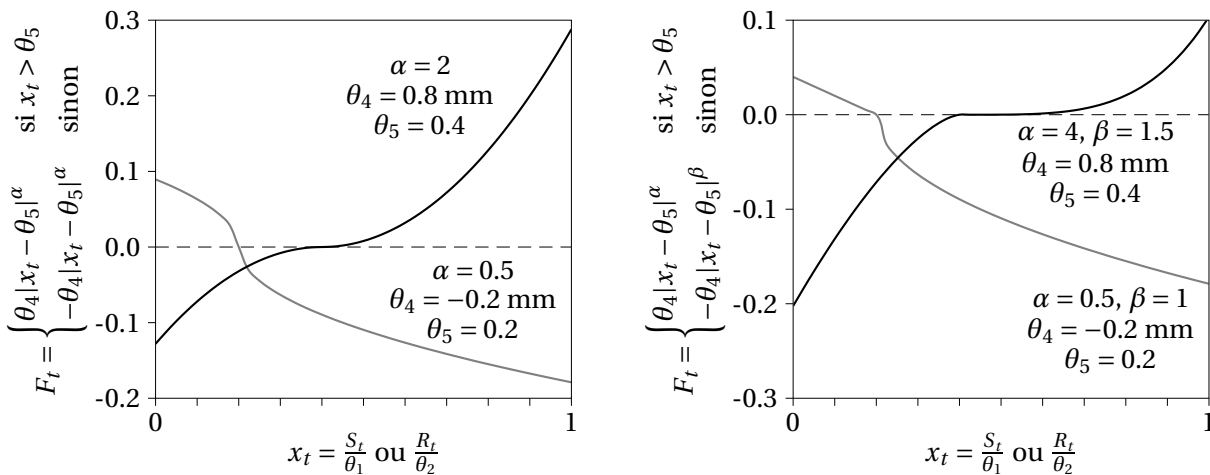


FIG. 8.3: Forme des fonctions d'échange du groupe **f2**, à gauche (même exposant à gauche et à droite du seuil de changement de signe θ_5) et du groupe **f3**, à droite (exposant différent pour $x_t > \theta_5$ ou $x_t < \theta_5$).

Points d'ajout

Différentes solutions d'ajout ont été testées. Nous utilisons la notation « (Point d'ajout si $F_t > 0$, Point d'ajout si $F_t < 0$) » pour les décrire :

- **p1** : point d'ajout invariant quelle que soit la valeur de F_t : $(S, S) =$ dans le réservoir S , $(R, R) =$ dans le réservoir R (et éventuellement à la branche rapide), $(Pr, Pr) =$ dans le terme de production Pr ,
- **p2** : point d'ajout dépendant du signe de F_t : (S, R) , (R, S) , (S, Pr) , (Pr, S) , (R, Pr) ou (Pr, R) .

Modification ou optimisation de certains paramètres fixes dans GR4J

Partant du principe que la modification de la forme de la fonction d'échange est susceptible de « déplacer » la valeur optimale de certaines autres valeurs fixes dans la structure, nous avons testé :

- **m1** : des modifications de la loi de « percolation »¹ du réservoir sol, introduite par Perrin (2000). Celle-ci s'écrit :

$$Perc_t = S_t \left(1 - \left[1 + \left(\frac{S_t}{A \cdot \theta_1} \right)^4 \right]^{-\frac{1}{4}} \right)$$

Nous avons donc cherché à faire varier la constante A (fixée à 2.25 dans GR4J) ou à l'optimiser, ainsi qu'à faire varier la puissance (4) de la loi de vidange,

- **m2** : des modifications du routage linéaire, notamment en remplaçant les deux hydrogrammes unitaires par un seul, de forme définie par un exposant fixe (à optimiser),
- **m3** : des modifications de la répartition entre les deux branches du routage : coefficient de partage $\alpha = \theta_6$, ou fonction du taux de remplissage du réservoir de routage : $\alpha = \theta_6 \left(\frac{R_t}{\theta_2} \right)^Y$

Toutes ces modifications n'ont pas été testées de façon systématiquement « croisée ». En particulier, nous n'avons pas croisé les solutions à point d'ajout variable et celles à puissance variable en fonction du signe de F_t , ce qui aurait multiplié les combinaisons (**f3**×**p2**) au-delà du raisonnable. Dans la suite, nous présentons les statistiques pour environ 80 structures issues de différentes combinaisons **f**×**p**×**m** (résumé non-exhaustif).

8.3 Résultats

8.3.1 Résultats sur l'échantillon complet (bassins français et australiens)

La Figure 8.4 résume les performances obtenues en contrôle pour ces 80 premières structures. Pour plus de lisibilité, les modèles ont été classés par performance moyenne décrois-

¹Nous avons repris le terme habituel pour désigner cette fonction, mais il crée une confusion fâcheuse avec le processus physique. Cette fonction ne correspond probablement à rien de tel.

sante pour chacun des critères utilisés ($C_{MM}(Q)$, $C_{MM}(\sqrt{Q})$ et $C_{MM}(\ln Q)$), cinq structures étant mises en lumière :

- la structure GR4J,
- les structures issues de GR3J auquel on a ajouté un paramètre de correction de surface (C_S), de pluie (C_P) ou d'ETP (C_P),
- la structure M_{043} , avec échange de forme $F_t = \theta_4 \left(\frac{R_t}{\theta_2} - \theta_5 \right)$, ajouté au réservoir R et à la branche directe quel que soit le signe de F_t , et ne possédant plus qu'un seul hydrogramme en tête des deux branches (la « fusion » des deux hydrogrammes unitaires n'occasionnant pas de diminution de performances). Cette structure se classe dans les quatre premières pour chacun des trois critères en contrôle, raison pour laquelle elle nous paraît la plus intéressante.

De façon générale, les quinze meilleures structures ont en commun d'avoir une fonction F_t dépendant du taux de remplissage du réservoir de routage, avec un ajout dans ce même réservoir quel que soit le signe de F_t . Les différences de performance entre les structures sont plus marquées sur le critère $C_{MM}(\ln Q)$, avec un écart-type de 0.032 (3.2 points) entre les moyennes des structures en contrôle, contre 1.8 points pour la moyenne des $C_{MM}(\sqrt{Q})$ et 1.3 point pour la moyenne des $C_{MM}(Q)$. Ceci est cohérent avec le fait que la fonction d'échange joue davantage sur les étiages que sur les crues ; cependant, les meilleures structures du point de vue du critère $C_{MM}(\ln Q)$ ne dégradent pas la simulation des crues et, au contraire, les améliorent même légèrement. Ainsi, le gain moyen en contrôle entre la structure GR4J et la structure M_{043} est 3.1 points de $C_{MM}(\ln Q)$ (significatif), de 0.9 point de $C_{MM}(\sqrt{Q})$ (en limite de significativité) et de 0.5 point de $C_{MM}(Q)$ (pas significatif). Au vu de ces résultats, nous nous baserons préférentiellement sur le critère $C_{MM}(\ln Q)$ pour évaluer la pertinence des différentes solutions, en prenant toutefois garde à ce que les deux autres critères ne soient pas dégradés.

Notons que les structures donnant les moins bons résultats sont celles faisant intervenir une fonction périodique dont les paramètres calés sont l'amplitude et la phase. Cela tient probablement à deux phénomènes :

- le fait d'avoir une moyenne annuelle nulle est une contrainte trop forte,
- le paramètre de phase présente de gros problèmes d'identifiabilité.

Dans le cas où la phase a été fixée (optimisation de la moyenne et de l'amplitude), les performances sont améliorées mais ne dépassent pas celles du modèle initial.

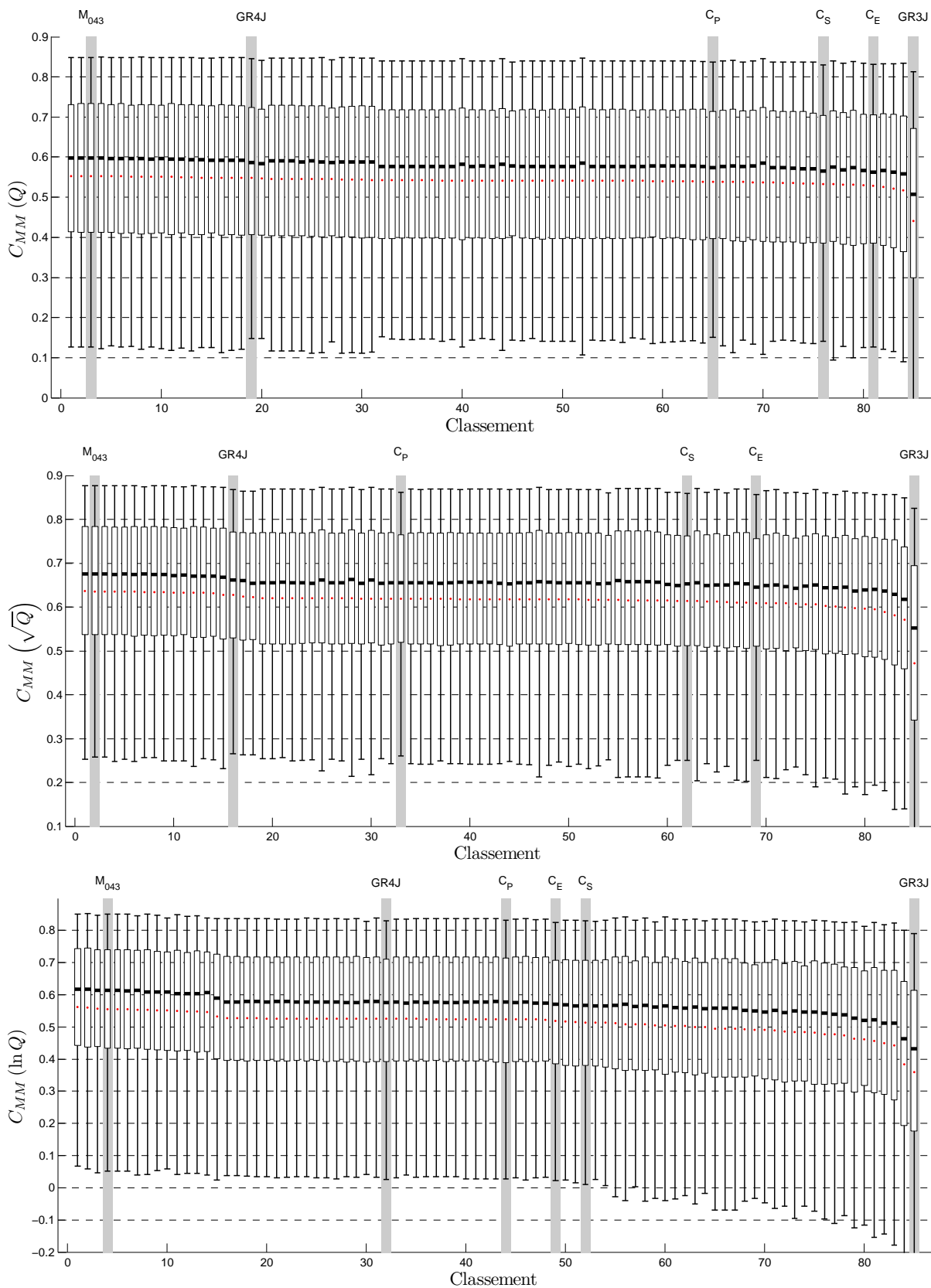


FIG. 8.4: Distributions de performances en contrôle pour 80 structures issues de la première série de tests journaliers. Résultats sur l'échantillon complet de bassins français et australien (1389 bassins versants, 4070 périodes de contrôle de 5 ans). GR3J : modèle GR4J dont la fonction de perte est désactivée (θ_4 fixé à zéro); C_P : GR3J + correction de pluie; C_E : GR3J + correction d'ETP; C_S : GR3J + correction de surface.

8.3.2 Résultats différenciés par sous-échantillon

Sur le graphique de la Figure 8.5, nous avons repéré chaque structure : en abscisse, par ses performances moyennes (toujours en contrôle) sur les bassins versants de l'échantillon français, et en ordonnées par ses performances moyennes sur les bassins australiens. Nous constatons, ce qui est plutôt rassurant, que les structures les plus performantes sur l'un des échantillons sont aussi les plus performantes sur l'autre échantillon. Sans surprise, compte tenu du plus grand contraste du climat australien, les performances moyennes du modèle sont plus faibles sur l'échantillon australien que sur l'échantillon français. Il est d'ailleurs intéressant de noter que les solutions « correctives » (surface, pluie, ETP) marchent relativement mieux sur l'échantillon australien que sur l'échantillon français, cette différence étant peut-être due au post-traitement de l'information pluviométrique dans l'échantillon français : en effet, l'information pluviométrique journalière des bassins australiens est donnée par une simple moyenne pondérée de pluviomètres, tandis que celle des bassins français est issue de la réanalyse SAFRAN (cf. Chapitre 2).

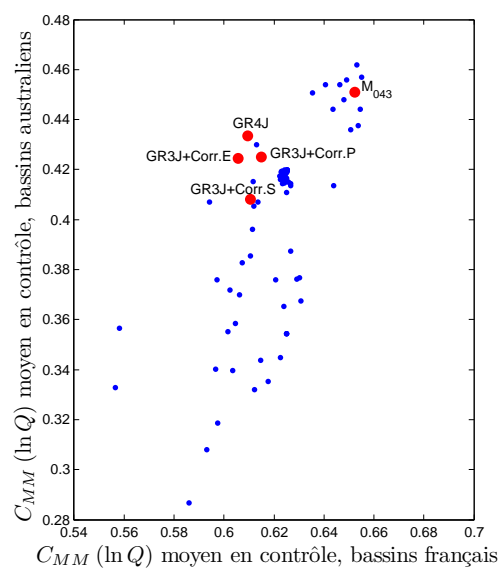
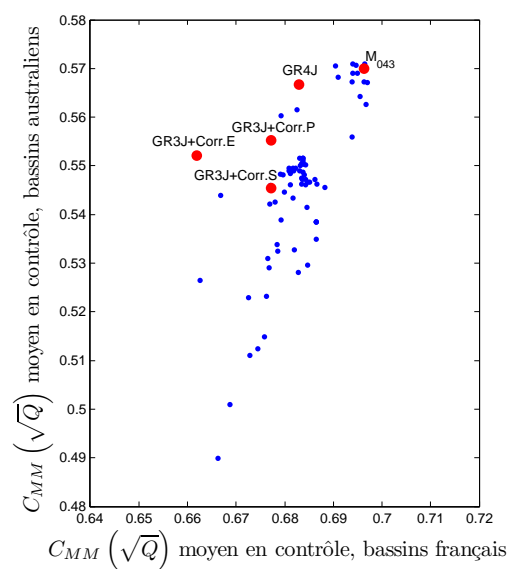
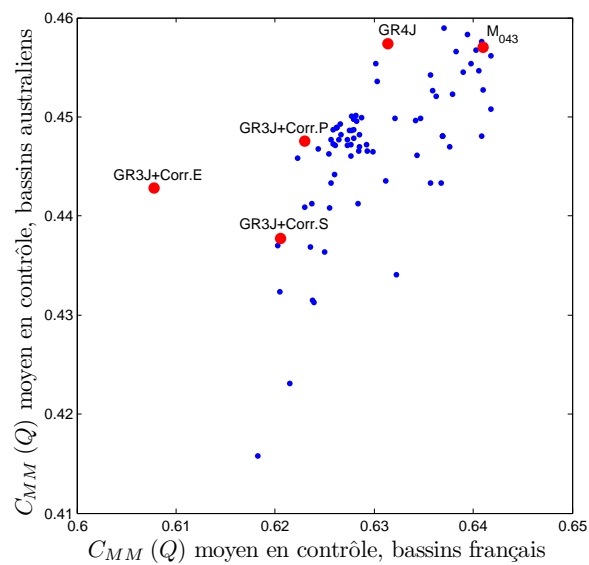


FIG. 8.5: Performances différenciées de chaque structure sur les deux sous-échantillons français et australien.

8.3.3 Cartographie de l'effet des modifications

Nous avons expliqué au Chapitre 3 que l'approche choisie comportait un certain nombre de risques, au premier rang desquels le risque que la fonction d'échange serve à corriger une multitude d'autres sources d'erreurs de bilan ou de dynamique. Pour vérifier s'il en est ainsi, nous avons cartographié sur la Figure 8.6, pour chaque bassin de l'échantillon français, l'écart-type des performances (en contrôle) de toutes les structures testées. Là où cet écart-type est faible, les modifications n'ont aucun effet ; là où il est fort, les résultats sont effectivement sensibles à la formulation des échanges. Nous constatons sur cette carte que les bassins sensibles à la formulation des échanges se regroupent selon une logique régionale que l'on peut relier assez précisément à la géologie. Ce résultat est donc plutôt rassurant : si nous sommes bien incapables de dire *pourquoi* telle ou telle formulation des échanges est supérieure à l'autre, nous pouvons cependant être relativement confiants sur le fait que cette fonction semble bien, dans la majorité des cas, jouer le rôle qu'on lui attribue conceptuellement. En particulier, les modifications testées semblent neutres dans les zones où l'on peut supposer des problèmes d'estimation des précipitations et/ou des influences nivales :

- Cévennes,
- Alpes du Nord (Savoie et Haute-Savoie),
- Corse,
- Massif Central

A contrario, la forme de la fonction d'échange influence fortement les performances dans les régions suivantes :

- Picardie et Haute Normandie (nappe de la Craie, tons verts pâles sur la carte géologique),
- Beauce (calcaires éocènes, en beige),
- Vosges (Nappe des Grès du Trias, tons mauves),
- Alsace (nappe alluviale du Rhin, tons très clairs),
- Régions karstiques (calcaires du Jurassique, franges bleues) : Doubs, Jura, Ain ; Quercy, Périgord, Charente ; Causses ; Vaucluse,
- secteur des vallées de Vienne (affluents rive Est du Rhône situés sur de grandes épaisseurs de molasses).

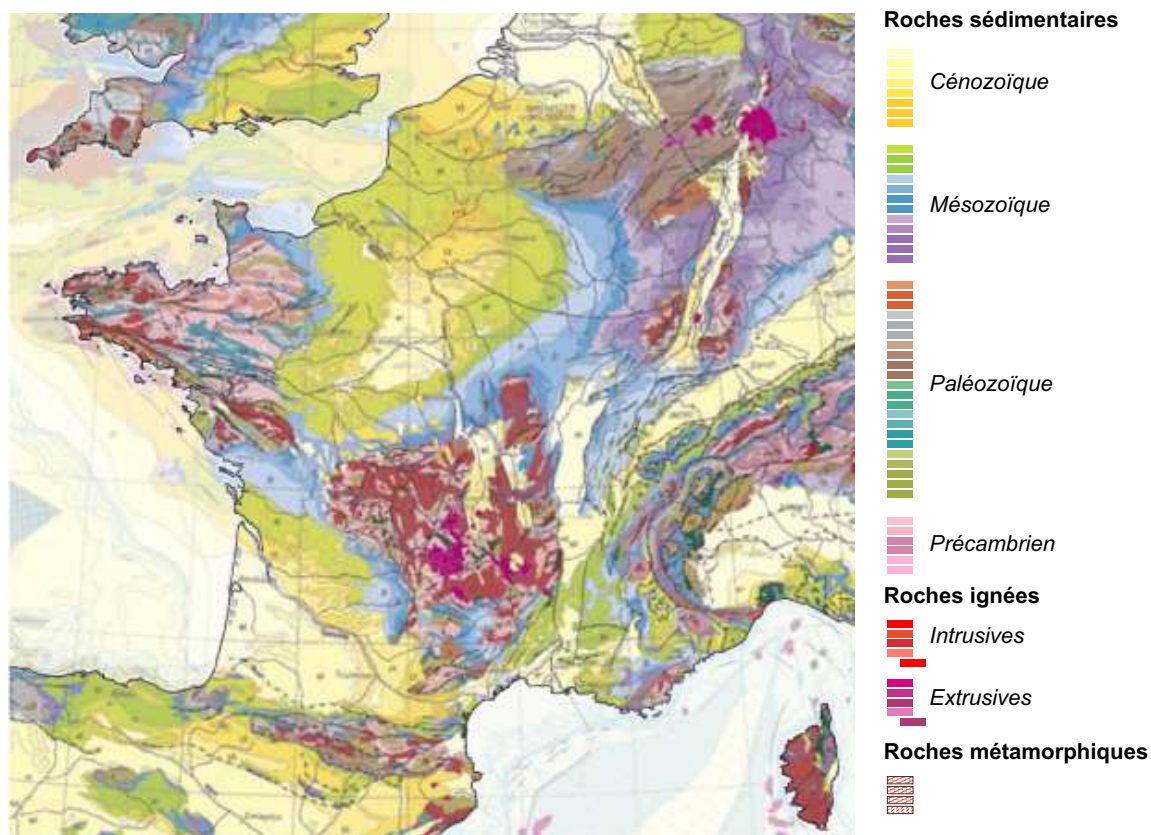
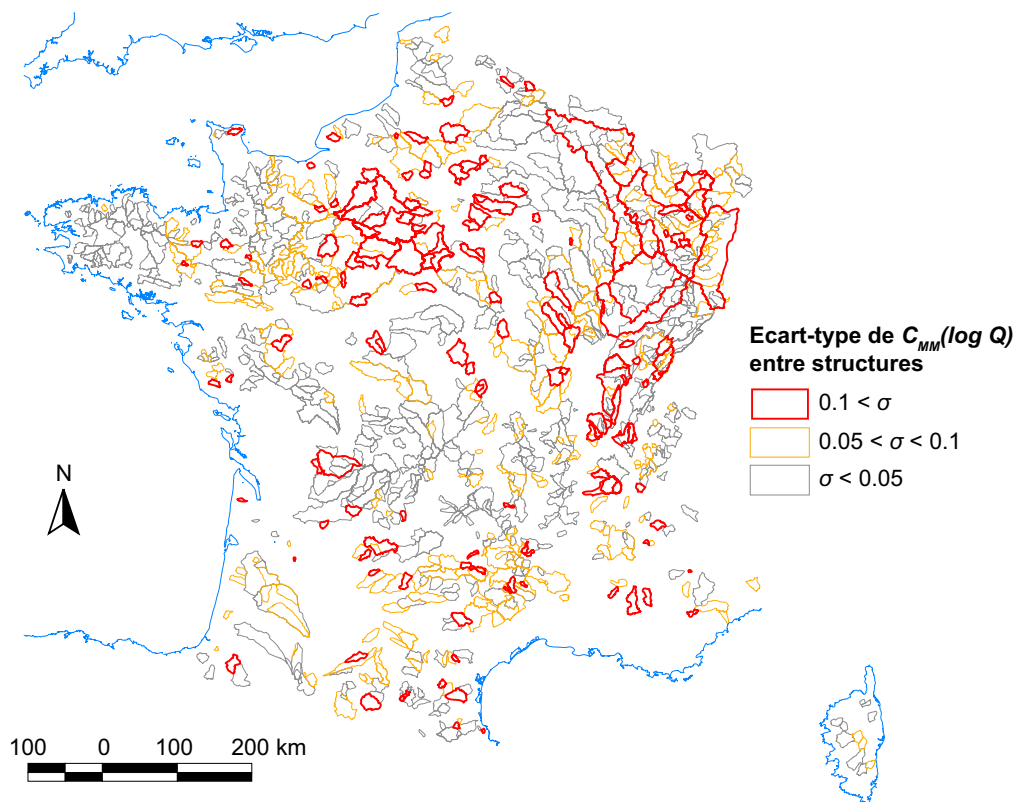


FIG. 8.6: En haut : cartographie de la sensibilité des bassins français à la formulation des échanges (écart-type des performances des 80 structures testées). Les contours gris indiquent un écart-type inférieur à 0.05 (5 points de C_{MM} en contrôle), les contours en orange indiquent un écart-type compris entre 0.05 et 0.1, et les contours rouges épais indiquent un écart-type supérieur à 0.1 (10 points de C_{MM}). On constate très clairement une logique régionale, très liée à la géologie (carte du bas, d'après *Asch*, 2005).

8.3.4 Recherche de complémentarités entre structures

Les Tableaux 8.1 et 8.2 présentent les performances moyennes des différentes combinaisons deux-à-deux de 16 structures potentiellement complémentaires (la meilleure dans chaque classe de modifications testées). Dans le Tableau 8.1, le choix du modèle est fait *a priori* (i.e. en calage), tandis que dans le Tableau 8.2 la sélection est effectuée en contrôle (multi-modèle *a posteriori*). Dans les deux cas, le gain est limité : il ne semble pas y avoir de complémentarités notables entre structures prises par paires.

	GR4J	M_{006}	M_{012}	M_{016}	M_{024}	M_{027}	M_{035}	M_{037}	M_{043}	M_{048}	M_{055}	M_{058}	M_{065}	M_{067}	M_{075}	M_{076}
GR4J	0.525	0.530	0.528	0.531	0.564	0.536	0.543	0.537	0.564	0.536	0.567	0.536	0.538	0.537	0.545	0.536
M_{006}		0.518	0.525	0.522	0.565	0.529	0.539	0.529	0.564	0.529	0.567	0.529	0.533	0.530	0.541	0.530
M_{012}			0.491	0.525	0.562	0.533	0.523	0.533	0.563	0.533	0.566	0.533	0.517	0.533	0.520	0.533
M_{016}				0.513	0.564	0.530	0.539	0.530	0.564	0.530	0.567	0.530	0.534	0.530	0.540	0.530
M_{024}					0.556	0.565	0.558	0.563	0.561	0.564	0.564	0.564	0.557	0.563	0.560	0.563
M_{027}						0.527	0.542	0.528	0.565	0.528	0.567	0.528	0.538	0.529	0.543	0.529
M_{035}							0.512	0.540	0.557	0.541	0.562	0.541	0.514	0.540	0.519	0.540
M_{037}								0.525	0.563	0.529	0.566	0.529	0.537	0.526	0.542	0.527
M_{043}									0.556	0.564	0.562	0.564	0.557	0.562	0.559	0.563
M_{048}										0.527	0.568	0.527	0.538	0.529	0.543	0.529
M_{055}											0.561	0.567	0.561	0.566	0.564	0.566
M_{058}												0.527	0.538	0.529	0.543	0.529
M_{065}													0.503	0.536	0.516	0.535
M_{067}														0.526	0.542	0.528
M_{075}															0.509	0.542
M_{076}																0.525

TAB. 8.1: Performances moyennes des différentes combinaisons *a priori* deux-à-deux des 16 meilleurs représentants des classes de modifications testées (résultats en $C_{MM}(\ln Q)$).

	GR4J	M_{006}	M_{012}	M_{016}	M_{024}	M_{027}	M_{035}	M_{037}	M_{043}	M_{048}	M_{055}	M_{058}	M_{065}	M_{067}	M_{075}	M_{076}
GR4J	0.525	0.537	0.544	0.538	0.587	0.549	0.568	0.550	0.588	0.548	0.588	0.548	0.566	0.549	0.571	0.549
M_{006}		0.518	0.538	0.525	0.587	0.540	0.564	0.541	0.587	0.539	0.587	0.539	0.562	0.541	0.566	0.541
M_{012}			0.491	0.541	0.583	0.552	0.541	0.553	0.583	0.552	0.583	0.552	0.538	0.553	0.541	0.553
M_{016}				0.513	0.587	0.541	0.564	0.542	0.588	0.540	0.587	0.540	0.562	0.542	0.566	0.541
M_{024}					0.556	0.583	0.571	0.582	0.567	0.583	0.571	0.583	0.570	0.583	0.571	0.582
M_{027}						0.527	0.564	0.534	0.584	0.531	0.583	0.531	0.563	0.534	0.565	0.535
M_{035}							0.512	0.563	0.571	0.564	0.574	0.564	0.522	0.563	0.529	0.562
M_{037}								0.525	0.583	0.534	0.583	0.534	0.562	0.529	0.565	0.531
M_{043}									0.556	0.584	0.571	0.584	0.570	0.584	0.570	0.583
M_{048}										0.527	0.584	0.527	0.563	0.534	0.566	0.534
M_{055}											0.561	0.584	0.575	0.583	0.573	0.583
M_{058}												0.527	0.563	0.534	0.566	0.534
M_{065}													0.503	0.562	0.530	0.561
M_{067}														0.526	0.565	0.532
M_{075}															0.509	0.564
M_{076}																0.525

TAB. 8.2: Performances moyennes des différentes combinaisons *a posteriori* deux-à-deux des 16 meilleurs représentants des classes de modifications testées (résultats en $C_{MM}(\ln Q)$).

Sur la Figure 8.7, nous avons représenté la distribution des performances des multi-modèles *a priori* et *a posteriori* quand on choisit non plus entre deux structures à chaque fois, mais parmi les 80 structures. La structure M_{043} est à 1.7 point seulement du multi-modèle *a priori*, mais la différence de 7.2 points avec le multi-modèle *a posteriori* indique une marge de progression importante au niveau de la robustesse (rehaussement important du cinquième centile (moustache inférieure), indiquant une limitation des forts échecs du modèle).

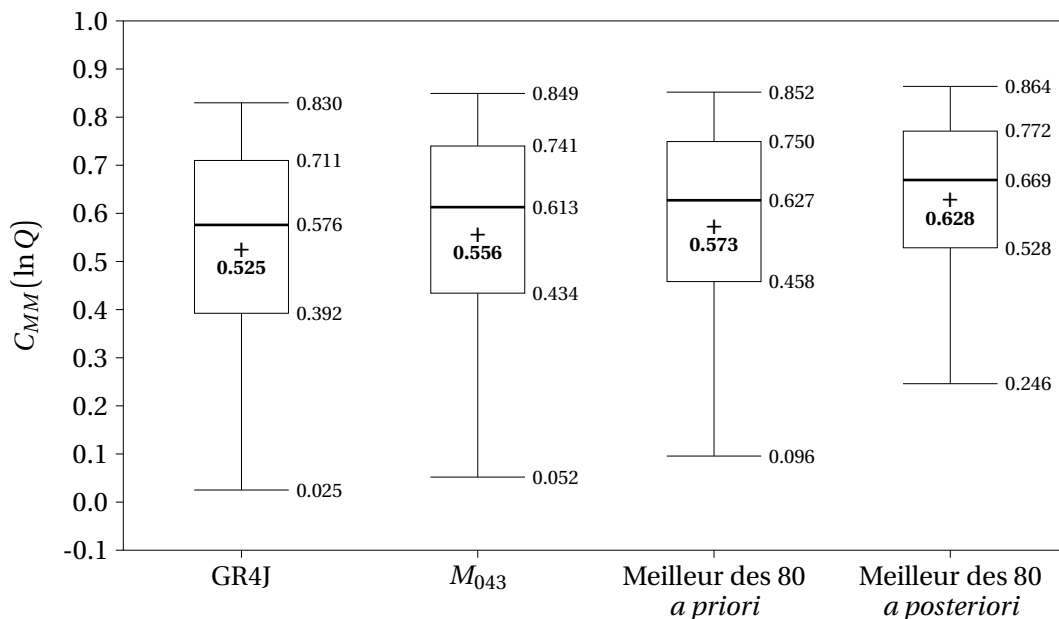


FIG. 8.7: Performances comparées des structures GR4J, M_{043} et des multi-modèles « totaux » : meilleure des 80 structures, sélectionnée soit en calage (*a priori*) soit en contrôle (*a priori*).

8.3.5 Un premier modèle intermédiaire : GR5J ?

Les tests précédents nous inciteraient à privilégier la structure M_{043} , faisant appel à une fonction d'échange linéaire dépendant du niveau du réservoir de routage. Le routage linéaire a également été simplifié, avec la fusion des deux hydrogrammes unitaires en un seul (symétrique), confirmant les tests de *Mathevet* (2005). Les différences de performances sur l'échantillon français sont cartographiées sur la Figure 8.8. On observe bien une amélioration moyenne, mais certains bassins sont tout de même très dégradés par rapport à la version GR4J. Le changement de structure semble « faire l'unanimité » sur les bassins de l'Est (Vosges et Alsace). En revanche les bassins de la Craie présentent à la fois de fortes améliorations et de fortes dégradations, dont l'interprétation est d'ailleurs rendue difficile (1) par la faiblesse des performances absolues sur ces bassins, et (2) par des problèmes d'initialisation des modèles dans certains cas, que nous détaillerons au Chapitre 10. De façon à nous assurer que le changement de fonction d'échange n'a pas « déplacé » l'optimum des paramètres fixes de la percolation, nous avons testé différentes valeurs de la constante et de l'exposant de la percolation autour de leurs valeurs dans la structure GR4J. Les résultats sont représentés sur la Figure 8.9 : le point (4, 2.25) semble rester un bon compromis pour maximiser les trois critères $\overline{C_{MM}}(Q)$, $\overline{C_{MM}}(\sqrt{Q})$ et $\overline{C_{MM}}(\ln Q)$.

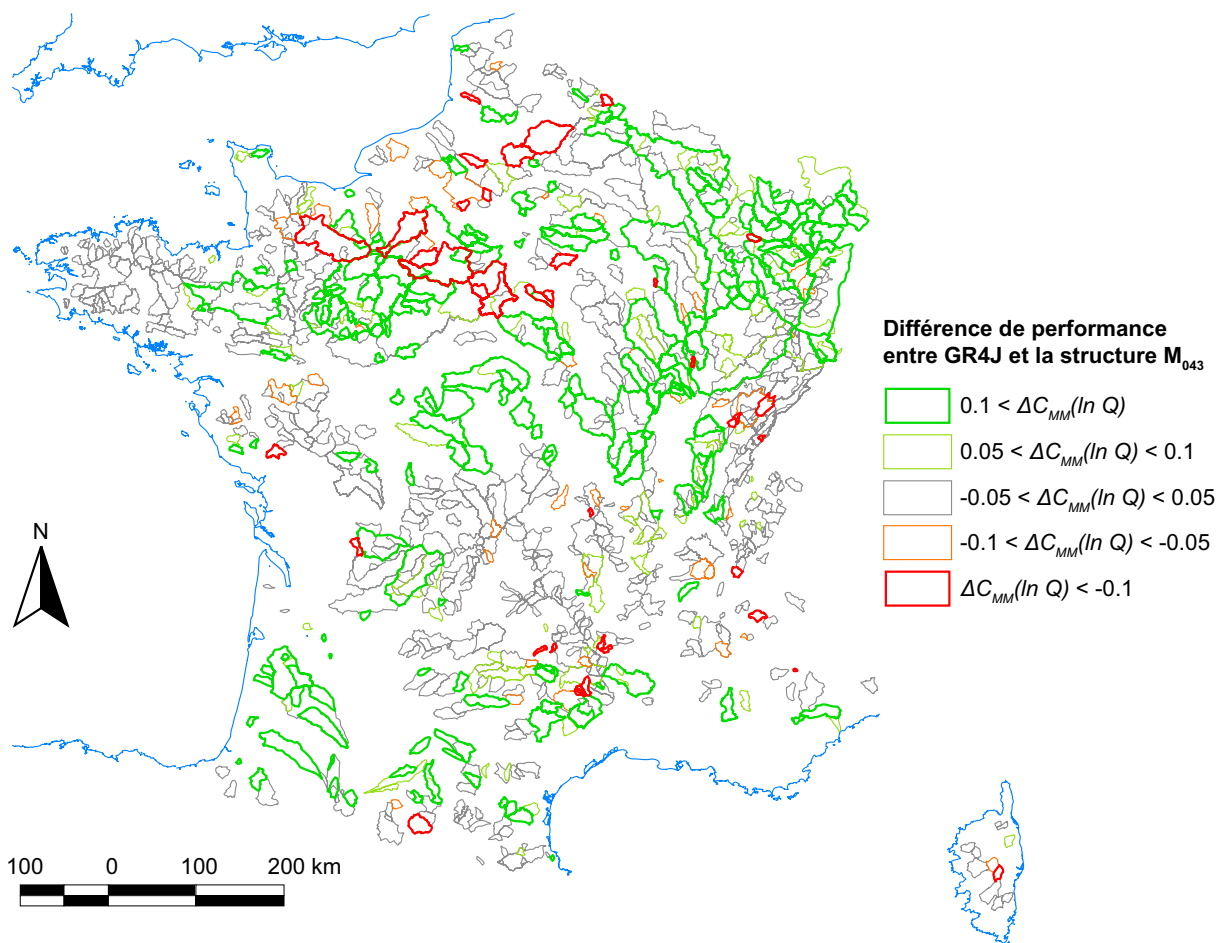


FIG. 8.8: Cartographie des différences de performance sur le critère $C_{MM}(\ln Q)$ en contrôle entre la version point de départ GR4J et la structure M_{043} .

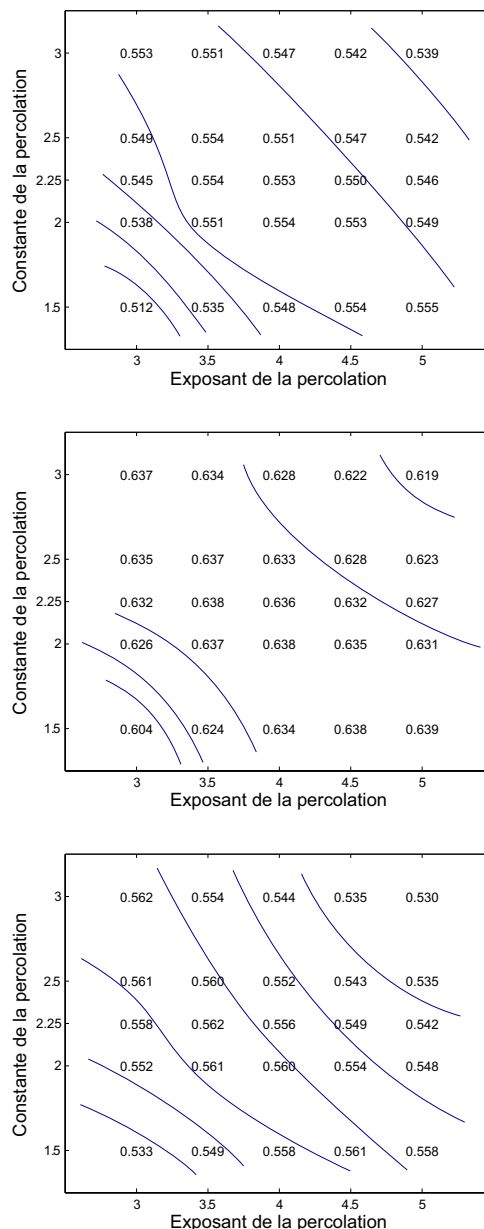


FIG. 8.9: Variation des critères moyens sur Q (haut), \sqrt{Q} (milieu) et $\ln Q$ (bas) de la structure M_{043} lorsque l'on change l'exposant et/ou la constante de la fonction de percolation autour de leurs valeurs d'origine (respectivement 4 et 2.25).

Nous testons encore deux modifications à la marge sur cette structure M_{043} :

- la suppression d'un degré de liberté, par la fixation du paramètre θ_5 (taux de remplissage du réservoir de routage pour lequel l'échange s'inverse) à la médiane sur l'échantillon,
- l'ajout d'un degré de liberté, par l'optimisation du coefficient de partage entre les deux branches du routage.

Les résultats sont représentés dans le Tableau 8.3. Nous constatons que la suppression du degré de liberté θ_5 (fixé à sa médiane, 0.34) pénalise beaucoup la structure, dont les performances retombent nettement en dessous de celles de GR4J ; à l'inverse, l'optimisation

du coefficient de partage ne permet pas un gain de performances significatif. À ce stade, nous pouvons donc conclure que la structure M_{043} offre un bon compromis entre complexité et performances. Dans la section 8.6, nous proposerons une évaluation plus approfondie de la différence entre M_{043} (notée GR5J) et GR4J.

	GR4J	M_{043}	M_{043} , θ_5 fixé	M_{043} + optimisation coeff. partage
$C_{MM}(Q)$ moyen	0.548	0.553	0.542	0.555
$C_{MM}(\sqrt{Q})$ moyen	0.627	0.636	0.615	0.638
$C_{MM}(\ln Q)$ moyen	0.525	0.556	0.498	0.561

TAB. 8.3: Tentatives de réduction ou d'augmentation d'un degré de liberté dans la structure M_{043} .

8.4 Vers une plus grande complexité ?

8.4.1 Peut-on limiter les flux échangés en apparence, en incluant un stockage longue mémoire ?

La sélection de la structure M_{043} , autorisant le bassin à perdre ou gagner de l'eau selon le taux de remplissage du réservoir de routage, nous amène au constat suivant : selon cette formulation, nous pouvons tout à fait avoir des échanges positifs pendant une partie de l'année, négatifs pendant l'autre partie, de sorte que le bilan net sur une année soit nul en moyenne. Il est tout à fait plausible d'imaginer que cette situation corresponde bien à la réalité : le test de semi-distribution du Chapitre 6 a montré qu'une telle solution n'était pas forcément incohérente du point de vue hydrologique (elle consiste à se placer rigoureusement dans le cadre de la Figure 3.12, page 66, et à compter tous les flux entrants et sortants du système strictement superficiel). Cependant, cette formulation pourrait aussi être critiquée, puisqu'il pourrait s'agir, dans certains cas, d'un artefact : le modèle ne « sachant » pas où stocker / déstocker certains volumes, il aurait tendance à exporter / importer ces volumes au gré de ses besoins pour reproduire l'hydrogramme observé. Dans un tel cas, il serait certainement préférable de « réintégrer » ces flux *dans* notre représentation du système, par exemple sous forme d'échanges entre le réservoir de routage et un réservoir à la dynamique plus lente, de sorte que l'échange soit toujours du même signe (voire négligeable pour la plupart des bassins) avec le compartiment « extérieur du bassin »¹.

Nous testons donc un nouveau groupe de fonctions d'échange, dépendant de l'état d'un réservoir L supplémentaire :

¹ou supposé comme tel, la prudence étant toujours de mise dès que l'on essaye de projeter des éléments du modèle sur des éléments de la réalité

- **f5** : F_t fonction d'un ou deux paramètre(s) et dépendant du niveau L_t du nouveau réservoir, les entrées et sorties de ce nouveau réservoir étant régies par d'autres paramètres supplémentaires.

Nous testons aussi par conséquent un nouveau point possible d'ajout du terme F_t :

- **p3** : ajout du terme d'échange dans le nouveau réservoir, i.e. ajout (L, L) .

Étant donnés les résultats de la première série de tests, nous ne testons pas autant de combinaisons (**f**×**p**). Nous supposons en particulier que la forme linéaire est la plus robuste, et nous testons donc des formes $F_t = \theta_4 (x_t - \theta_5)$ avec $x_t = \frac{S_t}{\theta_1}$ ou $x_t = \frac{R_t}{\theta_2}$ (échange dépendant du niveau du réservoir R ou S mais ajouté au réservoir L) ; ou alors $F_t = \theta_4 (L_t - \theta_5)$, i.e. échange contrôlé par L et ajouté à ce même réservoir L .

En pratique nous avons substitué le réservoir L à la branche de routage directe, et nous avons défini ce réservoir comme un réservoir linéaire de constante de temps θ_7 . Pour l'alimentation de ce réservoir, nous avons laissé libre le coefficient de partage entre les deux branches, désormais noté θ_6 . Les modifications testées sont résumées sur la Figure 8.10.

Cette série de tests porte donc sur des modèles à sept paramètres, ou six si le paramètre θ_5 est fixé à zéro (par exemple $F_t = \theta_4 L_t$ avec ajout à L , structure M_{081}).

8.4.2 Résultats de la deuxième série de tests

La Figure 8.11 montre les distributions de critères obtenus en contrôle pour 11 des structures de cette deuxième série de tests, classées par performances décroissantes. Nous rappelons sur la même figure les distributions obtenues pour GR4J et la structure M_{043} . Nous constatons sur cette figure qu'une seule structure à réservoir supplémentaire (M_{096} , $\overline{C_{MM}}(\sqrt{Q}) = 0.636$) parvient à faire jeu égal avec les deux modèles précédents. Il s'agit de la solution :

- fonction d'échange de même forme que la structure M_{043} : $F_t = \theta_4 \left(\frac{R_t}{\theta_2} - \theta_5 \right)$,
- ajout de F_t dans les deux réservoirs R et L .

Le point remarquable de cette série de test est que toutes les tentatives pour faire dépendre l'échange du niveau L_t du nouveau réservoir échouent. Il faut cependant nuancer ces échecs par le fait que cette série de structures à réservoir plus lent ont des performances potentiellement très dépendantes de la longueur d'initialisation (ce qui peut être vue, d'une certaine façon, comme un défaut).

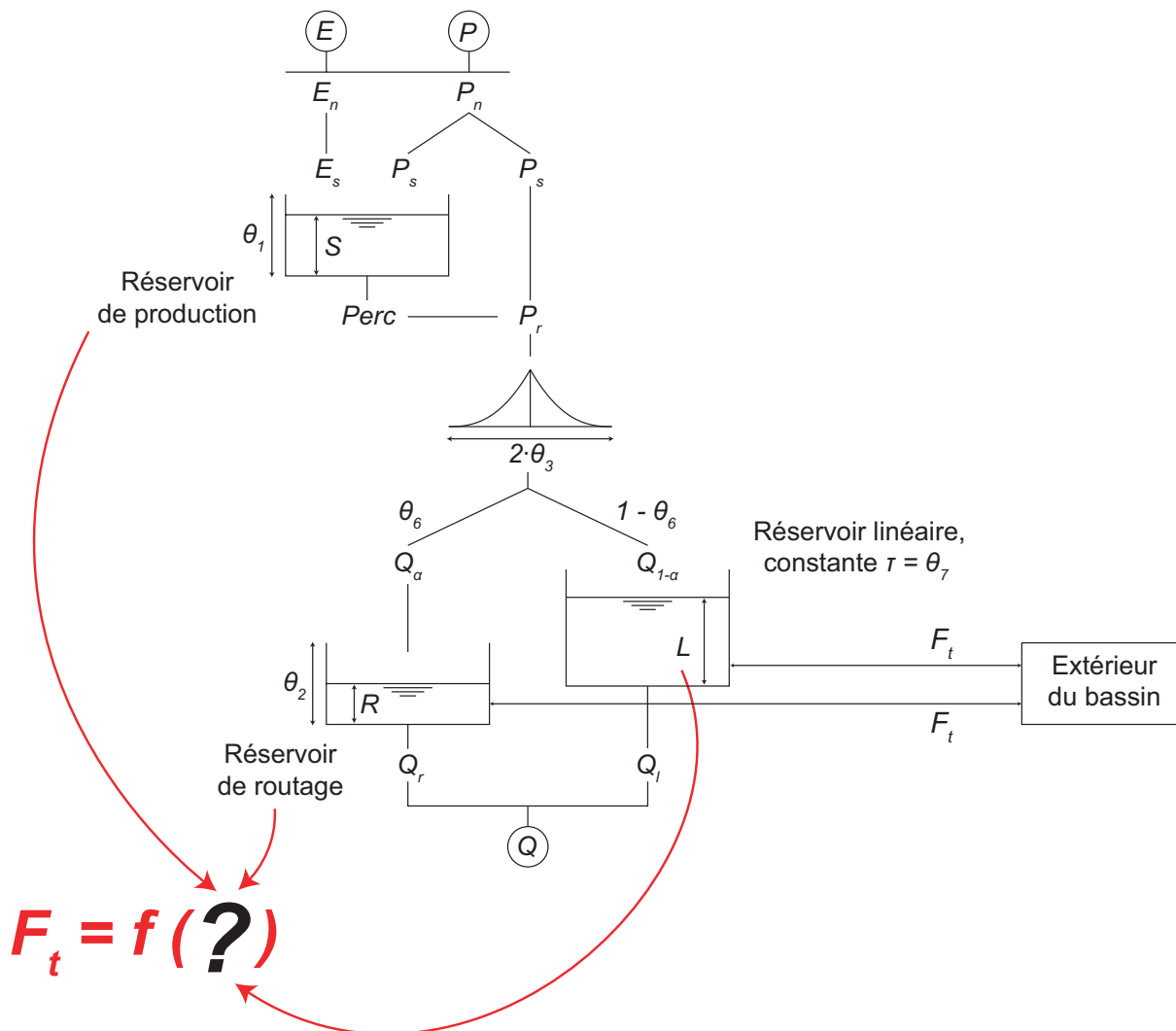


FIG. 8.10: Canevas des modifications basées sur l'introduction d'un réservoir supplémentaire pour prendre en compte les stockages longs au sein du bassin.

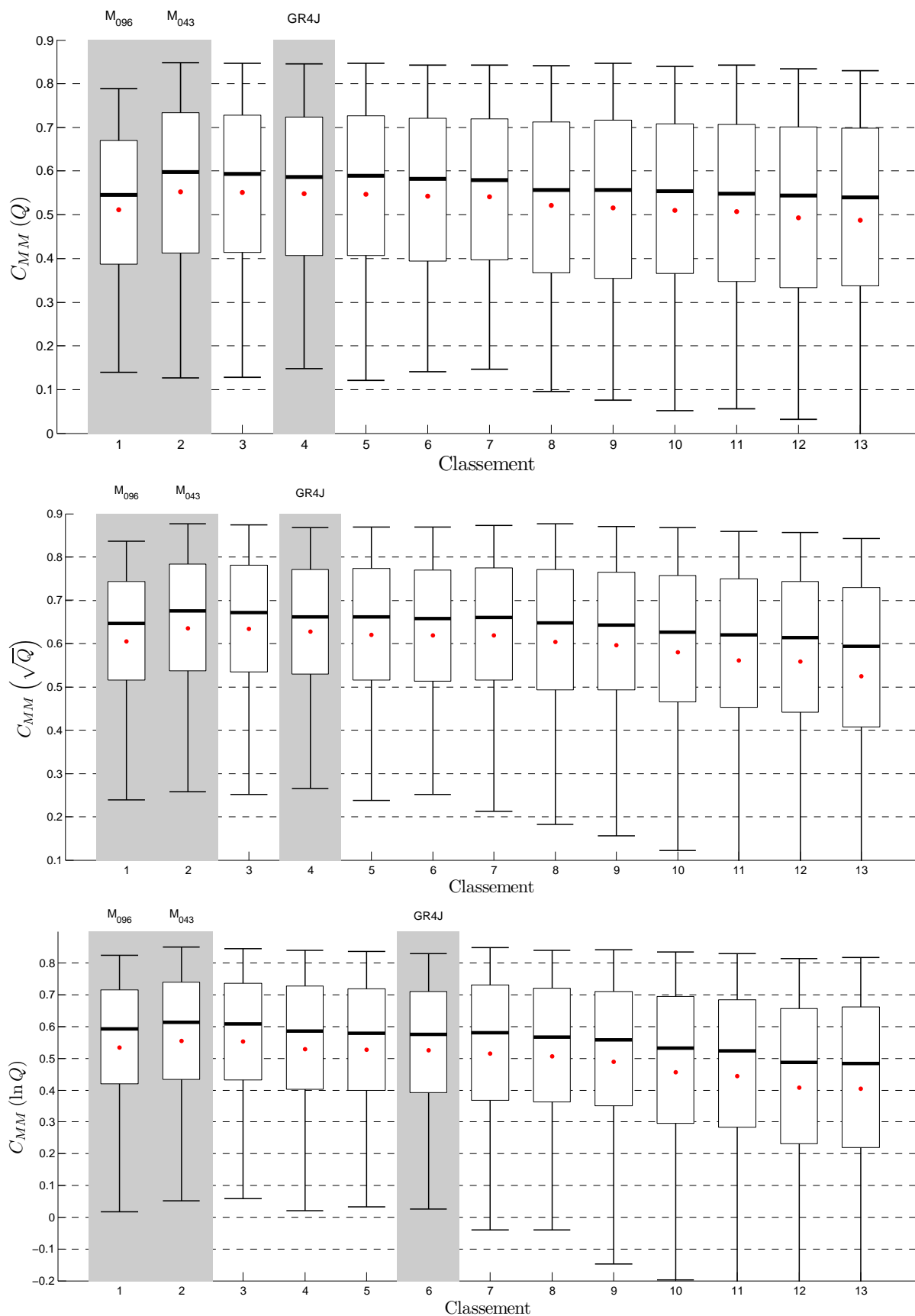


FIG. 8.11: Distributions de performances pour 11 structures issues de la deuxième série de tests journaliers (ajout d'un réservoir supplémentaire). Résultats sur l'échantillon complet de bassins français et australiens (1389 bassins versants, 4070 périodes de contrôle de 5 ans).

Comme précédemment, nous recherchons les complémentarités éventuelles entre structure par le biais des paires de modèles (choix *a priori*, Tableau 8.4, ou *a posteriori*, Tableau 8.5), puis par les multi-modèles totaux (ici 13 structures), sur la Figure 8.12. Là encore une marge de progression existe en terme de robustesse, étant donnée la différence de performances entre le multi-modèle *a priori* et le multi-modèle *a posteriori*.

	GR4J	M_{043}	M_{082}	M_{083}	M_{084}	M_{085}	M_{086}	M_{087}	M_{088}	M_{089}	M_{090}	M_{091}	M_{096}
GR4J	0.525	0.563	0.529	0.538	0.556	0.533	0.538	0.544	0.551	0.559	0.547	0.542	0.575
M_{043}		0.556	0.559	0.560	0.567	0.562	0.563	0.567	0.561	0.564	0.562	0.560	0.570
M_{082}			0.404	0.460	0.519	0.444	0.470	0.511	0.536	0.557	0.531	0.512	0.569
M_{083}				0.456	0.519	0.470	0.483	0.517	0.540	0.559	0.535	0.519	0.571
M_{084}					0.516	0.522	0.523	0.530	0.550	0.567	0.549	0.538	0.574
M_{085}						0.407	0.446	0.493	0.534	0.558	0.530	0.514	0.573
M_{086}							0.444	0.494	0.536	0.559	0.533	0.519	0.573
M_{087}								0.489	0.541	0.562	0.538	0.530	0.576
M_{088}									0.530	0.563	0.534	0.544	0.573
M_{089}										0.554	0.562	0.555	0.575
M_{090}											0.527	0.540	0.574
M_{091}												0.507	0.571
M_{096}													0.568

TAB. 8.4: Performances moyennes des différentes combinaisons *a priori* des 11 structures précédentes ainsi que des structures GR4J et M_{043} (résultats en $C_{MM}(\ln Q)$ moyen).

	GR4J	M_{043}	M_{082}	M_{083}	M_{084}	M_{085}	M_{086}	M_{087}	M_{088}	M_{089}	M_{090}	M_{091}	M_{096}
GR4J	0.525	0.587	0.551	0.559	0.576	0.553	0.556	0.560	0.570	0.580	0.566	0.564	0.598
M_{043}		0.556	0.578	0.581	0.586	0.583	0.587	0.591	0.575	0.576	0.579	0.576	0.577
M_{082}			0.404	0.472	0.535	0.495	0.513	0.541	0.562	0.575	0.559	0.534	0.589
M_{083}				0.456	0.533	0.514	0.524	0.547	0.565	0.579	0.562	0.543	0.591
M_{084}					0.516	0.546	0.551	0.559	0.572	0.587	0.572	0.561	0.594
M_{085}						0.407	0.449	0.500	0.551	0.575	0.548	0.540	0.593
M_{086}							0.444	0.499	0.555	0.578	0.551	0.546	0.597
M_{087}								0.489	0.559	0.583	0.557	0.558	0.600
M_{088}									0.530	0.581	0.542	0.567	0.586
M_{089}										0.554	0.580	0.566	0.586
M_{090}											0.527	0.564	0.589
M_{091}												0.507	0.587
M_{096}													0.568

TAB. 8.5: Performances moyennes des différentes combinaisons *a posteriori* des 11 structures précédentes ainsi que des structures GR4J et M_{043} (résultats en $C_{MM}(\ln Q)$ moyen).

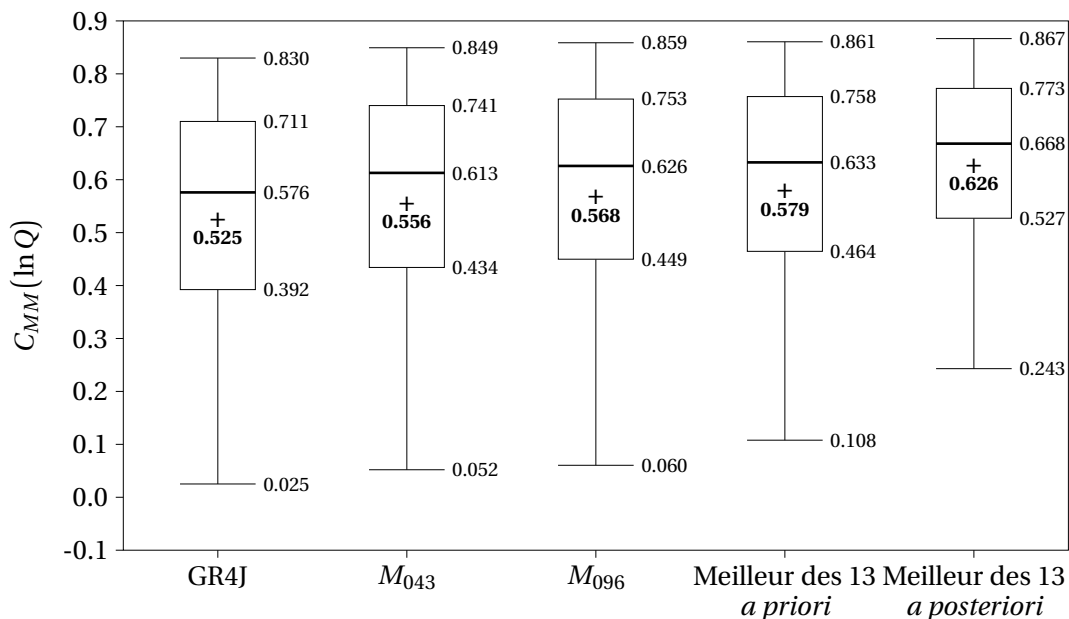


FIG. 8.12: Performances comparées des structures GR4J (quatre paramètres), M_{043} (cinq paramètres), M_{096} (sept paramètres), et des multi-modèles « totaux » : meilleure des 13 structures, sélectionnée soit en calage (*a priori*) soit en contrôle (*a priori*).

8.5 Résultats complémentaires sur les bilans simulés

Dans cette section, nous présentons quelques résultats relatifs aux bilans simulés par les modèles :

- tout d'abord, en donnant la distribution complémentaire d'un critère de bilan, pénalisant spécifiquement des écarts entre le volume total d'écoulement simulé par le modèle, et le volume réellement écoulé,
- ensuite, en analysant les flux échangés (importations et exportations) sous les différentes hypothèses testées.

8.5.1 Critères de bilan

En complément de critères mesurant l'adéquation des séries temporelles observées et simulées, Perrin (2000) propose une formulation de critère de bilan (volume total écoulé sur la période) qui permet de pénaliser de la même façon un biais de a ou de $\frac{1}{a}$:

$$C_B = 1 - \left| \left(\frac{\sum_t \hat{q}_t(\boldsymbol{\theta})}{\sum_t q_t} \right)^k - \left(\frac{\sum_t q_t}{\sum_t \hat{q}_t(\boldsymbol{\theta})} \right)^k \right| \quad (8.3)$$

où q_t et $\hat{q}_t(\boldsymbol{\theta})$ désignent respectivement le débit observé et le débit simulé au pas de temps t . En suivant ses recommandations nous utilisons ce critère avec $k = \frac{1}{2}$. La Figure 8.13 montre les distributions obtenues pour toutes les structures testées dans ce chapitre, classées par

critère moyen décroissant. Nous constatons que les structures M_{043} et M_{096} se classent en milieu de peloton, sans que les écarts soient toutefois très importants (mis à part pour les quatre plus mauvaises structures) : le critère moyen est de 0.909 pour le modèle GR4J, et de 0.913 pour les deux structures M_{043} et M_{096} .

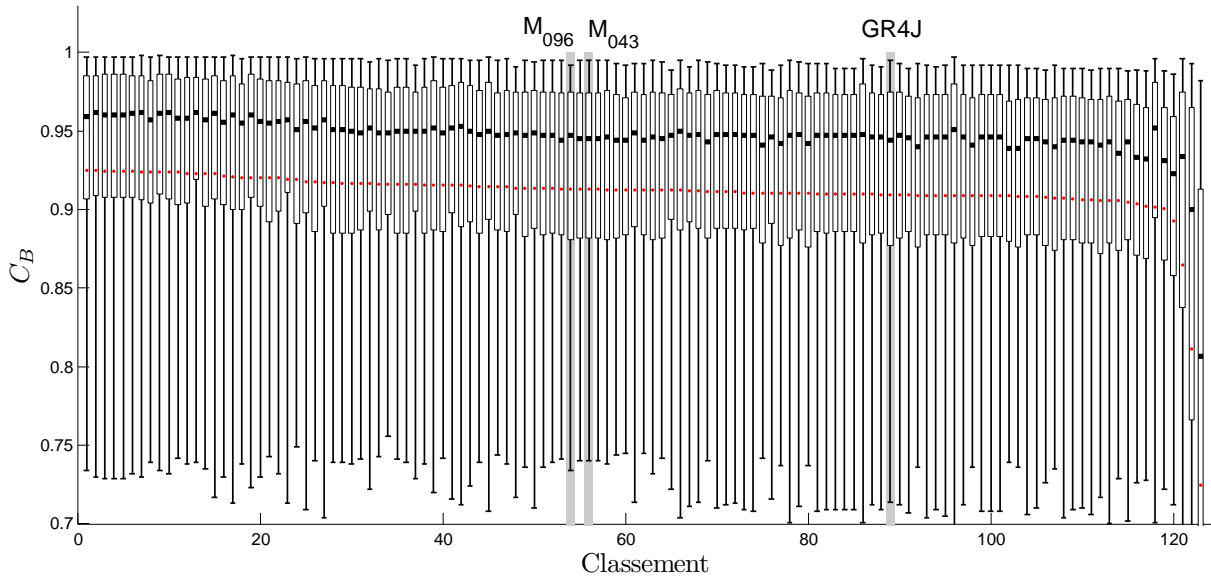


FIG. 8.13: Distributions du critère de bilan en contrôle pour les structures journalières testées dans ce chapitre.

8.5.2 Statistiques sur les flux d'échange

Sur la Figure 8.14 nous avons tracé les distributions des flux d'échange (moyennes interannuelles) simulés par les trois structures GR4J, M_{043} et M_{096} . Nous observons des distributions très proches pour les modèles GR4J et M_{043} , bien qu'il ne faille véritablement considérer que les valeurs supérieures à quelques dizaines de millimètres. La structure M_{096} simule globalement des pertes plus importantes, peut-être parce la présence du stockage intermédiaire pénalise moins une soustraction d'eau en période de crue que dans les deux autres structures. Une autre explication serait que ces pertes apparemment plus importantes correspondent en fait à une variation (positive) du stock des réservoirs. Les structures M_{043} et M_{096} autorisant le terme d'échange G_t à être de signe différent au cours de l'année, nous pouvons comparer $\sum_t |G_t|$ la somme des flux échangés (comptés positivement quel que soit le sens) pour ces deux structures. Le résultat de cette comparaison est montré sur la Figure 8.15. Nous constatons que l'ajout du réservoir supplémentaire ne diminue pas la somme des flux échangés, puisque l'on observe un écart d'une trentaine de millimètres annuellement, en moyenne sur l'ensemble des bassins.

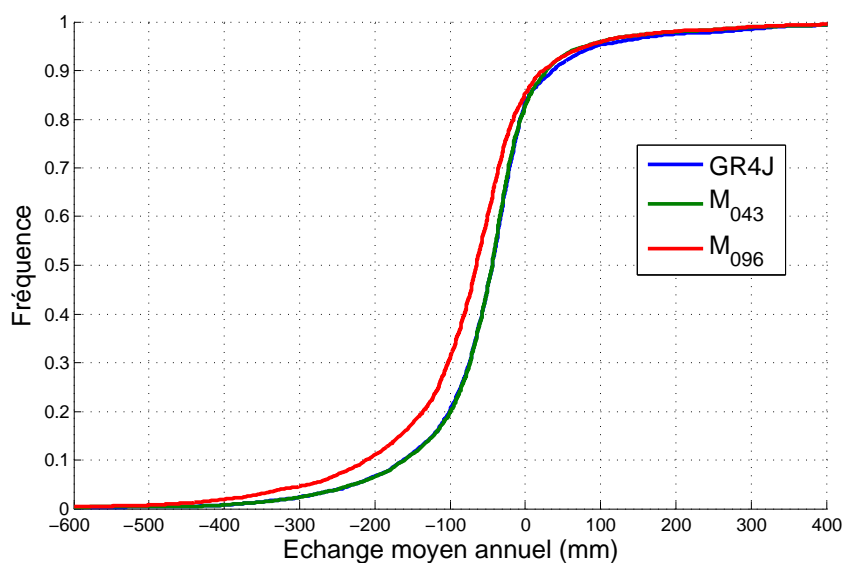


FIG. 8.14: Distributions des flux échangés (en moyenne annuelle) pour les 1389 bassins versants français et australiens.

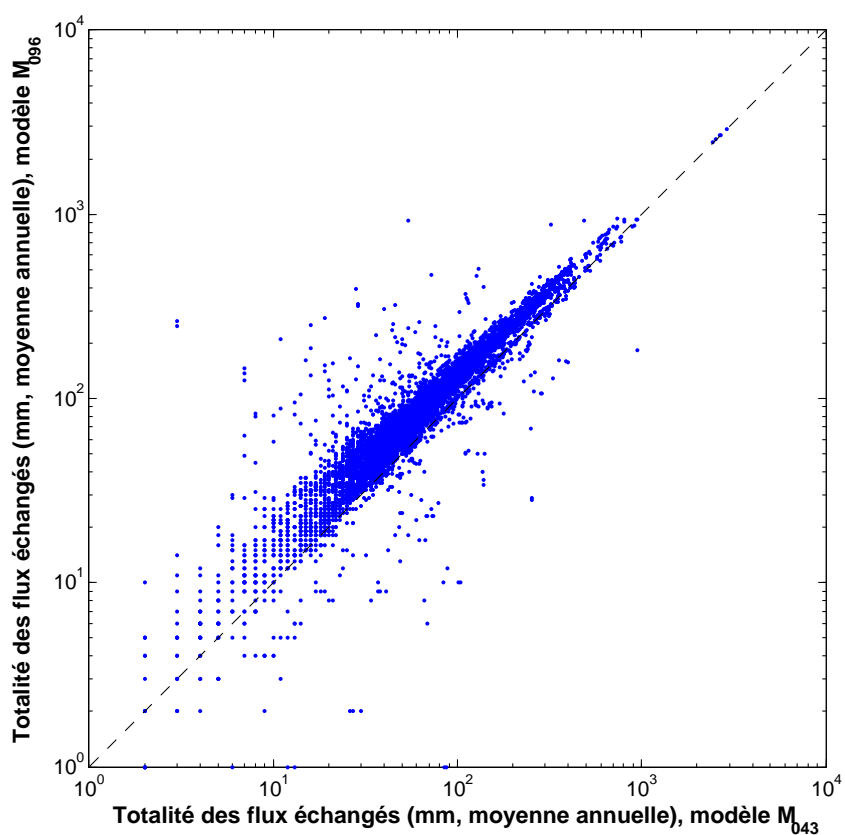


FIG. 8.15: Comparaison du total des flux échangés en valeur absolue (moyenne annuelle) pour les 1389 bassins versants français et australiens, avec les structures M₀₄₃ et M₀₉₆.

8.6 Un test extrême de robustesse

8.6.1 Les niveaux du test de *Klemeš* (1986)

Dans son article classique intitulé ‘Operational Testing of Hydrological Simulation Models’, *Klemeš* (1986) suggérait un schéma hiérarchisé pour tester des modèles, un schéma exigeant qui lierait « *la nature du test à la difficulté de la tâche de modélisation* ». Ce test comprenait quatre niveaux :

1. **Premier niveau : (*standard*) *split-sample test***

Ce test consiste à calibrer un modèle sur une période et à contrôler (*Klemeš* utilise le terme peut-être trop ambitieux de *validation*) les paramètres obtenus sur une période indépendante, afin d’évaluer la robustesse du modèle, et éviter que la procédure de calage ne fasse opter pour des paramètres qui traduiraient également des caractéristiques propre à la période de calage ;

2. **Deuxième niveau : *Proxy basin test***

Il s’agit d’imposer une difficulté supplémentaire avec un test de transférabilité des paramètres calés entre voisins : les paramètres calés sur le bassin A sont validés sur le bassin B et *vice versa* ;

3. **Troisième niveau : *Differential split-sample test***

Avec cette variante du test 1, *Klemeš* propose de tester la capacité de transposition entre deux périodes climatiquement contrastées (par exemple calage sur période humide, validation sur période sèche et *vice versa*) afin d’évaluer la capacité du modèle à faire face à des conditions climatiques non stationnaires ;

4. **Quatrième niveau : *Proxy-differential split-sample test***

Pour ce test ultime, le plus complet et le plus exigeant de tous, on transfère des paramètres à la fois entre bassins voisins et entre périodes climatiquement contrastées. On confronte ainsi le modèle à la nécessité d’extrapoler spatialement et climatiquement ;

Le premier niveau de test est aujourd’hui quasi-universellement appliqué en hydrologie (quoique l’on puisse trouver, ça et là dans la littérature hydrologique, des performances en calage). Pour notre part, nous l’avons systématiquement utilisé tout au long de cette thèse. Pour ce qui est les trois niveaux suivants, il faut bien reconnaître que rares sont les hydrologues qui s’y sont risqués, même si l’intérêt suscité par la décennie sur les bassins versants non jaugés de l’AISH (PUB) a récemment mené à un regain d’intérêt pour les *proxy basin tests* (voir par exemple *Oudin et al.*, 2008). Parmi les rares études ayant appliqué le test complet de *Klemeš*, on peut citer *Refsgaard et Knudsen* (1996) qui l’ont testé sur trois modèles (NAM, MIKE SHE et WATBAL) sur trois bassins au Zimbabwe, et *Donnelly-Makowecki et Moore* (1999) qui l’ont appliqué sur TOPMODEL et des modèles simples à réservoirs sur

deux bassins de Colombie Britannique, au Canada. A notre connaissance, le test complet de *Klemeš* n'a jamais été appliqué sur un grand nombre de bassins versants. Nous avons donc décidé de le faire pour essayer de départager qualitativement les structures GR4J et GR5J (M_{043}), et surtout pouvoir savoir si l'augmentation du nombre de paramètres ne risquait pas de réduire la robustesse du modèle.

Ce test de *Klemeš* complet, nous l'avons appliqué sans le raffiner (notamment pour le *proxy-basin test*, nous n'avons pas essayé de faire des ensembles de voisins afin d'améliorer les performances de transfert, comme cela est suggéré par *Oudin et al.*, 2008), en nous contentant de l'appliquer exactement de la même façon aux deux modèles. Le critère utilisé pour le calage est le critère $C_{MM}(\sqrt{Q})$, et nous utilisons les trois critères habituels d'évaluation en contrôle : $C_{MM}(Q)$, $C_{MM}(\sqrt{Q})$, et $C_{MM}(\ln Q)$.

Une difficulté pour appliquer le *differential test* est d'arriver à identifier des périodes contrastées suffisamment longues. Pour éviter ce problème et généraliser la procédure de *Klemeš*, nous avons choisi de caler et évaluer le modèle sur des périodes non continues, en sélectionnant les années en fonction de leurs caractéristiques climatiques. Nous avons pris la pluie annuelle comme critère de sélection, mais d'autres critères auraient pu être envisagés. Précisons que ce test a pu être réalisé grâce à l'extension de l'archive SAFRAN à la période 1970–2006, qui nous a été gracieusement fournie par Météo-France fin 2007, et à la récupération des débits journaliers correspondants par Jean-Louis Rosique.

D'un point de vue pratique, les différentes périodes ont été sélectionnées comme suit pour chaque bassin :

- a. dans la chronique disponible, nous avons identifié deux sous-périodes (P_1 et P_2) d'au moins dix ans chacune. Cette première division est destinée à constituer deux jeux de données **indépendants** ;
- b. nous avons constitué, pour chaque sous-période P_i , trois sélections de cinq années : une sélection avec pluviométrie moyenne (notée $P_{i,m}$), une autre avec pluviométrie forte (notée $P_{i,+}$), et la troisième avec pluviométrie faible (notée $P_{i,-}$). Les sélections d'années avec pluviométrie faible et forte sont strictement indépendantes, mais la sélection d'années moyennes peut recouvrir partiellement les deux autres. Ce second niveau de division est destiné à constituer des jeux de données **contrastés** ;
- c. nous avons calé le modèle sur la sélection d'années moyennes de la sous-période P_1 , puis contrôlé le modèle sur les sélections d'années moyennes, fortes et faibles de la sous-période P_2 . Nous avons fait de même en calant sur les années moyennes de P_2 , puis en faisant les trois contrôles sur P_1 . Nous avons ensuite renouvelé le test précédent en prenant pour le calage les années fortes, puis les années faibles, des sous-périodes P_1 et P_2 : on

couvre ainsi les niveaux 1 et 3 du test de *Klemeš*, et ce de manière étendue. On contrôle toujours sur un jeu de données qui est *indépendant* de celui du calage, avec un degré variable de *contraste* :

- un faible degré de contraste, lorsque l'on cale sur les années de pluviométrie moyenne d'une sous-période et que l'on contrôle sur les années de pluviométrie moyenne de l'autre sous-période ($P_{i,m} \rightarrow P_{j,m}$) ; ou que l'on cale sur les années fortes de l'une et que l'on contrôle sur les années fortes de l'autre ($P_{i,+} \rightarrow P_{j,+}$) ; ou enfin que l'on cale sur les années faibles de l'une et que l'on contrôle sur les années faibles de l'autre ($P_{i,-} \rightarrow P_{j,-}$) ;
 - un degré de contraste moyen : avec les mêmes notations, ce sont les calages-contrôle de type $P_{i,m} \rightarrow P_{j,+}$ (calage sur années moyennes et contrôle sur années fortes), $P_{i,+} \rightarrow P_{j,m}$, $P_{i,m} \rightarrow P_{j,-}$ ou encore $P_{i,-} \rightarrow P_{j,m}$;
 - un degré de contraste élevé, lorsque l'on cale sur les années de forte pluviométrie d'une sous-période et que l'on contrôle sur les années de faible pluviométrie de l'autre sous-période ($P_{i,+} \rightarrow P_{j,-}$), ou le contraire ($P_{i,-} \rightarrow P_{j,+}$)
- d. Pour les niveaux 2 et 4, on a procédé de la même façon, mais au lieu de faire un calage sur le bassin cible, on a transféré le jeu de paramètres du bassin le plus proche calé dans les mêmes conditions. Les tests les plus extrêmes sont donc les *proxy-differential split-sample tests* avec degré de contraste élevé, dans lequel le calage a lieu sur le bassin *A* et le contrôle sur le bassin *B*. En complétant nos notations, ce sont les tests $P_{i,+}^A \rightarrow P_{j,-}^B$ et $P_{i,-}^A \rightarrow P_{j,+}^B$, épreuves impitoyables dans lesquelles tous les indices changent !

Notons que l'échantillon de bassins test a été réduit, dans la mesure où nous avons posé comme condition la disponibilité d'au moins vingt années de débit (un peu plus de 600 bassins versants français satisfaisaient cette condition).

8.6.2 Résultats

Nous ne reviendrons pas sur les résultats du niveau 1 (*standard split sample test*), qui constituent la matière de ce chapitre. Commençons par les résultats présentés en Figure 8.16, qui correspondent au troisième niveau du test de *Klemeš* (*differential split-sample test*). Dans ce test, nous retrouvons les mêmes écarts entre GR4J et GR5J que ceux qui existaient dans les tests précédents : écarts très faibles (mais toujours positifs) sur les critères $C_{MM}(Q)$ et $C_{MM}(\sqrt{Q})$, écarts plus importants et stables sur les bas débits ($C_{MM}(\ln Q)$).

La Figure 8.17 présente les résultats du test niveau 2 et 4. On constate d'une part une chute brutale des performances des modèles, cohérentes avec les résultats publiés par *Oudin et al.* (2008). Par contre, on observe dans ces conditions une perte relative de l'avantage de GR5J par rapport à GR4J, voire un léger recul des performances. Ce résultat peut provenir

du fait que les échanges interbassins sont des phénomènes très locaux, qui ne peuvent être évalués que si des données sur le bassin sont disponibles. Les performances ne sont pas suffisamment différentes (en moyenne -0.2 points sur $C_{MM}(Q)$, -0.7 points sur $C_{MM}(\sqrt{Q})$, et -1.7 points sur $C_{MM}(\ln Q)$), et surtout trop faibles en valeur absolue, pour attester de la supériorité d'un modèle sur l'autre dans ce test du *proxy*. Pour mettre les modèles dans des conditions moins périlleuses, il faudrait peut-être renouveler ce test en utilisant des ensembles de voisins (plutôt qu'un seul voisin comme ici), condition de test probablement moins défavorable pour les modèles.

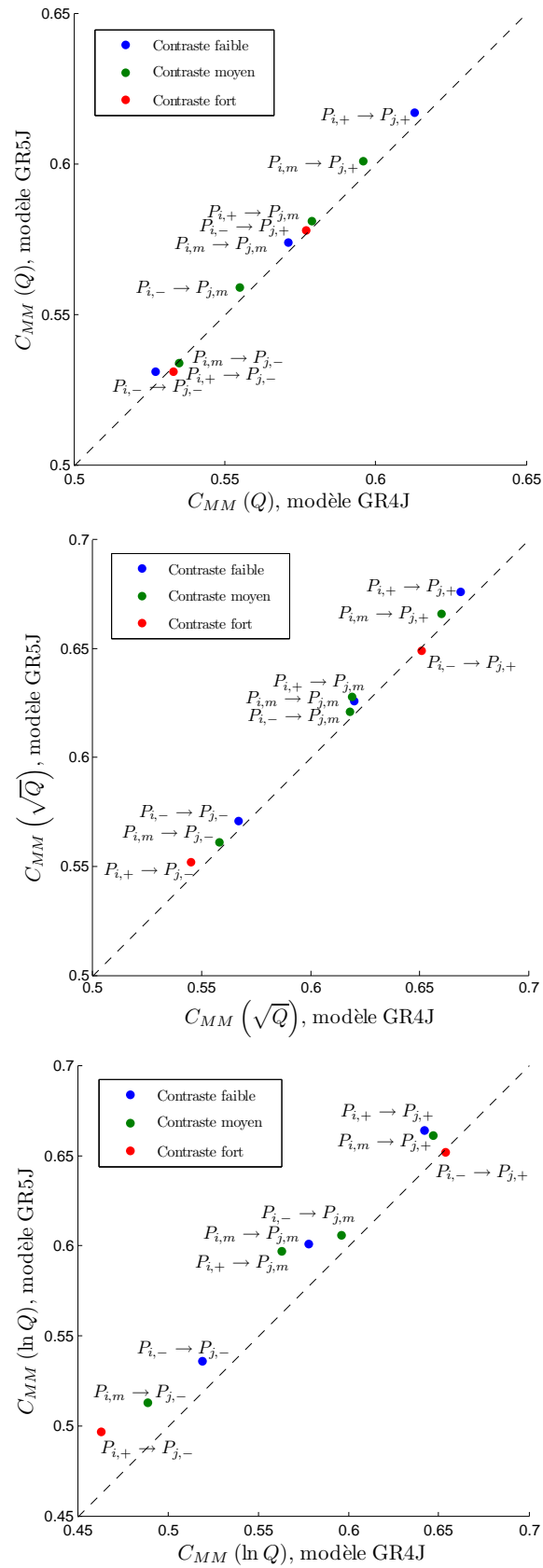


FIG. 8.16: Résultats du *differential split sample test* appliqué aux modèles GR4J et GR5J (structure M_{043}). Les coordonnées de chaque point représentent la moyenne des critères sur 612 bassins versants français.

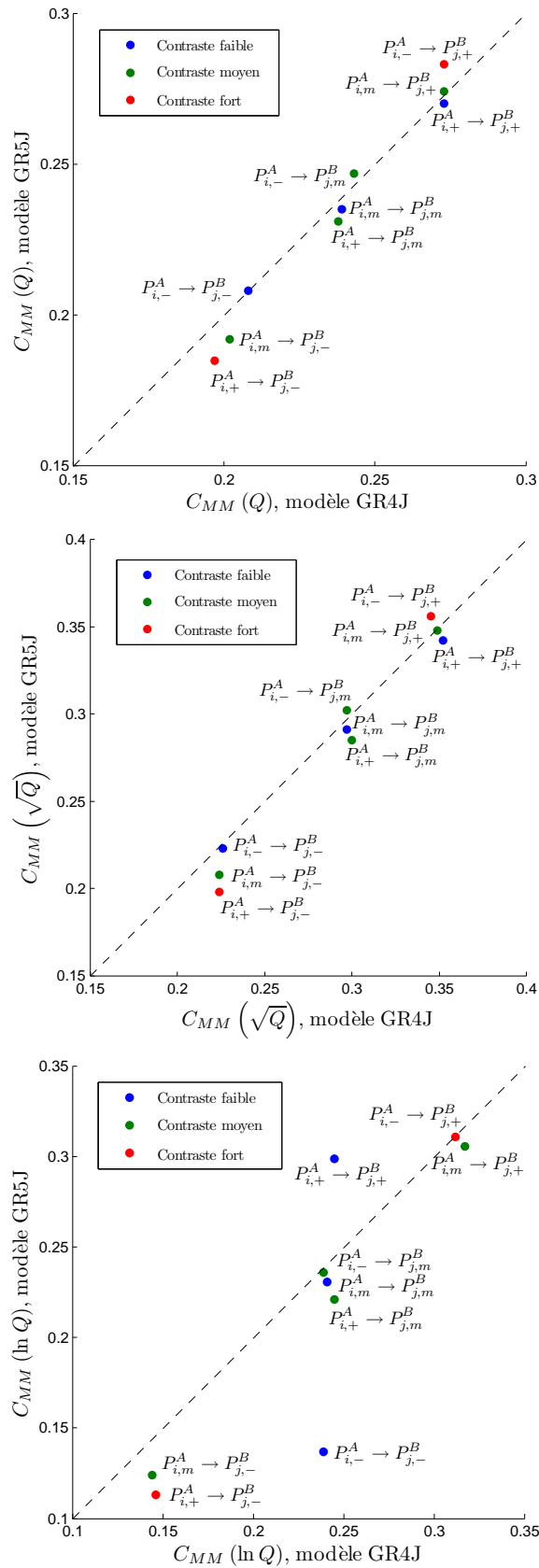


FIG. 8.17: Résultats du *proxy basin differential split sample test* appliqué aux modèles GR4J et GR5J (structure M_{043}). Les coordonnées de chaque point représentent la moyenne des critères sur 612 bassins versants français.

8.7 Synthèse

Dans ce chapitre, nous avons testé un grand nombre de formulations empiriques pour la fonction d'échange du modèle journalier GR4J. Ces tests, sans prétendre à l'exhaustivité (ce qui est impossible dans ce genre d'approche!), ont toutefois permis de confirmer certaines idées *a priori* et de dégager quelques tendances.

Tout d'abord, tous les bassins ne se sont pas avérés sensibles aux modifications que nous avons testées. Nous avons pu mettre en évidence une logique régionale à cette sensibilité, liée à la géologie (cette sensibilité resterait à vérifier pour les bassins australiens, pour lesquels nous ne disposons pas d'informations suffisantes). Ce résultat, qui va dans le même sens que ceux obtenus par l'analyse des facteurs correctifs de pluie, d'ETP et de surface au Chapitre 4, nous permet d'accréditer l'hypothèse selon laquelle la fonction d'échange joue bien le rôle qu'on lui attribue, et non le rôle exclusif de *fudge factor* (Mouelhi et al., 2005). Il s'agit là d'un élément de réalisme — lointain — que nous pouvons mettre au crédit du modèle : la fonction conceptuelle introduite par Nascimento (1995) et améliorée par Perrin (2000) semble bien *posséder une contrepartie dans le monde réel*.

Sur le plan quantitatif, les différentes solutions testées ont permis de retenir deux structures potentiellement d'intérêt :

- la structure M_{043} , à cinq paramètres, dont la fonction d'échange est contrôlée linéairement par le niveau du réservoir de routage et peut s'inverser selon que l'on est au-dessus ou en-dessous d'un taux de remplissage seuil. Ce seuil est le paramètre supplémentaire ;
- la structure M_{096} , à sept paramètres, de même formulation mais possédant un réservoir linéaire supplémentaire. La constante de temps de ce réservoir constitue le sixième paramètre, le septième étant le coefficient de partage entre le réservoir de routage et ce nouveau réservoir.

L'intérêt de ces modifications a été principalement mesuré par un critère quadratique sur le logarithme des débits ($C_{MM}(\ln Q)$). Toutefois, nous avons veillé à ce que ces améliorations de la simulation des étiages ne se traduisent pas par une diminution des performances en restitution des crues (critères $C_{MM}(\sqrt{Q})$ et surtout $C_{MM}(Q)$) : nous avons montré qu'il n'était rien, et que nous améliorions même légèrement les performances sur les forts débits.

Nous considérons que les deux paramètres supplémentaires (voire trois par rapport à GR4J) de la structure M_{096} ne présentent pas un avantage suffisant pour qu'elle soit intéressante au pas de temps journalier : nous conserverons la structure M_{043} pour ce pas de temps. Toutefois, comme nous cherchons également une structure pour le pas de temps horaire

(cf. chapitre suivant), nous garderons la structure à sept paramètres à l'esprit : en effet, la dynamique pluie-débit au pas de temps horaire est plus riche en information, ce qui pourrait éventuellement permettre d'identifier plus de paramètres, ou de faire en sorte qu'ils soient mieux définis.

9

Développements au pas de temps horaire

9.1 Point de départ

Au cours de cette thèse, nous avons d'abord traité la modélisation pluie-débit au pas de temps journalier. En abordant le pas de temps horaire, nous avons donc déjà bien « débroussaillé » le terrain, avec environ 120 versions testées et comparées.

De la même façon que nos travaux au pas de temps journalier reprenaient le modèle au point où l'avait amené *Perrin* (2000), les développements au pas de temps horaire venaient à la suite de la thèse de *Mathevet* (2005), basée sur un échantillon de 313 bassins versants français et étrangers. C'est donc sur des structures « encore chaudes » que nous avons travaillé, en appliquant une méthode assez semblable à celle suivie au pas de temps journalier, mais de moindre ampleur, essentiellement à cause de la durée des calculs sur mille bassins à un pas de temps aussi fin.

Précisons à nouveau que la disponibilité des données au pas de temps horaire était un peu plus réduite, de telle sorte que nous n'avons pu travailler « que » sur l'échantillon de 1040 bassins versants français (les travaux préliminaires pour constituer la base de données horaires ont été décrits au Chapitre 2), les données australiennes n'étant pas disponibles à ce pas de temps. L'échantillon sur lequel les tests ont porté est donc plus large que celui de *Mathevet*, tant du point de vue du nombre de bassins versants que de la durée des chroniques disponibles, mais il est malheureusement moins varié climatiquement. Il faut toutefois souligner que la fusion des données pluviographiques (horaires) et pluviométriques (journalières) a permis une bien meilleure estimation des lames d'eau de pluie horaire et que cela, ajouté à la meilleure disponibilité des données, se traduit par une bien meilleure stabilité des performances entre la phase de calage et celle de contrôle.

Concrètement, nous sommes repartis des structures sélectionnées par *Mathevet* :

- La structure GR4H, dérivée du modèle GR4J journalier,
- Une version simplifiée à six paramètres du modèle MORDOR (EDF-DTG, voir *Garçon*, 1999 ; *Paquet*, 2004 ; *Paquet et al.*, 2006), représentée sur la Figure 9.1.

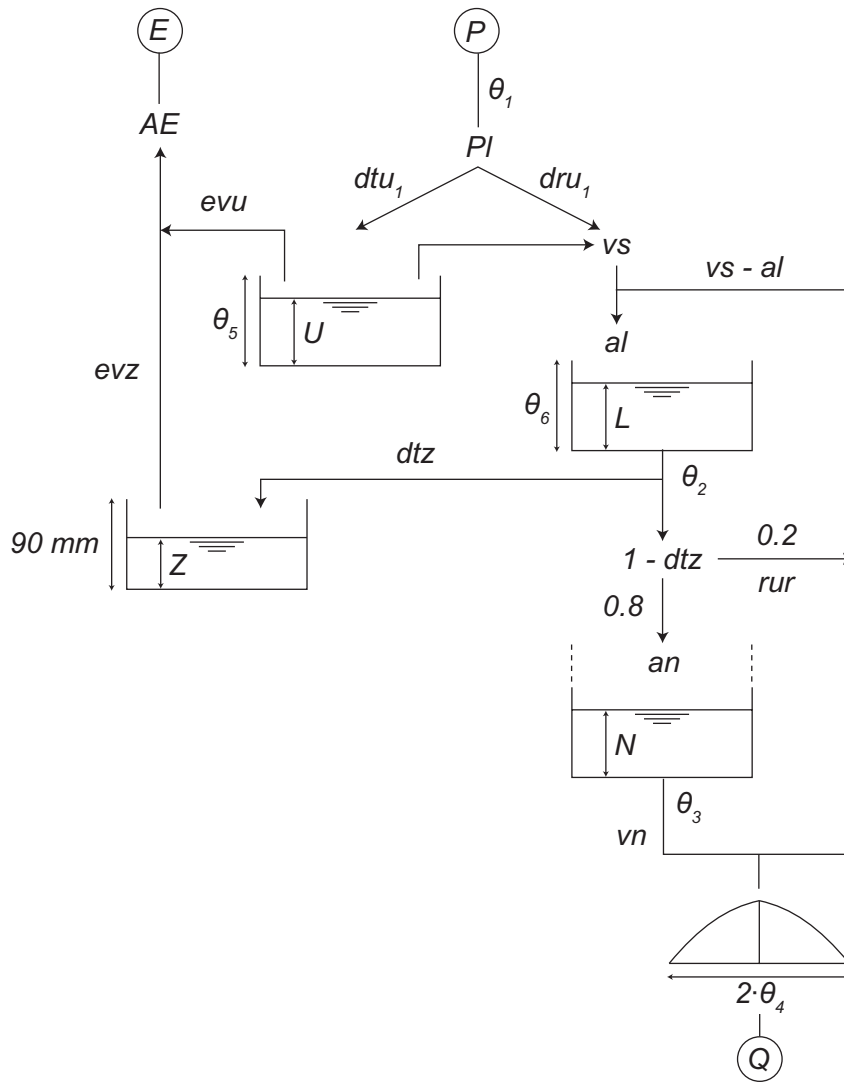


FIG. 9.1: Structure du modèle horaire MORDOR à six paramètres (Mathevet, 2005).

9.2 Test au pas de temps horaire des structures identifiées au pas de temps journalier

Dans cette section, nous commençons par essayer de « transférer » au pas de temps horaire les structures d'intérêt à cinq et sept paramètres identifiées au pas de temps journalier. Comme l'explique Mathevet (2005), le changement de pas de temps occasionne des transformations de certains paramètres, transformations que l'on peut essayer d'estimer *a priori* d'après les formes des différentes lois de vidange (cf. Annexe C) :

- pour la percolation, la capacité équivalente est $\beta\theta_1$ (où θ_1 est la capacité du réservoir de production, dont la loi de production n'est pas dépendante du pas de temps¹). La puissance de la loi de percolation étant égale à 4, on doit théoriquement multiplier θ_1 par $24^{\frac{1}{4}}\beta \approx 2.21\beta$ pour obtenir la constante de percolation. C'est le résultat trouvé par

¹elle en dépend implicitement, cf. Chapitre 10.

Mathevet, que nos premiers tests au pas de temps horaire ont confirmé pour l'échantillon de 1040 bassins français : la valeur $\beta_J = \frac{4}{9}$ du pas de temps journalier donne $\beta_H = 4.98 \approx 5$, ce qui est la valeur optimale obtenue empiriquement au pas de temps horaire ;

- la capacité θ_2 du réservoir de routage, dont la loi de vidange est également de puissance égale à 4, doit théoriquement être multipliée par 2.21 elle aussi ;
- En supposant que la fonction d'échange a une dynamique suffisamment lente pour pouvoir écrire $F_{H,t} \approx \frac{1}{24}F_{J,t}$, et que « l'encours » R_t du réservoir au temps t est comparable aux deux pas de temps, on doit avoir approximativement :

$$\begin{aligned} F_{H,t} &= \frac{1}{24}\theta_{J,4} \left(\frac{R_t}{\theta_{J,2}} - \theta_{J,5} \right)^\alpha \\ F_{H,t} &= \frac{1}{24}\theta_{J,4} \left(\frac{\theta_{H,2}}{\theta_{J,2}} \right)^\alpha \cdot \left(\frac{R_t}{\theta_{H,2}} - \frac{\theta_{J,2}}{\theta_{H,2}} \cdot \theta_{J,5} \right)^\alpha \\ F_{H,t} &= \frac{2.21^\alpha}{24}\theta_{J,4} \cdot \left(\frac{R_t}{\theta_{H,2}} - \frac{1}{2.21}\theta_{J,5} \right)^\alpha \end{aligned}$$

Cette forme inclut les formulations à quatre paramètres ($\alpha = \frac{7}{2}$ et $\theta_5 = 0$) et cinq paramètres ($\alpha = 1$, θ_5 optimisé).

- pour le routage linéaire, nous sommes repartis de la forme fixée par *Mathevet* : un hydrogramme unitaire unique en tête des deux branches, de forme puissance fixée à 1.25. Nous essaierons de vérifier si la base de temps vérifie $\theta_{H,3} \approx 24\theta_{J,3}$ (pour les $\theta_{J,3} > 0.5$).

9.2.1 Résultats en calage-contrôle

Sur les Figures 9.2 et 9.3, nous montrons les distributions de performance des structures à quatre, cinq et sept paramètres ainsi que celles de MORDOR, pour les critères $C_{MM}(Q)$, $C_{MM}(\sqrt{Q})$ et $C_{MM}(\ln Q)$: en calage sur la Figure 9.2, et en contrôle sur la Figure 9.3. Dans les structures à quatre paramètres nous avons ajouté, comme en journalier, celles utilisant un coefficient correctif plutôt que la fonction d'échange. Tout au long de ce chapitre, nous présentons des résultats obtenus en calant les modèles à l'aide d'une *fonction objectif sur le débit Q non transformé*. Comme nous le verrons au Chapitre 11, ce choix n'était rétrospectivement pas le meilleur : une fonction objectif sur \sqrt{Q} , comme au pas de temps journalier, donne à la fois des performances plus élevées et des écarts plus grands entre structures, sans toutefois changer le classement des modèles.

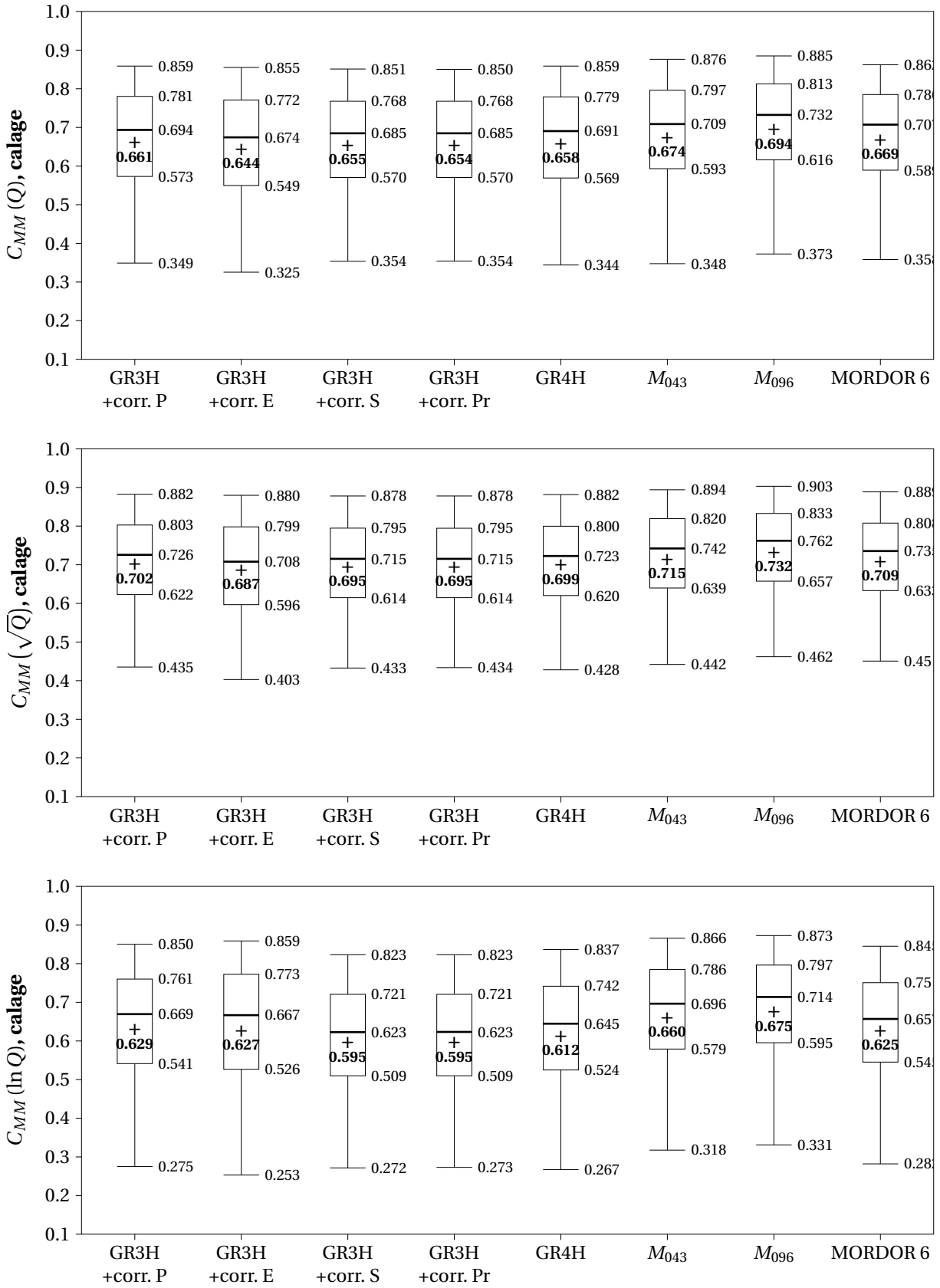


FIG. 9.2: Performances en **calage**, au pas de temps horaire, de huit premières structures : GR4H, MORDOR version six paramètres (*Mathevet, 2005*), ainsi que les structures M_{043} et M_{096} sélectionnées au pas de temps journalier. Nous ajoutons également les structures de type GR sans fonction d'échange mais avec un coefficient correctif (soit de la pluie, soit de l'ETP, soit de la surface).

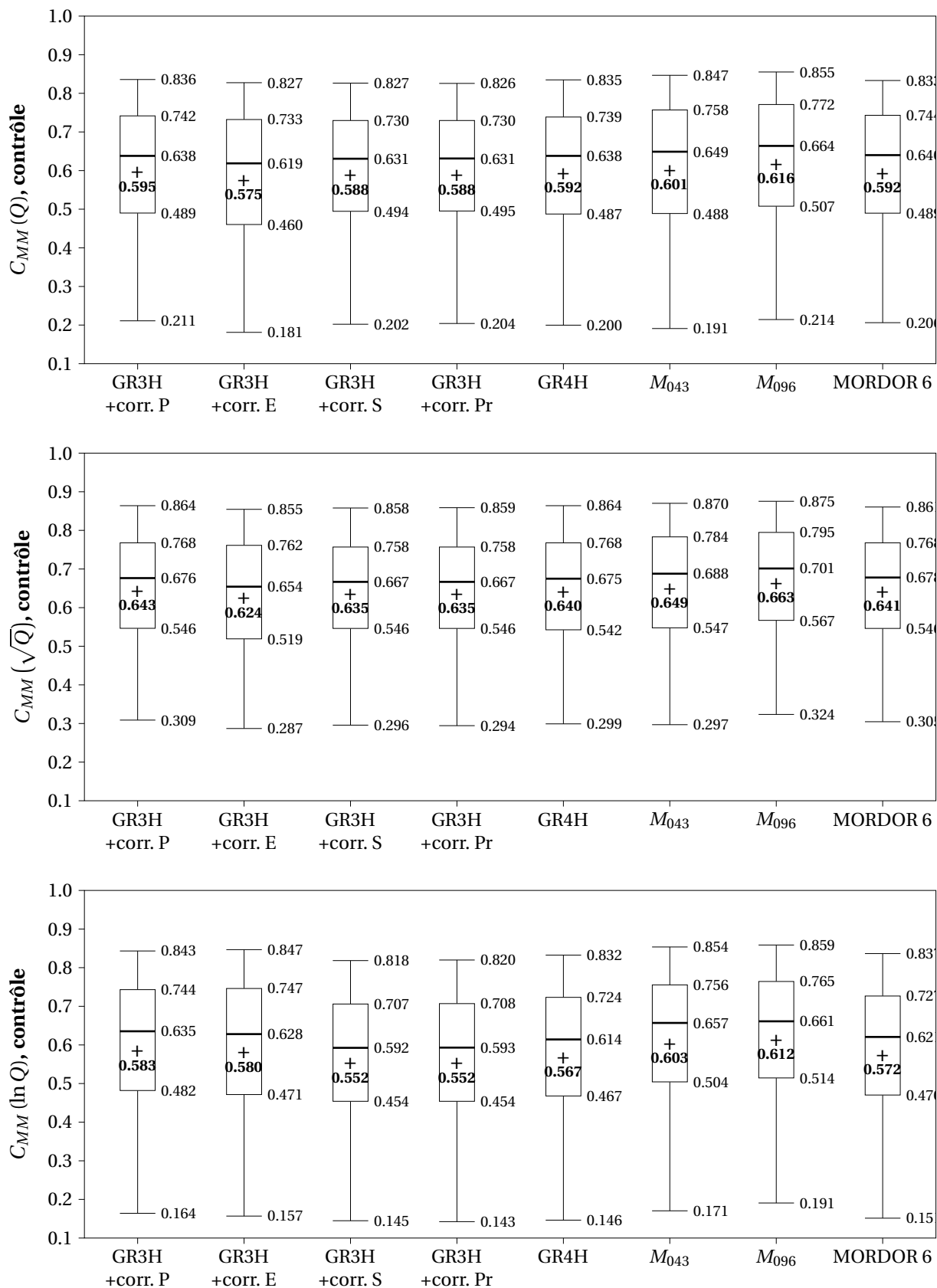


FIG. 9.3: Performances en **contrôle**, au pas de temps horaire, de huit premières structures : GR4H, MORDOR version six paramètres (Mathevet, 2005), ainsi que les structures M_{043} et M_{096} sélectionnées au pas de temps journalier. Nous ajoutons également les structures de type GR sans fonction d'échange mais avec un coefficient correctif (soit de la pluie, soit de l'ETP, soit de la surface, soit de la production).

Quoiqu'il en soit, les distributions de critères obtenues confirment les résultats obtenus au pas de temps journalier. Nous retrouvons notamment les mêmes ordres de grandeur pour l'écart entre GR4H et la structure à cinq paramètres M_{043} :

- des différences peu significatives mais positives pour les critères insistant sur les crues, avec +0.8 point sur la moyenne des $C_{MM}(Q)$ et +0.9 point sur la moyenne des $C_{MM}(\sqrt{Q})$,
- des différences plus significatives en étiage, avec +3.6 points sur le $C_{MM}(\ln Q)$.

Notons que les performances moyennes pour les structures GR4H et MORDOR 6 sont nettement supérieures à celles obtenues par *Mathevet* : de l'ordre de 0.6 à 0.65 pour les critères $C_{MM}(Q)$ et $C_{MM}(\sqrt{Q})$ en contrôle, contre 0.35 à 0.45 précédemment. Les chutes de performances observées entre calage et contrôle sont également bien moins grandes : elles étaient de 20 à 30 points en moyenne, contre seulement 5 à 8 points ici. Comme nous l'avons expliqué, ce résultat n'a rien de surprenant et tient à la disponibilité des données. Les chroniques utilisées sont nettement plus longues, et l'information pluviométrique est globalement de bien meilleure qualité, par le croisement des données pluviographiques (horaires) et pluviométriques (journalières). Contrairement à ce que nous avons trouvé au pas de temps journalier, l'utilisation d'un coefficient correctif de la pluie reste statistiquement une solution performante, puisqu'elle est au niveau de GR4H (et même meilleure sur les logarithmes). Cependant, nous retrouvons le même résultat que celui obtenu en journalier :

- les distributions des coefficients correctifs ne sont pas centrées en 1, comme pourrait le laisser présager une correction d'erreurs. Leur moyenne se situe entre 0.83 et 0.85, ce qui signifierait une *sur-estimation* systématique de la pluie de 15% (cf. Figure 9.4). Ce résultat est contre-intuitif, les facteurs invoqués pour expliquer l'incertitude dans la mesure des précipitations jouant quasiment tous dans le sens d'une *sous-estimation*, qu'ils concernent les instruments individuels (déficit de captation) ou les réseaux (faible densité sur les reliefs),
- les coefficients correctifs de la pluie pour les modèles GR3H+Corr.P et MORDOR présentent la même logique régionale que celle constatée en journalier (Figure 9.5).

D'autre part, étant donné leur faible nombre, nous pouvons examiner les bassins pour lesquels le coefficient correctif de pluie est supérieur à un seuil donné, par exemple 1.25 (sous-estimation théorique de la pluie supérieure à 25%), pour l'une ou l'autre de deux structures utilisant un tel coefficient. Le tableau 9.1 donne la liste de ces 22 bassins. Plus de la moitié d'entre eux sont alimentés par une ou plusieurs sources karstiques importantes, seule l'autre moitié pouvant effectivement correspondre à des problèmes d'estimation des précipitations, notamment nivales (Allier, Tarn, bassins corses, etc.).

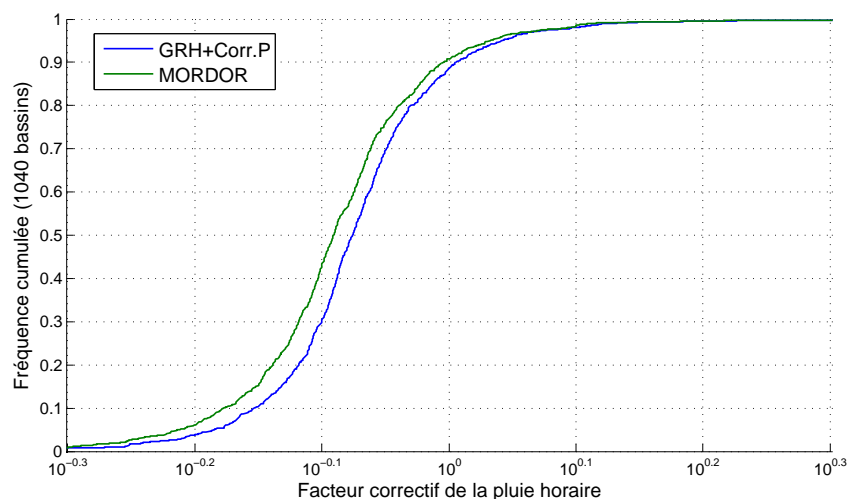


FIG. 9.4: Distributions des coefficients correctifs de pluie, pour les deux structures utilisant un tel paramètre.

CODE	STATION	SURFACE (km ²)	Corr. P, GRH	Corr. P MORDOR
U1115020	La Bèze à Bèze	33	5.05	5.05
H1333010	La Laine à Soulaines-Dhuys	24	2.40	2.38
V1425010	Le Groin à Artemare	42	2.04	1.98
U2604030	La Loue à Vuillafans	326	1.68	1.65
O1484320	La Touyre à Montferrier	29	1.56	1.56
K2010820	L'Allier à Laveyrune	49	1.53	1.52
O3064010	Le Tarnon à Florac	124	1.42	1.43
P0115010	La Burande à la Tour-d'Auvergne	20	1.41	1.38
V5425210	La Gagnière à Banne	55	1.38	1.31
A5622010	La Bouvade à Bicqueley	71	1.32	1.31
U2002010	Le Doubs à Mouthe	30	1.36	1.24
Y0115410	La Massane à Argelès-sur-Mer	17	1.33	1.26
O1484310	La Touyre à Lavelanet	59	1.29	1.29
O3314010	La Dourbie à Dourbies	43	1.32	1.26
P0885020	Le Mars au Falgoux	21	1.31	1.26
O0234020	Le Ger à Aspet	95	1.28	1.28
U2616410	Le Verneau à Nans-sous-Sainte-Anne	11	1.29	1.25
O3011010	Le Tarn au Pont-de-Montvert	67	1.27	1.27
U2655030	La Cuisance à Mesnay	70	1.28	1.26
O0362510	Le Salat à Soueix-Rogalle	379	1.25	1.27
Y9414020	Le Fium-Orbo à Ghisoni	115	1.26	1.23
Y9025010	Le Vecchio à Venaco	147	1.14	1.27

TAB. 9.1: Liste des bassins dont la correction de pluie dépasse 1.25 pour l'un ou l'autre des modèles GR3H+Corr.P et MORDOR. Les bassins versants indiqués en gras sont ceux pour lesquels une alimentation karstique est avérée : chacun d'eux est souvent le membre « gagnant » d'un doublet de bassins : systèmes Tille–(Venelle)–Bèze, Voire–Laine, Sérans–Groin, Doubs–Loue, Aroffe–Bouvade, etc.

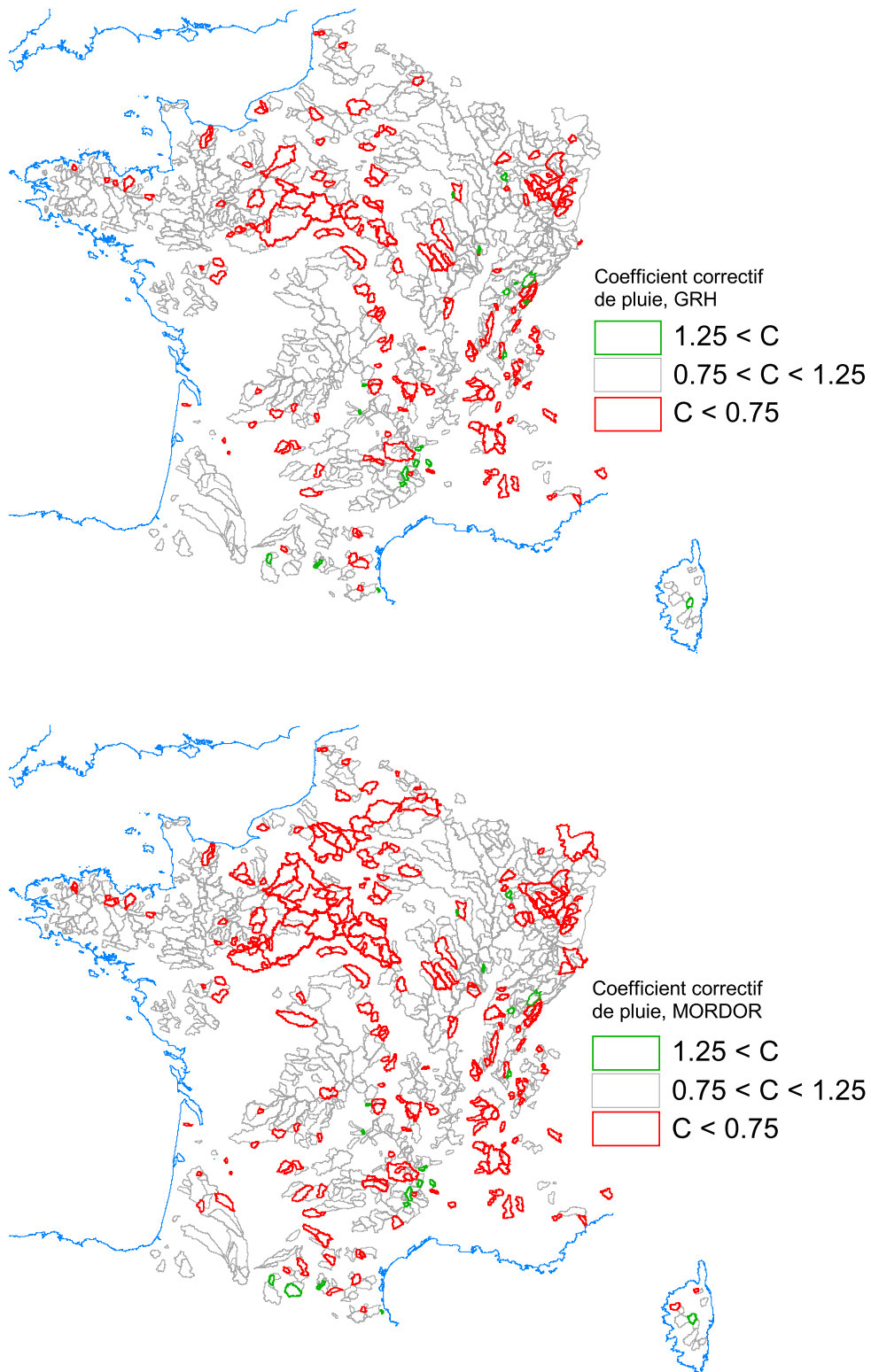


FIG. 9.5: Cartographie des coefficients correctifs de pluie, pour les deux structures utilisant un tel paramètre.

9.2.2 Relations entre les paramètres des structures aux pas de temps journalier et horaire

Les Figures 9.6, 9.7 et 9.8 montrent les relations entre les paramètres optimisés des structures GR4, M_{043} et M_{096} au pas de temps horaire, et les valeurs « théoriques » données par le changement de pas de temps d'intégration des équations d'évolution. Les paramètres du modèle journalier sont obtenus en calant sur les débits journaliers *avec comme entrée les pluies journalières*, le modèle horaire étant bien sûr calé sur les débits horaires et en utilisant des *pluies horaires* : il s'agit donc bien d'un transfert totalement *a priori* des paramètres, sans utilisation d'aucune donnée horaire en calage.

Nous constatons une assez bonne cohérence des paramètres obtenus aux deux pas de temps pour les structures à quatre et cinq paramètres, hormis pour le paramètre θ_3 (base de temps de l'hydrogramme unitaire) qui devient logiquement plus sensible au pas de temps horaire pour des valeurs inférieures à 24 h (au pas de temps journalier, toute valeur de θ_3 inférieure à 0.5 jour donne le même hydrogramme unitaire, celui-ci s'étalant entre 0 et $2 \cdot \theta_3$).

Les paramètres de la structure à sept paramètres (M_{096}) semblent en revanche moins cohérents, avec notamment des interactions probables entre les deux composants linéaires de routage (le réservoir L et l'hydrogramme unitaire) lorsque la constante de temps θ_7 du réservoir L est du même ordre de grandeur que la base de temps θ_3 .

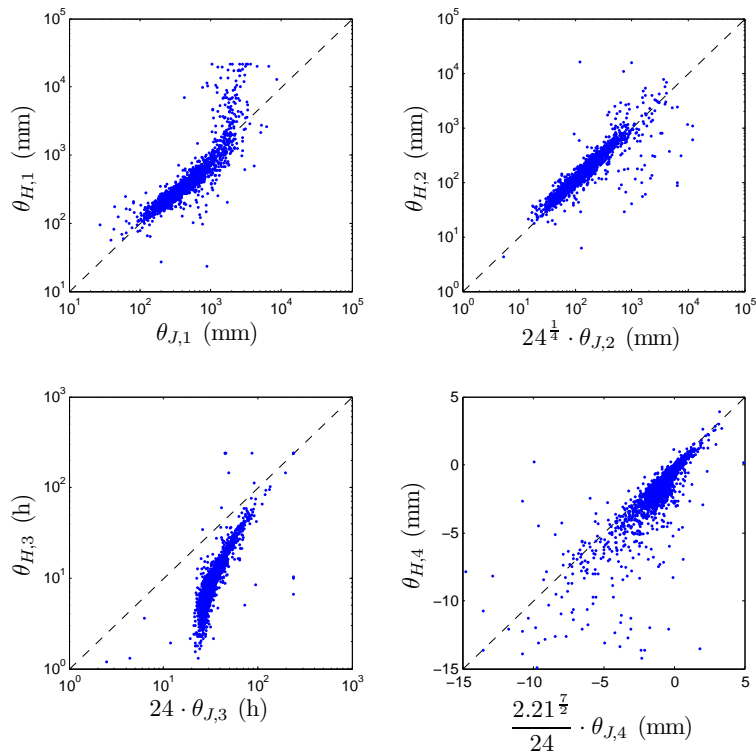


FIG. 9.6: Relation entre les paramètres optimisés du modèle horaire GR4H ($\theta_{H,i}$, en ordonnées) et les valeurs théoriques déduites des valeurs journalières ($f(\theta_{J,i})$, en abscisses).

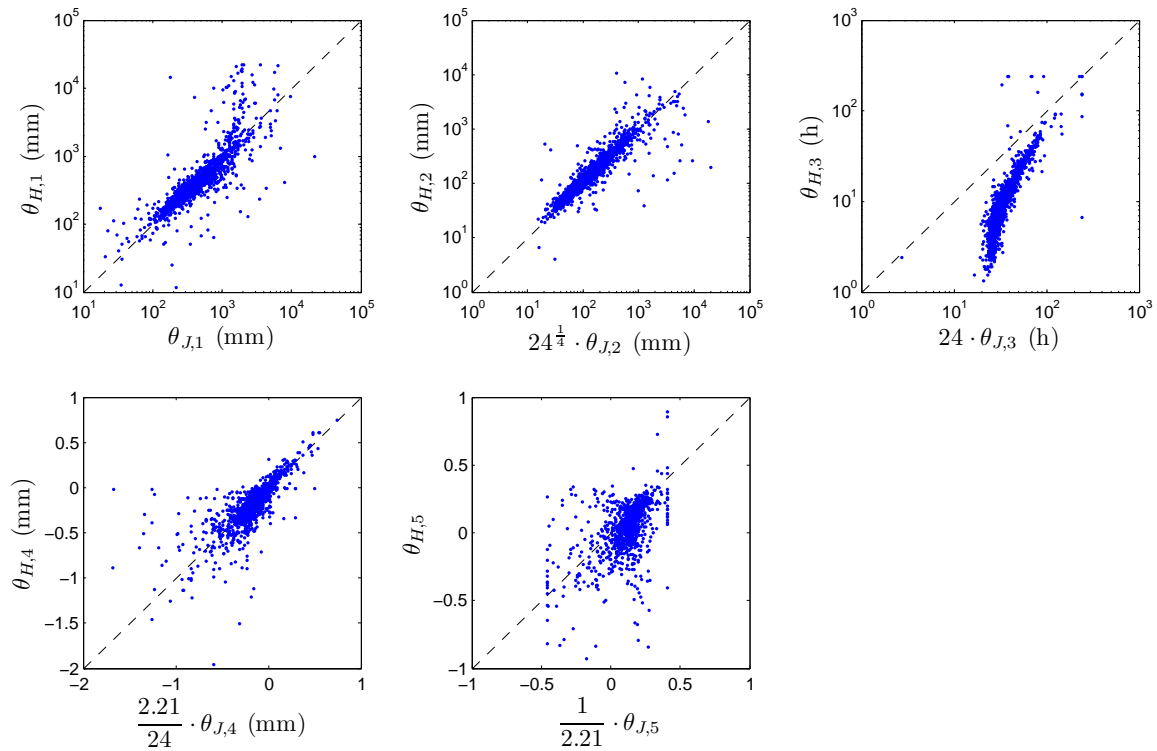


FIG. 9.7: Relation entre les paramètres optimisés de la structure M_{043} en horaire ($\theta_{H,i}$, en ordonnées) et les valeurs théoriques déduites des valeurs journalières ($f(\theta_{J,i})$, en abscisses).

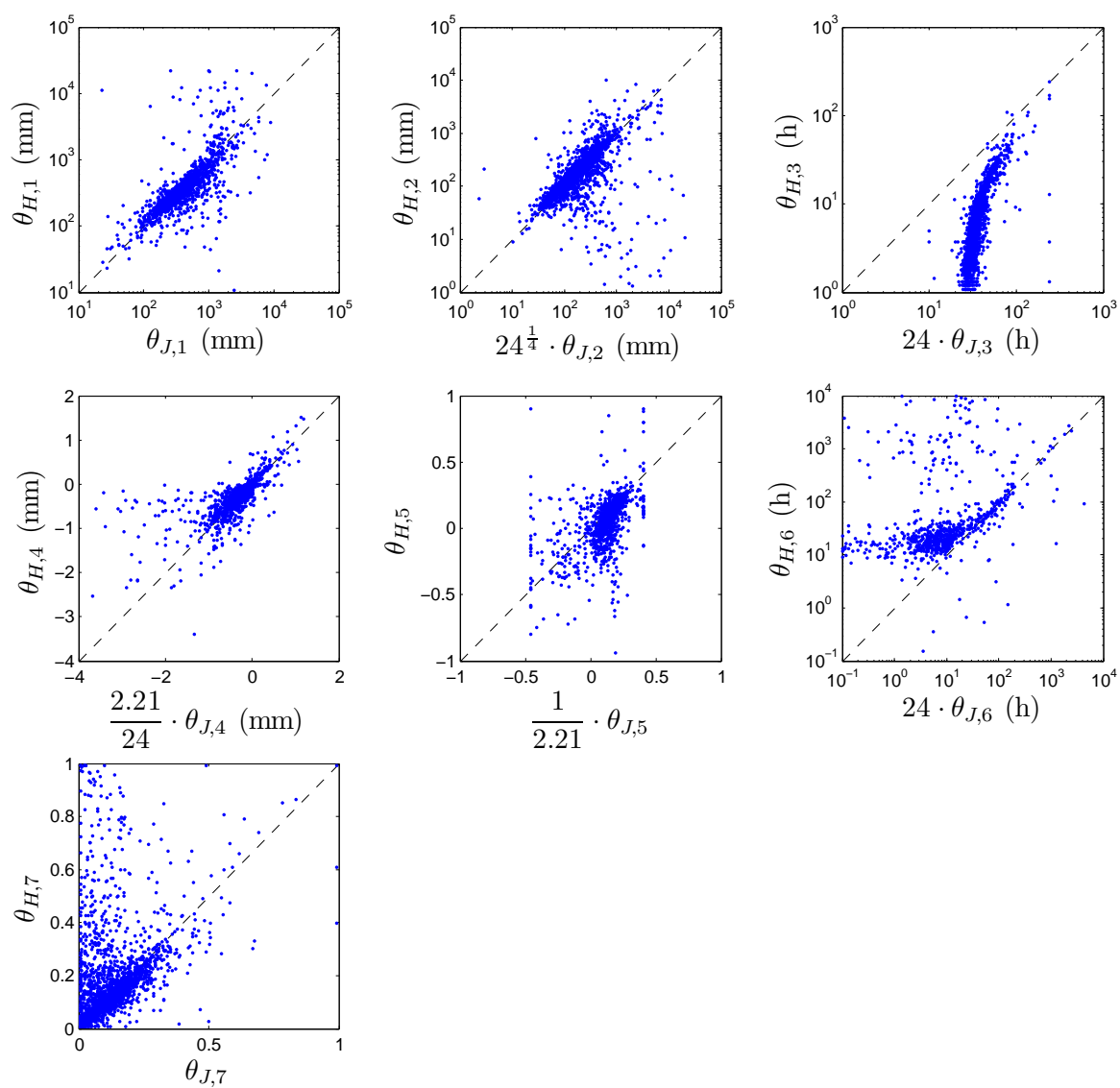


FIG. 9.8: Relation entre les paramètres optimisés de la structure M_{096} en horaire ($\theta_{H,i}$, en ordonnées) et les valeurs théoriques déduites des valeurs journalières ($f(\theta_{J,i})$, en abscisses).

9.3 Recherche d'améliorations spécifiques au pas de temps horaire

La transposition des résultats du pas de temps journalier nous a donc permis d'améliorer les performances du modèle GR4H initial. Cependant, les structures correspondantes (notamment la structure à cinq paramètres M_{043}) peuvent peut-être encore être améliorées.

9.3.1 Vérification de l'optimalité locale des structures M_{043}

Dans cette section, nous présentons les résultats de tests consistant à faire varier certains paramètres fixes des structures M_{043} et M_{096} , de façon à vérifier qu'elles sont bien localement optimales. Par « localement optimales », nous voulons signifier que nous ne recommençons pas à tester *de façon croisée* une *large* gamme d'exposants de la fonctions d'échange, de points d'ajout, de coefficient de percolation, etc. Au contraire nous faisons varier ces paramètres indépendamment les uns des autres autour de la valeur initiale pour vérifier que la structure optimale identifiée au pas de temps journalier l'est bien aussi au pas de temps horaire (l'expérience nous apprend qu'il y a de fortes chances pour que ce soit le cas !).

- variation des points d'ajout et du réservoir contrôlant l'échange : nous testons les formes linéaires $F_t = \theta_4 \left(\frac{S_t}{\theta_1} - \theta_5 \right)$ (contrôle par le réservoir de production) et $F_t = \theta_4 \left(\frac{R_t}{\theta_2} - \theta_5 \right)$ (contrôle par le réservoir de routage) croisées avec les solutions d'ajout (S, S), (S, R), (R, S) ou (R, R) en fonction du signe de F_t (huit variantes en incluant la structure M_{043}),
- variation de l'exposant de l'échange : $F_t = \theta_4 \left(\frac{R_t}{\theta_2} - \theta_5 \right)^\alpha$,
- essais d'optimisation de la constante de percolation, pour trois valeurs d'exposant (3, 4 et 5), soit trois variantes,
- variations du routage linéaire : (1) tests d'hydrogrammes unitaires, qui ont été beaucoup plus développés qu'au chapitre précédent étant donné la finesse du pas de temps, (2) modification du coefficient de partage entre les deux branches.

La question de l'hydrogramme unitaire optimal pour le pas de temps horaire n'était pas spécifique à notre sujet, mais comme les essais de *Mathevet* n'avaient pas montré une grande sensibilité à la forme de cet hydrogramme, nous avons repris ses tests sur notre échantillon élargi. Dans ce chapitre nous présentons les tests de formes paramétriques, mais on trouvera également en Annexe D une méthode d'identification par déconvolution adaptée de *Croke* (2005). Cette méthode, très raffinée sur le plan mathématique, ne permet malheureusement pas de dépasser les performances du simple hydrogramme de forme puissance...

Variation du point d'ajout et du réservoir contrôlant l'échange

Le Tableau 9.2 montre les performances obtenues en croisant deux formulations de la fonction d'échange (fonction affine du taux de remplissage du réservoir de routage, ou du réservoir de production) avec quatre solution d'ajout / soustraction (toujours dans R quel que soit le signe de F_t , toujours dans S , alternativement dans S ou R selon que F_t est positif ou négatif).

	$F_t = \theta_4 \left(\frac{R_t}{\theta_2} - \theta_5 \right)$				$F_t = \theta_4 \left(\frac{S_t}{\theta_1} - \theta_5 \right)$			
	(S, S)	(S, R)	(R, S)	(R, R)	(S, S)	(S, R)	(R, S)	(R, R)
$C_{MM}(Q)$ moyen	0.588	0.599	0.589	0.601	0.588	0.582	0.587	0.592
$C_{MM}(\sqrt{Q})$ moyen	0.636	0.648	0.637	0.649	0.637	0.633	0.636	0.635
$C_{MM}(\ln Q)$ moyen	0.572	0.593	0.576	0.603	0.572	0.560	0.571	0.558

TAB. 9.2: Performances moyennes de différentes formulations et différents points d'ajout du terme d'échange dans le modèle horaire.

Nous retrouvons sensiblement la même logique qu'au pas de temps journalier, la meilleure solution restant un contrôle par le niveau du réservoir de routage avec ajout et retrait dans ce réservoir quelque soit le signe de F_t .

Essais d'optimisation de la constante de percolation

Dans cet essai, nous vérifions que la modification de la fonction d'échange ne change pas la forme optimale de la fonction de percolation du réservoir de production (capacité au pas de temps égale à $\beta\theta_1$, où $\beta \simeq 5$ dans la version GR4H). Nous avons donc testé trois versions où la constante est optimisée ($\beta = \theta_6$), avec un exposant p de la percolation égal à 3, 4 ou 5.

	M_{043}	β optimisé		
		$p = 3$	$p = 4$	$p = 5$
$C_{MM}(Q)$ moyen	0.601	0.600	0.600	0.598
$C_{MM}(\sqrt{Q})$ moyen	0.649	0.647	0.645	0.643
$C_{MM}(\ln Q)$ moyen	0.603	0.599	0.596	0.593

TAB. 9.3: Performances en contrôle des structures où la constante de percolation fait partie des paramètres optimisés. On ne constate pas d'amélioration comparativement à la structure M_{043} .

Comme au pas de temps journalier, la modification de la fonction d'échange ne semble pas nécessiter d'adaptation de la fonction de percolation (i.e., de vidange du réservoir sol). Ces changements s'avérant peu sensibles, et la médiane du paramètre θ_6 étant proche de la valeur de GR4H (4.61 contre 4.98, cf. Figure 9.9), nous avons choisi de laisser cette fonction en l'état.

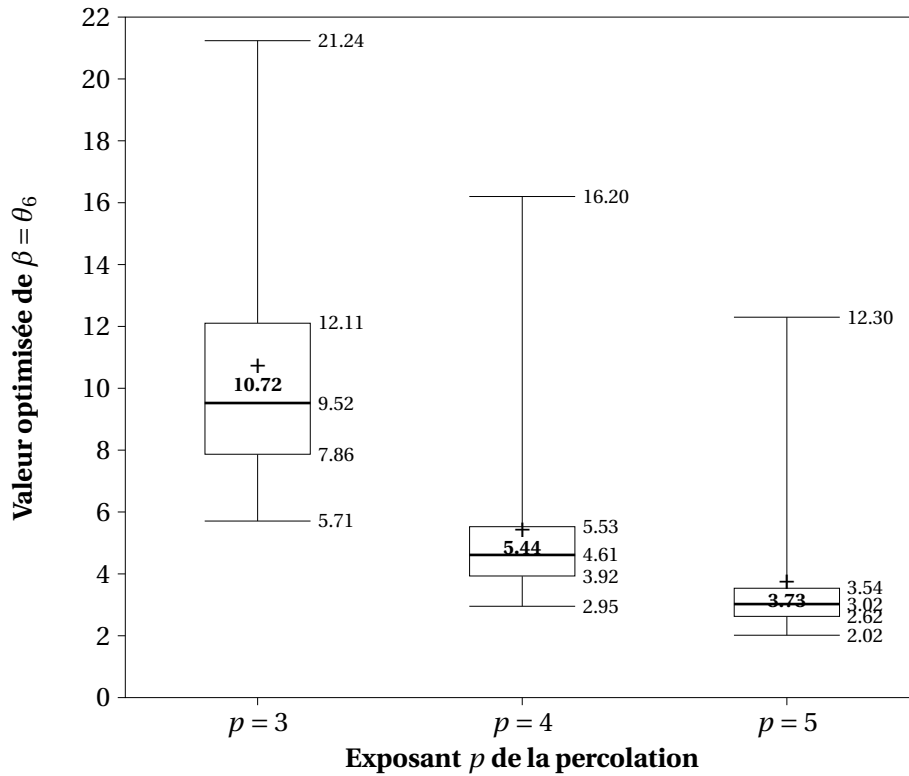


FIG. 9.9: Distribution du paramètre θ_6 laissé libre tel que la capacité au pas de temps de la percolation soit égale à $C_{\text{perc}} = \theta_1 \cdot \theta_6$. Résultats pour trois valeurs de l'exposant de la percolation ($p = 3, 4$ ou 5).

Tests d'hydrogrammes unitaires (formes paramétriques)

Le Tableau 9.4 résume les formes paramétriques testées, ainsi que leurs performances moyennes. Nous pouvons regrouper ces formes en deux catégories :

- des hydrogrammes symétriques, dont le maximum est localisé au paramètre d'échelle T (base de temps),
- des hydrogrammes asymétriques caractérisés par un paramètre d'échelle T et un paramètre de forme.

À chaque fois, nous donnons l'équation de l'hydrogramme unitaire $h(t)$ et celle de son intégrale $H(t)$, qui est utilisée pour calculer les ordonnées de l'hydrogramme discrétisé au pas de temps Δt , comme illustré sur la Figure 9.10. Dans le cas d'hydrogrammes symétriques (entre 0 et $2T$), nous ne donnons que l'expression entre 0 et T .

n°	Paramètres	Forme	Symétrie	$C_{MM}(Q)$ moyen, contrôle
#0	—	Pas d'hydrogramme unitaire (délai nul)	—	0.427
# 1	T	Puissance $\alpha = 2.5$ (idem journalier) $h(t \leq T) = \frac{\alpha}{2T} \left(\frac{t}{T}\right)^{\alpha-1}$ $H(t \leq T) = \frac{1}{2} \left(\frac{t}{T}\right)^\alpha$	oui	0.600
# 2	T	Puissance $\alpha = 2$	oui	0.602
# 3	T	Puissance $\alpha = 3$	oui	0.598
# 4	T	Puissance $\alpha = 1.5$	oui	0.603
# 5	T	Puissance $\alpha = 1.25$	oui	0.603
# 6	T	Sinusoïdal $h(t \leq T) = \frac{1}{2T} \left(1 - \cos\left(\pi \frac{t}{T}\right)\right)$ $H(t \leq T) = \frac{t}{2T} - \frac{1}{\pi} \sin\left(\pi \frac{t}{T}\right)$	oui	0.601
# 7	T	Parabolique $h(t \leq T) = \frac{3}{4T} \frac{t}{T} \left(2 - \frac{t}{T}\right)$ $H(t \leq T) = \frac{3}{4} \left(\frac{t}{T}\right)^2 - \frac{1}{2} \left(\frac{t}{T}\right)^3$	oui	0.603
# 8	T	Elliptique $h(t \leq T) = \frac{2}{\pi T} \sqrt{\frac{t}{T} \left(2 - \frac{t}{T}\right)}$ $H(t \leq T) = \frac{1}{\pi} \left[\arccos\left(1 - \frac{t}{T}\right) - \left(1 - \frac{t}{T}\right) \sqrt{\frac{t}{T} \left(2 - \frac{t}{T}\right)} \right]$	oui	0.604
# 9	T	Type GRP (<i>Tangara</i> , 2005), $\alpha = 1.25$ $H(t) = \frac{\left(\frac{t}{2T}\right)^\alpha}{\left(\frac{t}{2T}\right)^\alpha + \left(1 - \frac{t}{2T}\right)^\alpha}$	oui	0.603
# 10	T	Type GRP, $\alpha = 1.50$	oui	0.603
# 11	T	Type GRP, $\alpha = 1.75$	oui	0.602
# 12	T	Type GRP, $\alpha = 2.00$	oui	0.601
# 13	T	Plat $H(t) = \frac{t}{2T}$ $h(t) = \frac{1}{2T}$	oui	0.602
# 14	T	Pur délai $H(t) = \begin{cases} 0 & \text{si } t < T, \\ 1 & \text{si } t \geq T. \end{cases}$	oui	0.538
# 15	(T, α)	Gamma (forme α et échelle T)	non	0.603
# 16	T	Gamma, régression $\alpha(T)$	non	0.598
# 17	T	Gamma, α fixé à la médiane	non	0.602
# 18	(T, α)	Weibull (forme α et échelle T)	non	0.606
# 19	T	Weibull, α fixé à la médiane	non	0.603

TAB. 9.4: Forme mathématique et performances en contrôle des hydrogrammes unitaires testés.

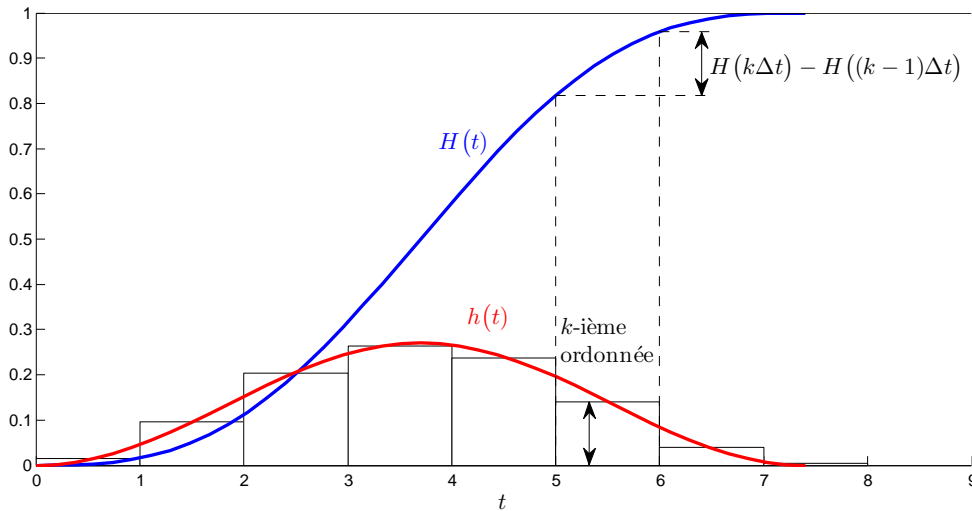


FIG. 9.10: Principe de calcul des ordonnées de l'hydrogramme unitaire discretisé à un pas de temps Δt . $h(t)$ et $H(t)$ désignent respectivement l'hydrogramme unitaire et son intégrale temporelle.

Nous ne constatons pas de forte sensibilité à la forme de l'hydrogramme unitaire, les seules solutions vraiment pénalisantes étant le pur délai et, bien sûr, la suppression de l'hydrogramme unitaire. Les performances des formes asymétriques (Gamma et Weibull) ne semblent pas justifier leur degré de liberté supplémentaire. Il semblerait que les formes « plates » (hydrogrammes puissance avec $\alpha < 1$ ou elliptique) soient légèrement plus performantes au pas de temps horaire, comme l'avait déjà noté *Mathevet* (2005) : nous retenons donc les solutions 5, 8 ou 9. Ces trois solutions sont d'ailleurs graphiquement très proches les unes des autres. En effet, l'hydrogramme elliptique a pour maximum (en $t = T$) la valeur $\frac{2}{\pi T} \simeq \frac{0.637}{T}$, tandis que les hydrogrammes « Puissance α » et « GRP α » ont tous deux une valeur maximum de $\frac{\alpha}{2T}$. Pour $\alpha = 1.25$, nous trouvons la valeur $\frac{0.625}{T}$ ce qui est très proche de la forme elliptique (cf. Figure 9.11).

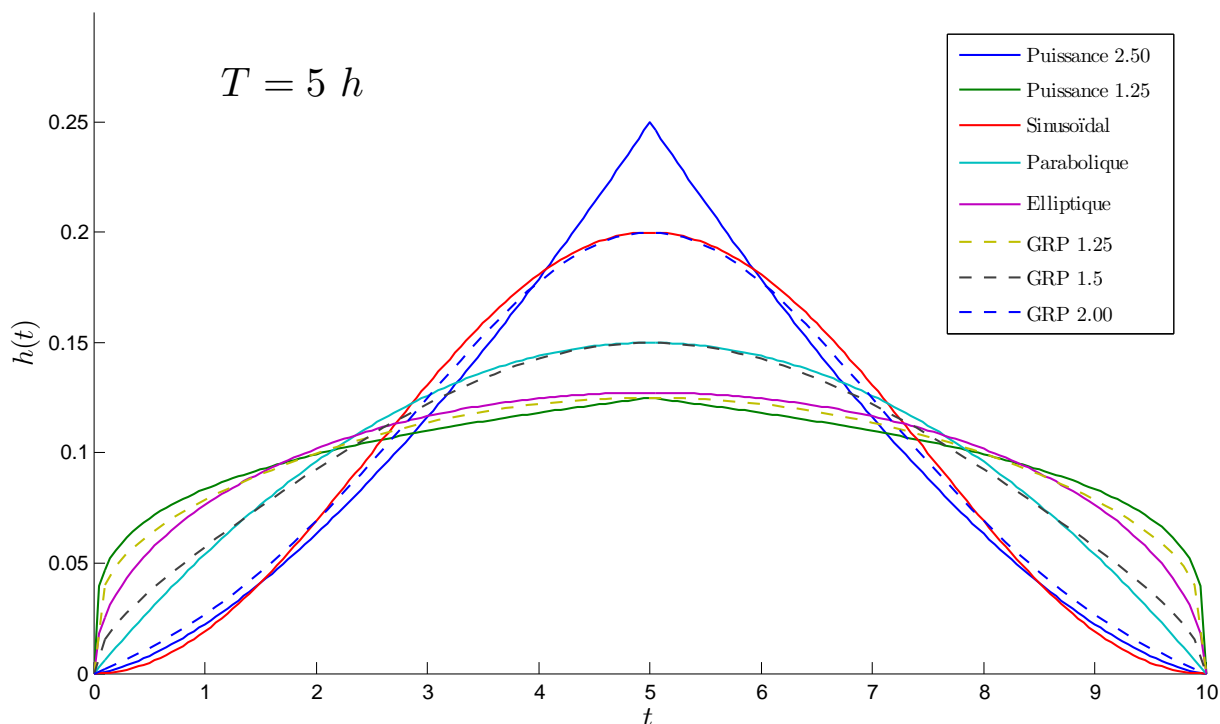


FIG. 9.11: Allure des formes symétriques d'hydrogrammes unitaires testées.

9.3.2 Essais de simplification de la structure M_{096}

Nous avons déjà mentionné au chapitre précédent le dilemme posé par la structure M_{096} . En effet, elle se classe première en termes de performances mais elle comporte également un nombre important de paramètres supplémentaires à identifier (sept, contre cinq pour la structure M_{043}). Nous avons donc cherché à réduire le nombre de paramètres de cette structure sans perte de performances, principalement en cherchant une relation entre le coefficient de partage θ_6 et la constante de temps θ_7 du réservoir linéaire.

Sur la Figure 9.12, nous avons placé les couples (θ_6, θ_7) obtenus en calage sur l'échantillon horaire. En fait, nous pouvons définir toute une zone de ce plan où il y a une équifinalité évidente : lorsque la fraction $\alpha = \theta_6$ alimentant le réservoir linéaire est très petite, la constante de temps $\tau = \theta_7$ peut prendre n'importe quelle valeur sans que cela ne modifie la simulation (zone entourée en pointillés). De même, lorsque la constante de temps est faible (jusqu'à quelques dizaines d'heures), elle risque d'interagir avec l'autre composant linéaire du modèle, l'hydrogramme unitaire. Nous cherchons donc à imposer une relation de la forme $\alpha = f(\tau)$ (en rouge sur la Figure 9.12), c'est-à-dire à imposer qu'une augmentation de la constante de temps du réservoir s'accompagne d'une alimentation accrue de ce réservoir, de façon à ce que cette constante soit mieux définie.

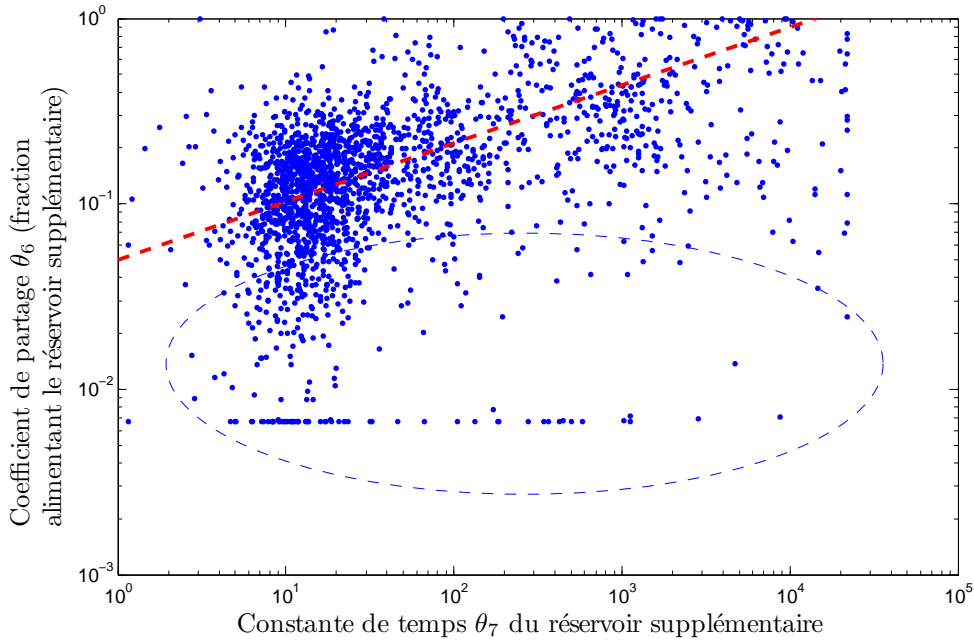


FIG. 9.12: Problèmes d'identification de la constante de temps θ_7 du réservoir linéaire lorsque l'alimentation de ce réservoir est faible.

En imposant une relation *a priori* entre ces deux paramètres (i.e. en supprimant un degré de liberté), on obtient la structure notée M'_{096} à six paramètres seulement, dont les performances sont données dans le Tableau 9.5. Nous constatons que les performances sont à peine dégradées par rapport à la structure à sept paramètres.

	GR4H (4 par.)	MORDOR 6 (6 par.)	M_{043} (5 par.)	M'_{096} (6 par.)	M_{096} (7 par.)
$C_{MM}(Q)$ moyen, contrôle	0.592	0.592	0.601	0.613	0.616
$C_{MM}(\sqrt{Q})$ moyen, contrôle	0.640	0.641	0.649	0.660	0.663
$C_{MM}(\ln Q)$ moyen, contrôle	0.567	0.572	0.603	0.607	0.612

TAB. 9.5: Classement de la structure M'_{096} parmi les modèles d'intérêt.

La contrainte ajoutée modifie assez fortement les distributions des paramètres correspondants, comme l'illustre la Figure 9.13. En particulier, la fréquence des fortes valeurs de la constante de temps est nettement diminuée, ce qui semble indiquer que les réservoirs très lents dans la structure initiale n'étaient en fait que très peu alimentés : l'écoulement ainsi retardé ne contribuant donc que très peu au débit total simulé.

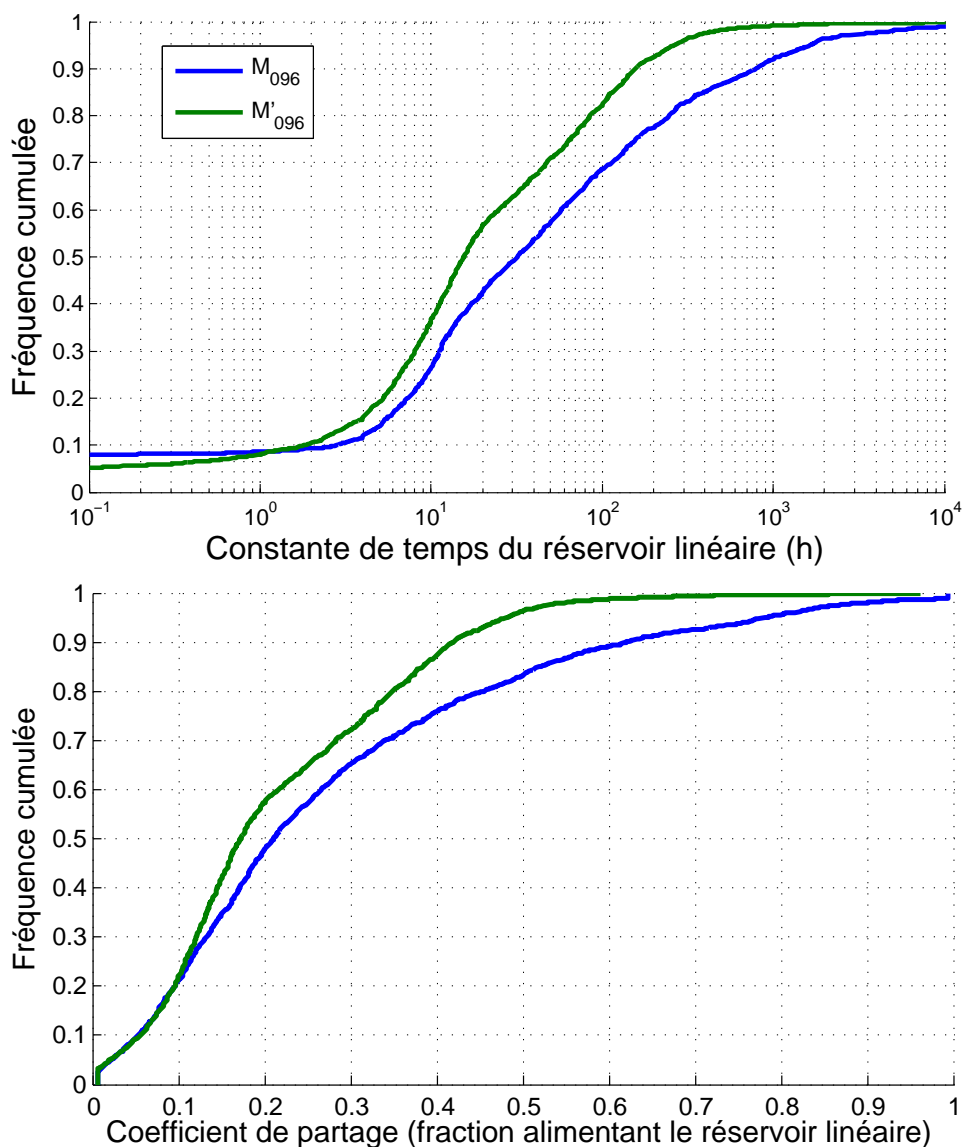


FIG. 9.13: Distributions des paramètres « coefficient de partage » et « constante de temps du réservoir linéaire » dans la structure M_{096} , où ces deux paramètres sont libres, et dans la structure M'_{096} où ils sont liés.

9.4 Résultats détaillés sur quelques bassins versants

Dans cette section, nous présentons quelques hydrogrammes pour essayer de donner un aperçu global de la qualité des simulations. Le choix d'une courte période sur quelques bassins pour résumer les performances d'un modèle est un jeu dangereux, et peut facilement (que ce soit volontaire ou non) donner une vision biaisée des faits. Pour cela, nous présentons les résultats pour quatre bassins versants contrastés :

- le premier (Figures 9.14 et 9.15) est le bassin versant de la Moder à Schweighouse-sur-Moder (A3301010, 622 km²), qui a un comportement plutôt classique hormis ses étiages qui sont soutenus de façon importante par les eaux souterraines ;
- le second bassin versant (Figures 9.16 et 9.17) est un bassin de la Craie : l'Hallue à

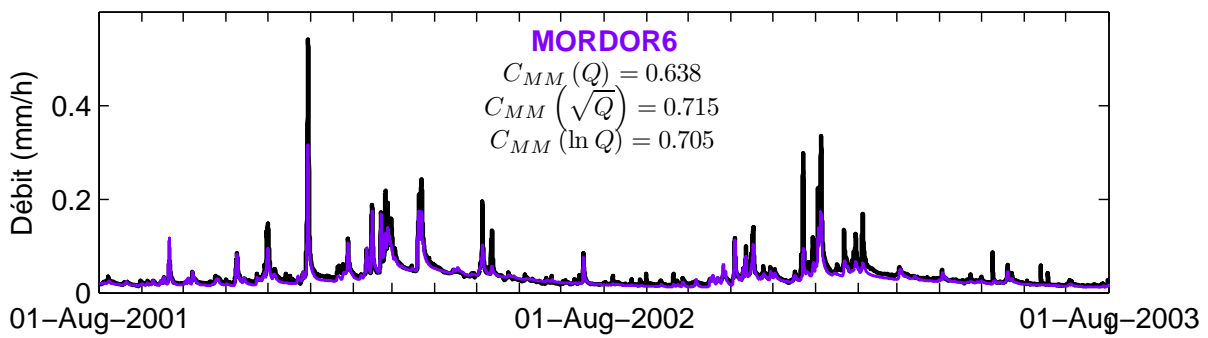
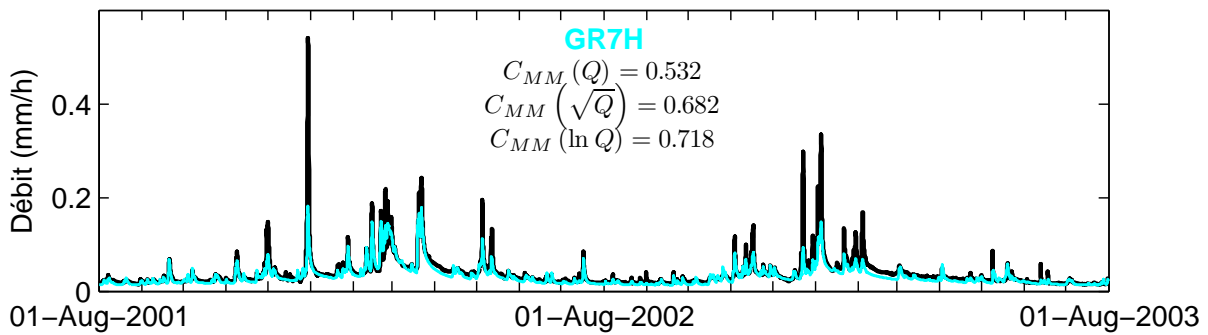
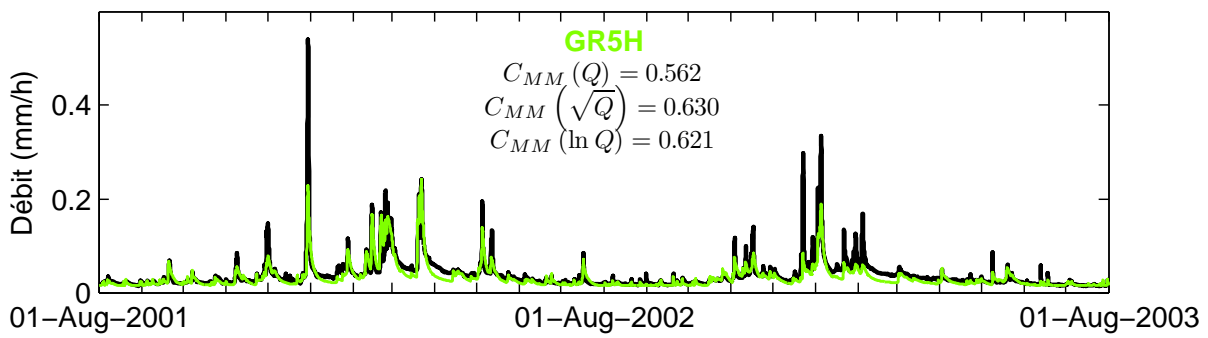
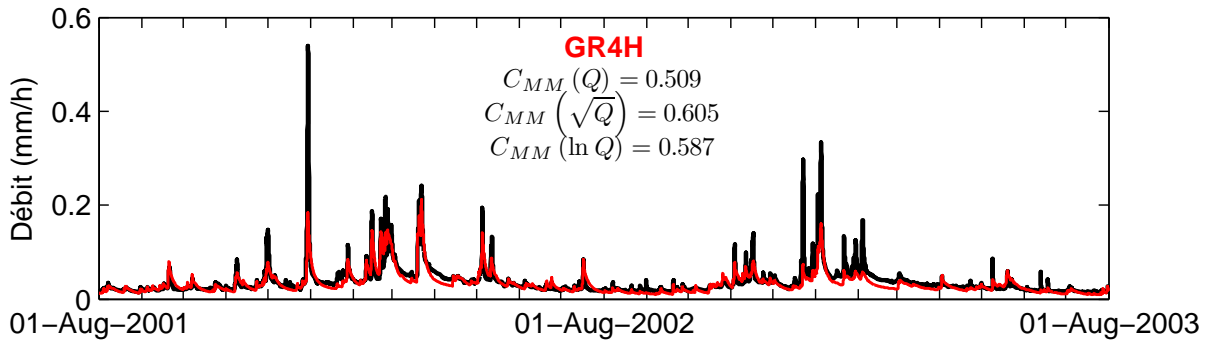
Bavelincourt (E6397010, 115 km²), un affluent de la Somme dont la dynamique est typique de cette région ;

- le troisième bassin (Figures 9.18 et 9.19) est un côtier breton, la Penzé (à Taulé, J2723010, 141 km²) ;
- le quatrième (Figures 9.20 et 9.21) est un petit bassin intermittent des Pyrénées Orientales, la Canterrane à Terrats (Y0325010, 31 km²).

Nous représentons à chaque fois une période de deux ans en contrôle pour chaque modèles (pages de gauche), ainsi que les distributions sur les dix années en contrôle (pages de droite).

Sans être trop pessimiste, ces résultats permettent tout de même de dresser un constat lucide de l'état des modèles :

- même lorsque les critères sont tous au vert, comme sur le troisième bassin, des erreurs d'un facteur deux en crue restent la règle ;
- on constate une difficulté à saisir toute la gamme de fréquences présentes dans les signaux de débit : ce cas classique est particulièrement flagrant sur le bassin de l'Hallue, où les modèles peinent à simuler la dynamique interannuelle et ignorent totalement les hautes fréquences (petites crues sur les enveloppes de récession par exemple) ;
- les bassins intermittents restent extrêmement difficiles à modéliser, aussi bien au moment de la cessation des écoulements que de leur reprise : dans le quatrième cas, nous constatons qu'à l'exception de MORDOR tous les modèles surestiment grossièrement la première crue de l'année 2002, et tous sans exception sous-estiment largement la seconde.



1

FIG. 9.14: Simulations horaires en contrôle pour la Moder à Schweighouse-sur-Moder (A3301010, 622 km²).

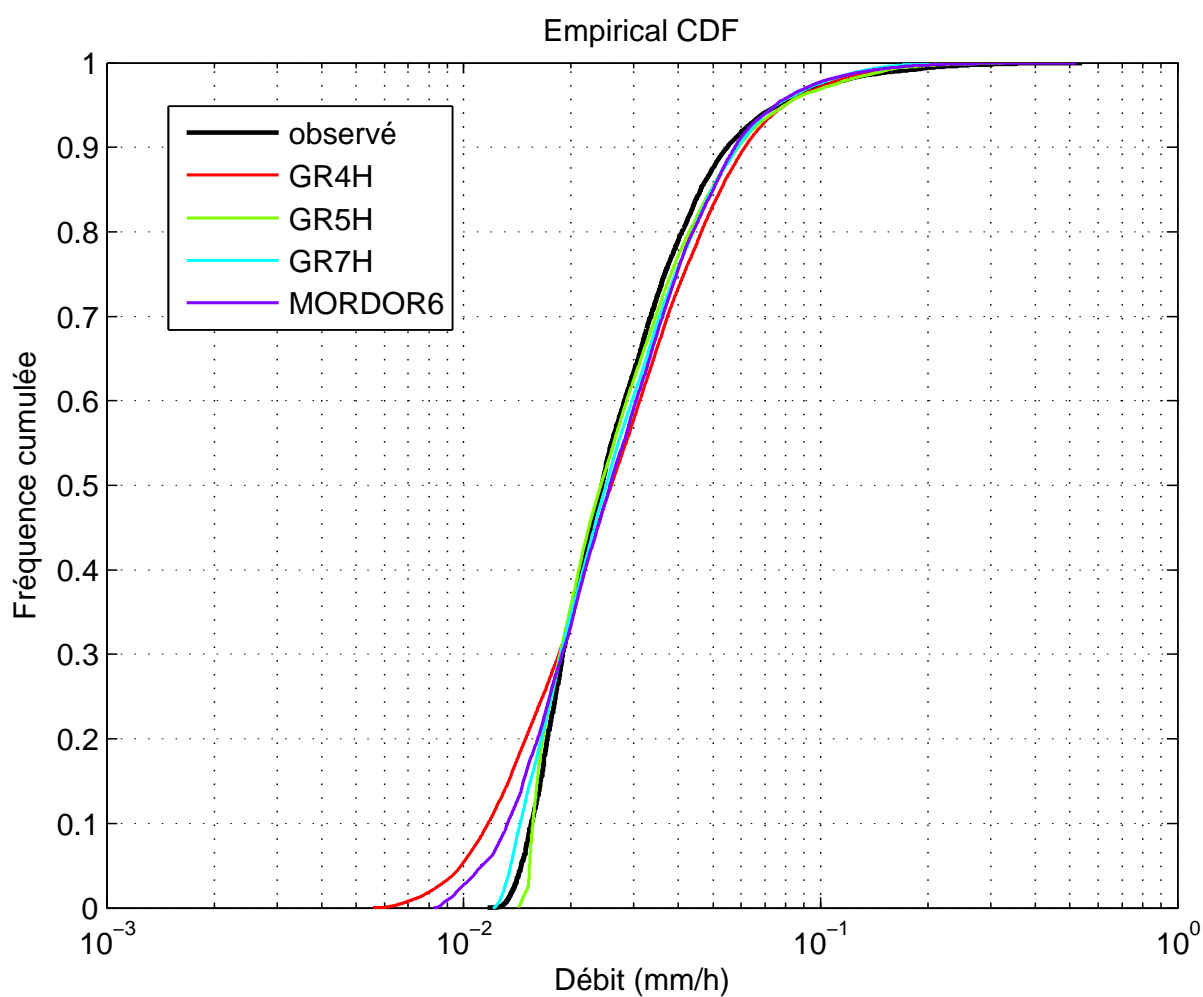


FIG. 9.15: Distributions des débits horaires observés et simulés pour la Moder à Schweighouse-sur-Moder (A3301010, 622 km²), sur la période 1995–2005 en contrôle.

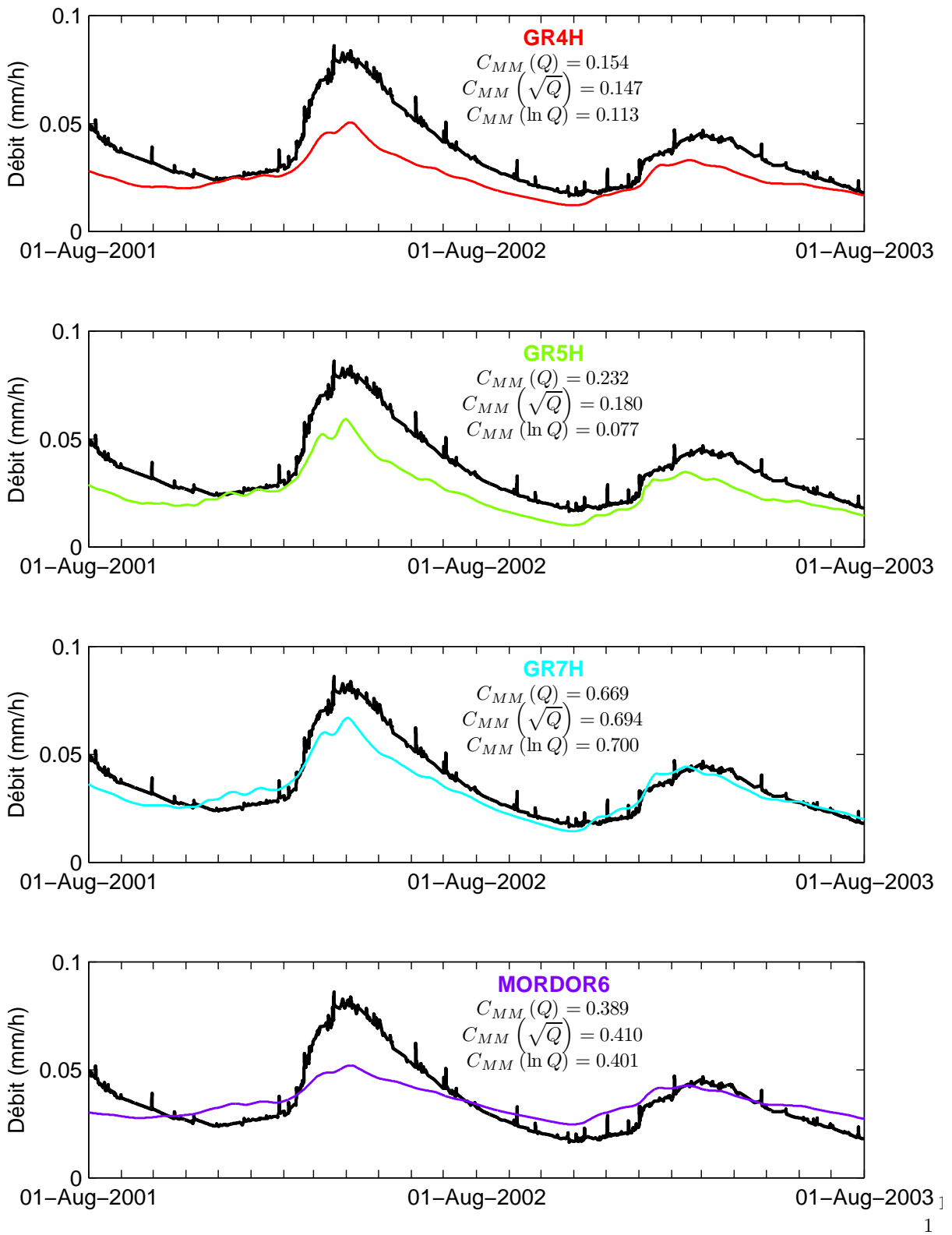


FIG. 9.16: Simulations horaires en contrôle pour l'Hallue à Bavelincourt (E6397010, 115 km²).

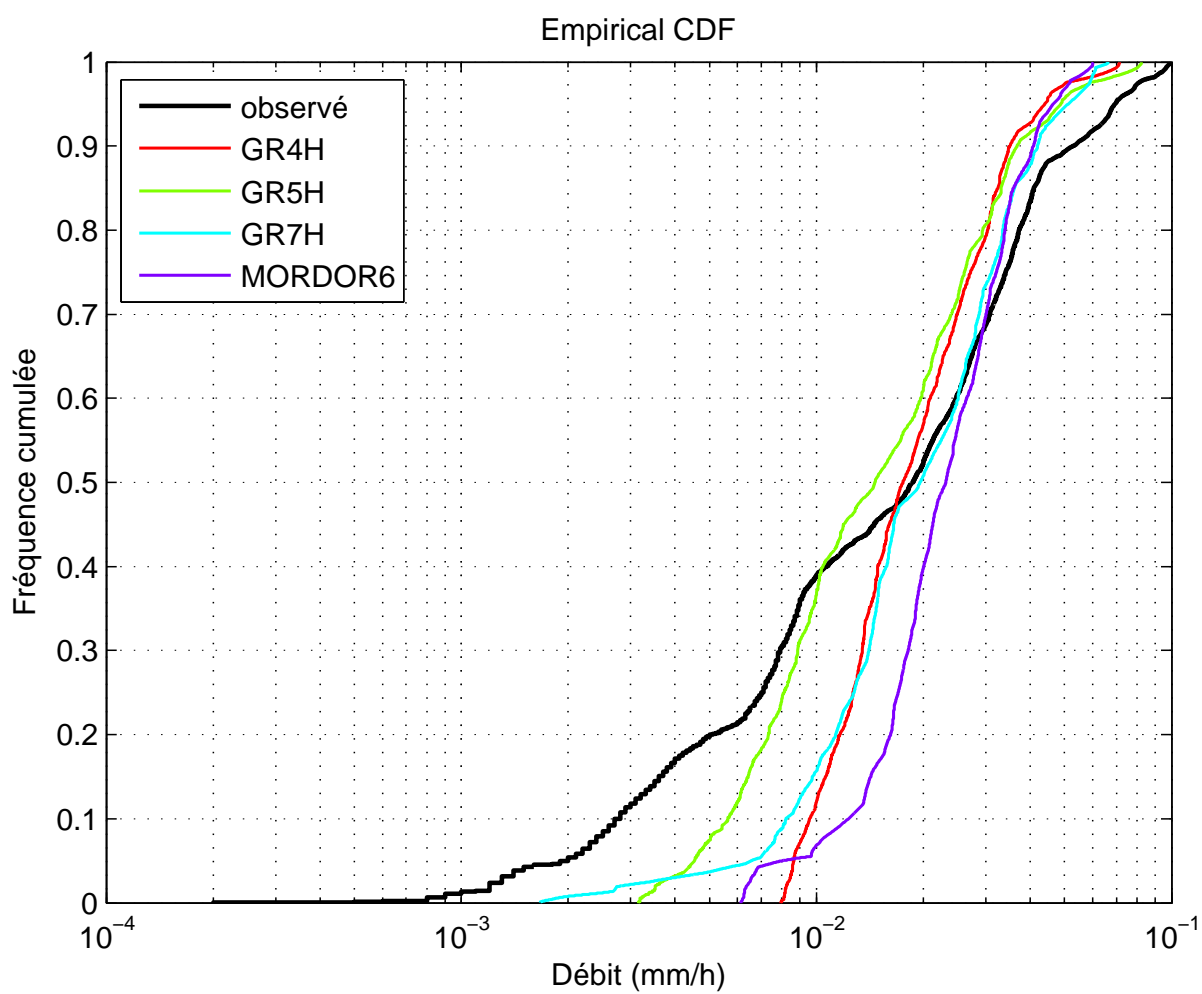


FIG. 9.17: Distributions des débits horaires observés et simulés pour l'Hallue à Bavelincourt (E6397010, 115 km²), sur la période 1995–2005 en contrôle.

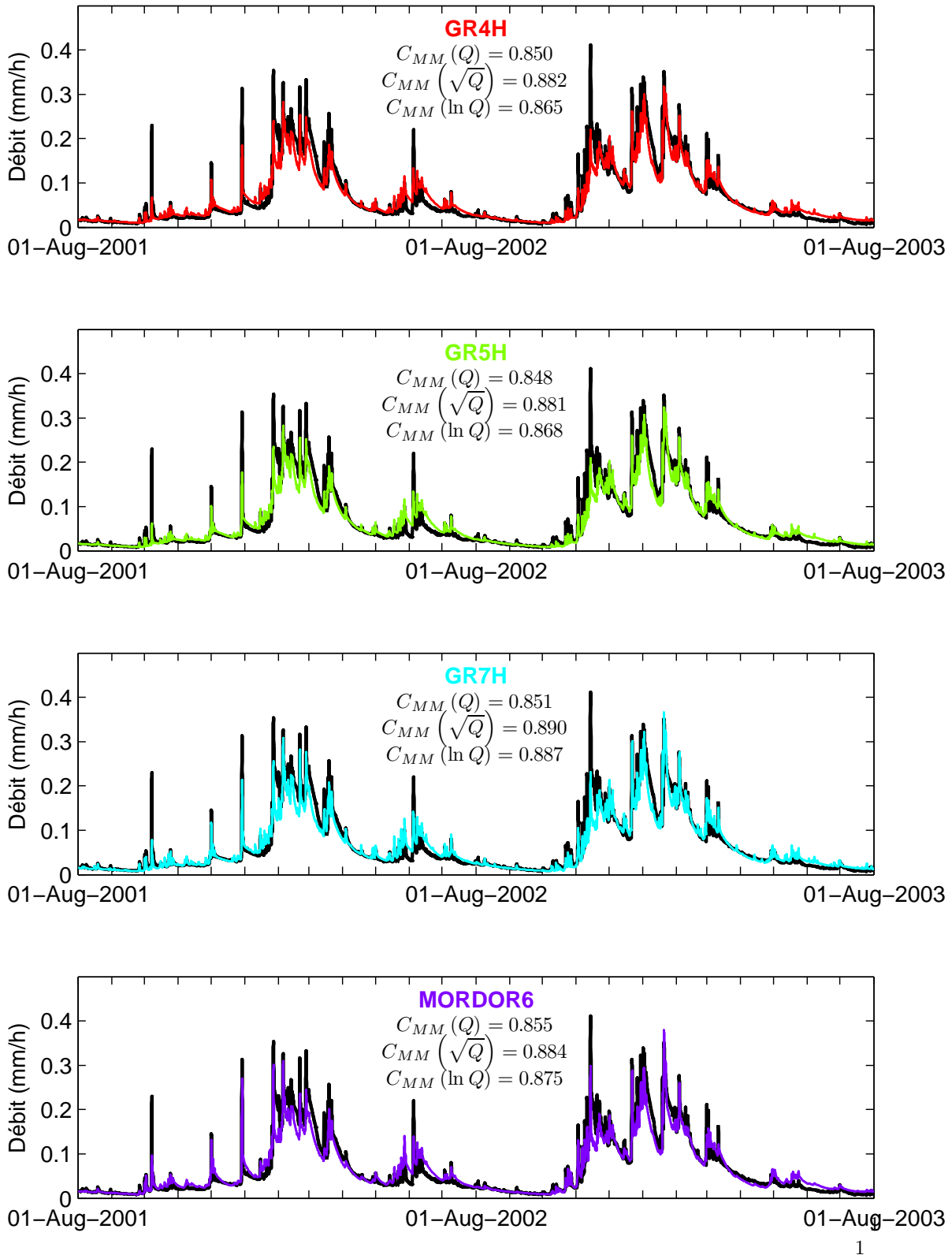


FIG. 9.18: Simulations horaires en contrôle pour la Penzé à Taulé (J2723010, 141 km²).

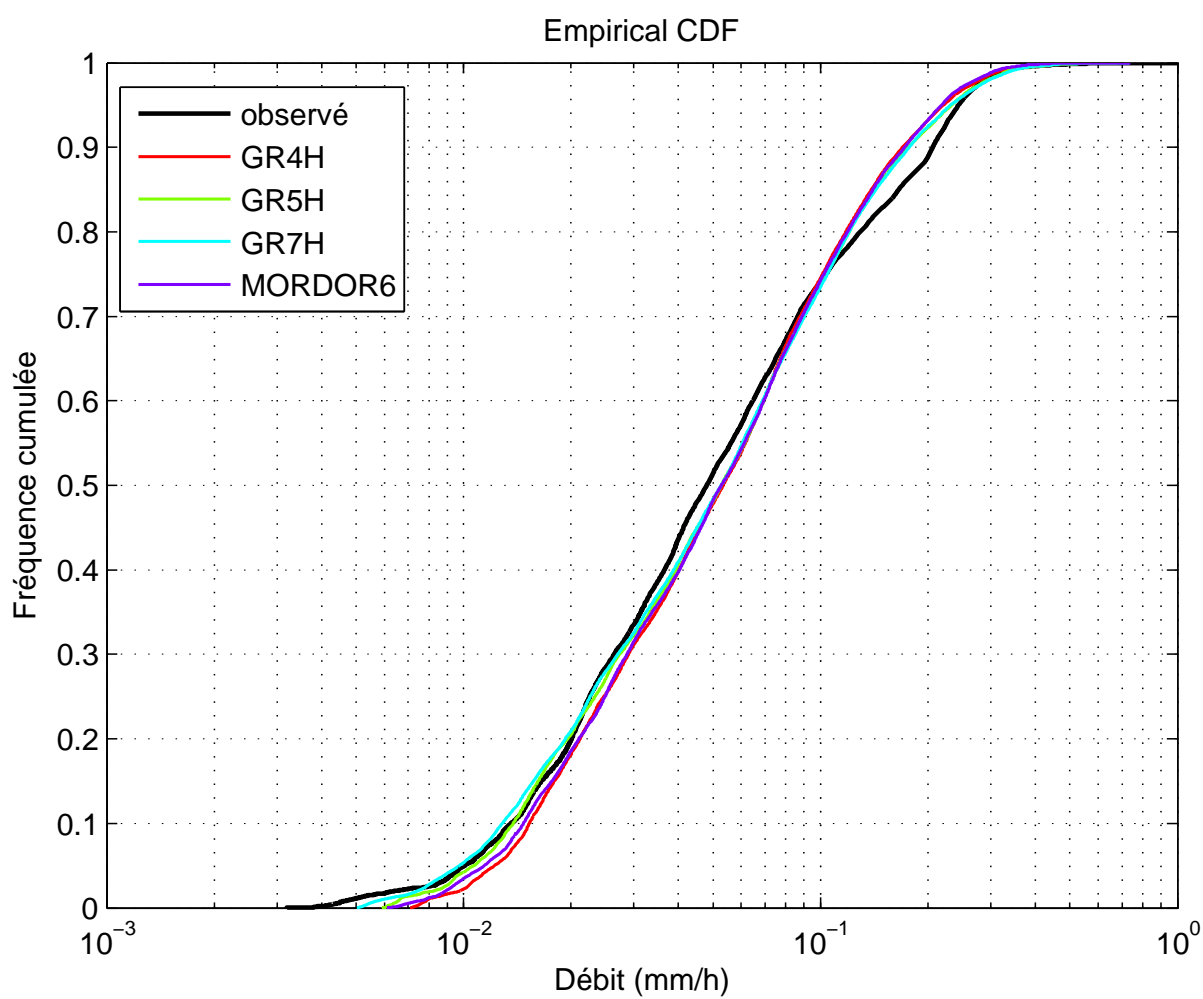


FIG. 9.19: Distributions des débits horaires observés et simulés pour la Penzé à Taulé (J2723010, 141 km²), sur la période 1995–2005 en contrôle.

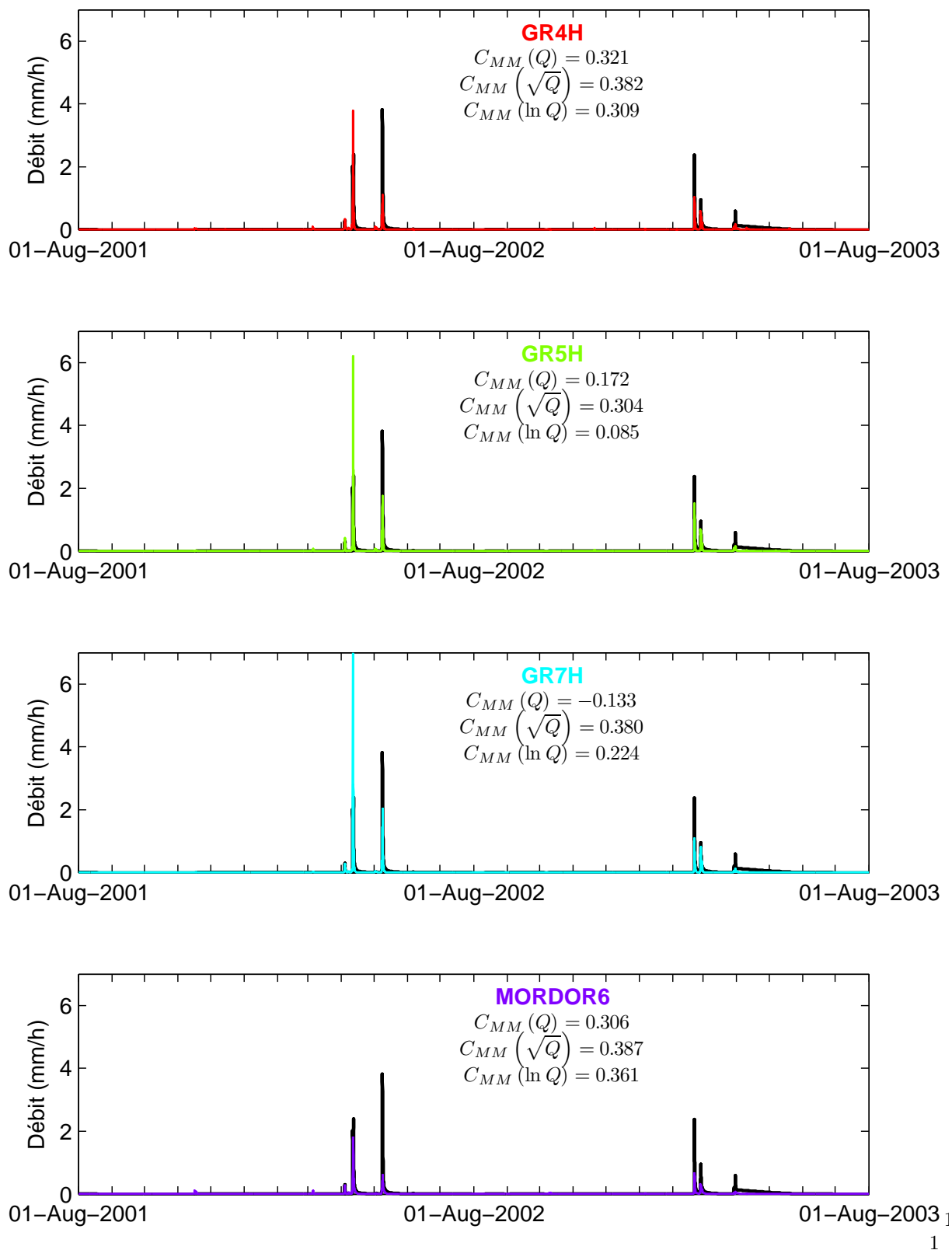


FIG. 9.20: Simulations horaires en contrôle pour la Canterrane à Terrats (Y0325010, 31 km²).

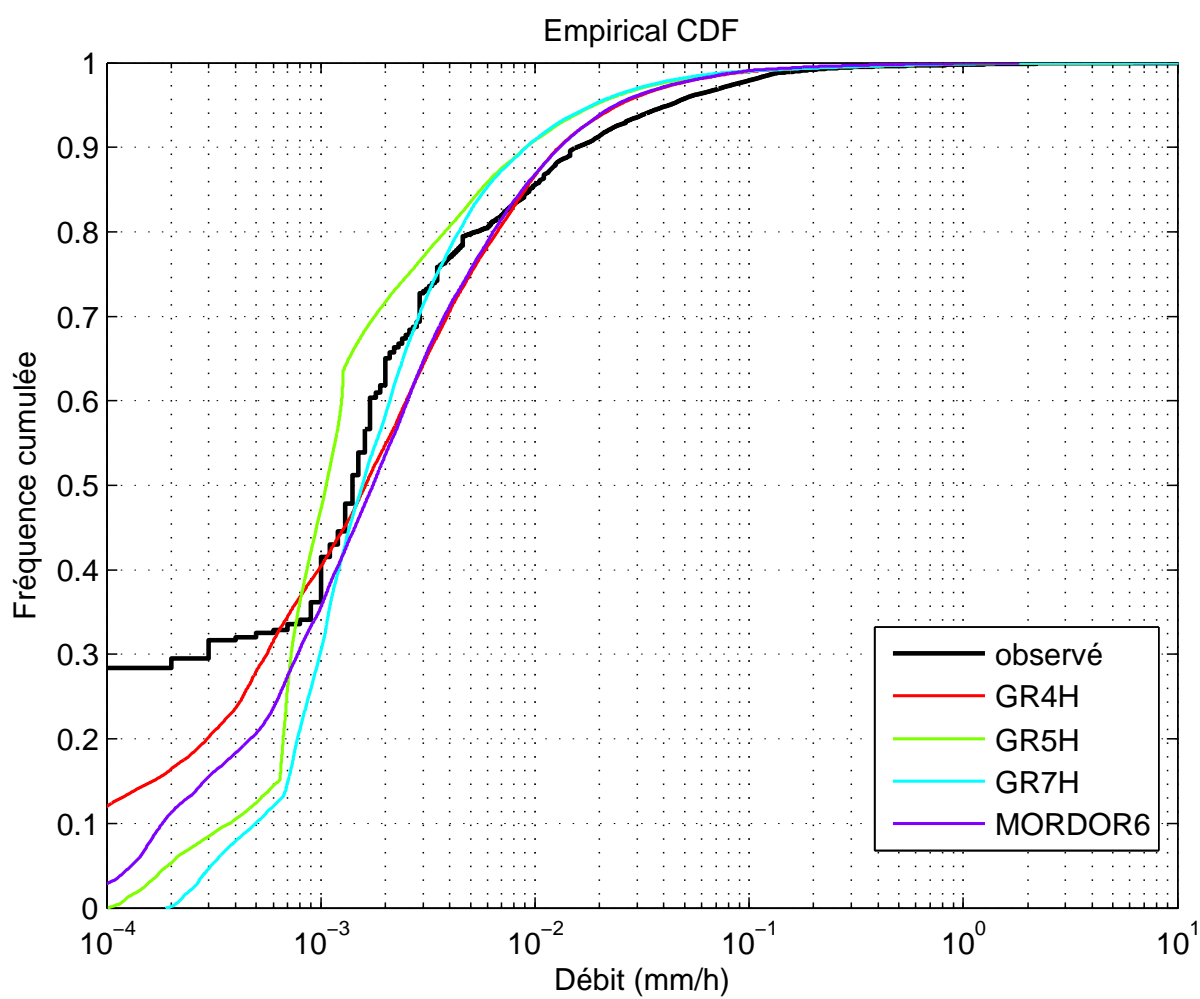


FIG. 9.21: Distributions des débits horaires observés et simulés pour la Canterrane à Terrats (Y0325010, 31 km²), sur la période 1995–2005 en contrôle.

9.5 Synthèse

Dans ce chapitre, nous avons cherché à améliorer une structure de modèle pluie-débit au pas de temps horaire, en nous basant à la fois sur les travaux de *Mathevet* (2005) et sur nos propres résultats au pas de temps journalier.

Il est apparu que la structure à cinq paramètres (notée M_{043}), qui se distinguait au pas de temps journalier, a donné des résultats également intéressants au pas de temps horaire : comme au pas de temps journalier, cette structure apporte peu d'améliorations à la simulation des hautes eaux mais fait progresser assez nettement la simulation des basses eaux et des étiages (+3.6 points de $C_{MM}(\ln Q)$).

Contrairement à ce que l'on pouvait imaginer *a priori*, l'ajout d'une composante linéaire lente (qui « représenterait » le stockage souterrain sans avoir nécessairement besoin de le considérer comme extérieur au système) ne semble pas constituer une voie d'amélioration évidente, et les paramètres correspondants ne sont pas mieux identifiables au pas de temps horaire qu'au pas de temps journalier. De façon générale, nous avons été confrontés au même problème que celui rencontré par *Mathevet* : la complexification de la fonction de routage (hydrogramme unitaire et routage non-linéaire confondus) ne semble pas être une voie directe d'amélioration des performances. Deux hypothèses au moins peuvent être avancées pour expliquer les limites rencontrées :

- une hypothèse sur la relation pluie-débit, selon laquelle la représentation du bassin n'aurait pas besoin d'être plus complexe au pas de temps horaire qu'au pas de temps journalier, toute la richesse provenant de l'information de pluie. Il y aurait ainsi peut-être moins de différences entre la relation pluie-débit au pas de temps journalier et celle au pas de temps horaire, qu'entre la relation pluie-débit au pas de temps mensuel et celle au pas de temps journalier, bien que le rapport des échelles de temps soit presque le même ;
- une hypothèse sur l'identifiabilité des paramètres : la structure des erreurs d'un modèle pluie-débit pourrait rendre difficile l'identification simultanée de paramètres simulant une vaste gamme de fréquences.

Les deux derniers chapitres seront l'occasion d'aborder ces questions, à travers deux aspects apparemment « périphériques » de la modélisation proprement dite : l'initialisation des modèles, et les critères utilisés pour caler et contrôler les modèles.

10

Stratégies d'initialisation des simulations pluie-débit

“ I have spoken of the rich years when the water was plentiful. But there were dry years too, and they put a terror on the valley. The water came in a thirty-year cycle. There would be five or six wet and wonderful years when there might be nineteen to twenty-five inches of rain, and the land would shout with grass. Then would come six or seven pretty good years of twelve to sixteen inches of rain. And then the dry years would come, and sometimes there would be only seven or eight inches of rain. The land dried up and the grasses headed out miserably a few inches high and great bare scabby places appeared in the valley. The live oaks got a crusty look and the sagebrush was gray. The land cracked and the springs dried up and the cattle listlessly nibbled dry twigs. Then the farmers and the ranchers would be filled with disgust for the Salinas Valley. The cows would grow thin and sometimes starve to death. People would have to haul water in barrels to their farms just for drinking. Some families would sell out for nearly nothing and move away. And it never failed that during the dry years the people forgot about the rich years, and during the wet years they lost all memory of the dry years. It was always that way.”

John STEINBECK — *East of Eden*.

10.1 Problématique

Au cours des deux chapitres précédents, nous avons testé un grand nombre de structures de modèles pluie-débit différant par la paramétrisation des échanges surface-souterrain. Nous avons en particulier discuté de l'intérêt d'introduire des dynamiques très lentes dans la structure des modèles (par exemple dans la structure à sept paramètres). Ce point nous a amené à nous pencher sur une difficulté souvent sous-estimée dans la modélisation (aussi bien conceptuelle que physique) : le problème de l'initialisation des simulations.

On peut s'accorder pour définir une simulation pluie-débit comme un ensemble de *lois de comportement* (la structure du modèle et un jeu de paramètres défini), de *forçages* (pluie et évapotranspiration) et de *conditions initiales* (principalement, les taux de remplissage des réservoirs en début de simulation). L'inadéquation d'un de ces trois éléments entraîne une diminution des performances du modèle (adéquation entre les sorties observées et simulées) : ceci est communément accepté pour les lois de comportement et les forçages, mais on part souvent du principe que les conditions initiales n'ont que peu de poids pourvu que l'on prévoise une courte période de « préchauffage » (initialisation), au cours de laquelle on laisse évoluer les états du modèle sans calculer la fonction d'erreur. Typiquement, on a coutume de « sacrifier » une à deux années de forçages climatiques, c'est-à-dire de jouer sur des stocks de quelques centaines de millimètres en plus ou en moins.

Mouelhi (2003) a étudié cette question de l'état initial pertinent au pas de temps annuel et mensuel. La solution qu'il propose se base sur l'établissement d'un « régime permanent », suivi d'une période de forçages réels, le régime permanent étant obtenu en forçant le modèle pendant une longue durée en répétant n fois une année moyenne (il propose $n = 20$ ans). Ce principe est à la fois simple et séduisant : le régime permanent sert à calculer l'état initial « le plus probable » en dehors de toute connaissance du contexte « présent » (i.e., est-on plutôt dans une période sèche ou humide en début de simulation), puis la période de forçages réels corrige cet état (d'autant mieux que la période est longue) avant de commencer à comparer les sorties.

Cependant, la transposition de cette méthode à des pas de temps plus fins (journalier et horaire) se heurte à deux problèmes :

- un problème purement numérique : le régime permanent devant être recalculé à chaque fois que l'on teste un nouveau jeu de paramètres (notamment, à chaque itération de la méthode de calage « pas-à-pas »), le coût en temps de calcul augmente vite lorsque l'on raffine le pas de temps. Ainsi, avec le même nombre d'années et la même complexité de structure, la durée de calcul du régime permanent est multipliée par 30 quand on passe du

pas de temps mensuel au pas de temps journalier, et encore par 24 quand on passe au pas de temps horaire¹,

- un problème de définition d'une année moyenne à la résolution de travail. En effet, si la précipitation moyenne du mois de Janvier sur un bassin donné conserve un sens, cela devient moins vrai pour la précipitation moyenne du 15 Janvier, et franchement absurde pour la précipitation moyenne du 15 Janvier entre 18 et 19 heures... Ainsi, la série constituée par les 365 valeurs de précipitations journalières moyennes successives ne ressemble en rien à une année de pluie moyenne : sa variance est bien plus faible que celle d'une année réelle, et cela empire au pas de temps horaire. Il faut aussi prendre en compte le fait que plus on raffine le pas de temps, moins on a en général de données pour calculer ces moyennes. Ainsi, avec cette méthode, on est obligé d'utiliser des années de forçages réels pour le calcul du régime permanent, le cas échéant plusieurs fois de suite pour assurer la convergence de l'état moyen : la technique devient extrêmement coûteuse, et l'on doit vite se limiter à une ou deux années d'initialisation (ce que nous avons fait aux chapitres précédents).

Ainsi, tout en conservant son esprit, la méthode élégante proposée par *Mouelhi* au pas de temps annuel et mensuel nécessite d'être adaptée pour des pas de temps plus fin, spécialement en ce qui concerne l'établissement des valeurs de régime permanent.

10.2 Impact d'une mauvaise spécification des conditions initiales sur les performances des modèles pluie-débit

Pour illustrer notre problème, nous montrons dans cette section les résultats d'une étude de sensibilité aux conditions initiales, au pas de temps journalier.

10.2.1 Principe

La Figure 10.1 illustre le principe de l'étude, qui se décompose en deux parties :

- Dans le split-sample test A1, les états initiaux sont considérés comme des paramètres et sont optimisés au même titre que les paramètres de structure en période de calage. En contrôle, nous transférons les paramètres de structure mais nous autorisons le recalage des états initiaux (niveau initial s_0 du réservoir sol et r_0 du réservoir de routage) : ce n'est donc pas un contrôle au sens strict du terme, même s'il est difficile de compenser un mauvais jeu de paramètres de structure uniquement en jouant sur deux états initiaux,

¹L'effet n'est pas très sensible pour la simulation d'un bassin, mais il le devient pour un millier...

- Dans le split-sample test A2, nous procédons exactement de la même façon que précédemment en calage, mais nous n'autorisons plus le recalage des états initiaux. À la place, nous utilisons la valeur moyenne des états au début de chaque année hydrologique (1^{er} Août). Ainsi, si la période de calage est 2000–2005, nous calons les paramètres $\{\theta_i\}$ et les états initiaux s_0 et r_0 au 01/08/2000, puis nous calculons la moyenne des quatre valeurs de s et r au 01/08/2001, 01/08/2002, 01/08/2003 et 01/08/2004. Ces moyennes sont utilisées pour initialiser les réservoirs au début de la période de contrôle 1995–2000 (i.e. au 01/08/1995). Nous aurions également pu aussi utiliser les valeurs s_0 et r_0 optimisées en calage, mais cela aurait été moins robuste car lorsque la première année de la période de calage est lacunaire, ces niveaux initiaux peuvent prendre des valeurs extrêmes sans que cela ait nécessairement un impact sur le critère.

Dans les deux cas, nous n'utilisons pas de période de préchauffage et nous commençons le calcul de la fonction objectif dès le premier jour de la période de calage ou de contrôle. Par construction du test, dans les périodes de contrôle, les paramètres de structure sont exactement les mêmes entre les cas A1 et A2 (puisque les calages sont identiques) : seuls diffèrent les états initiaux, ce qui permet d'évaluer l'impact de leur modification.

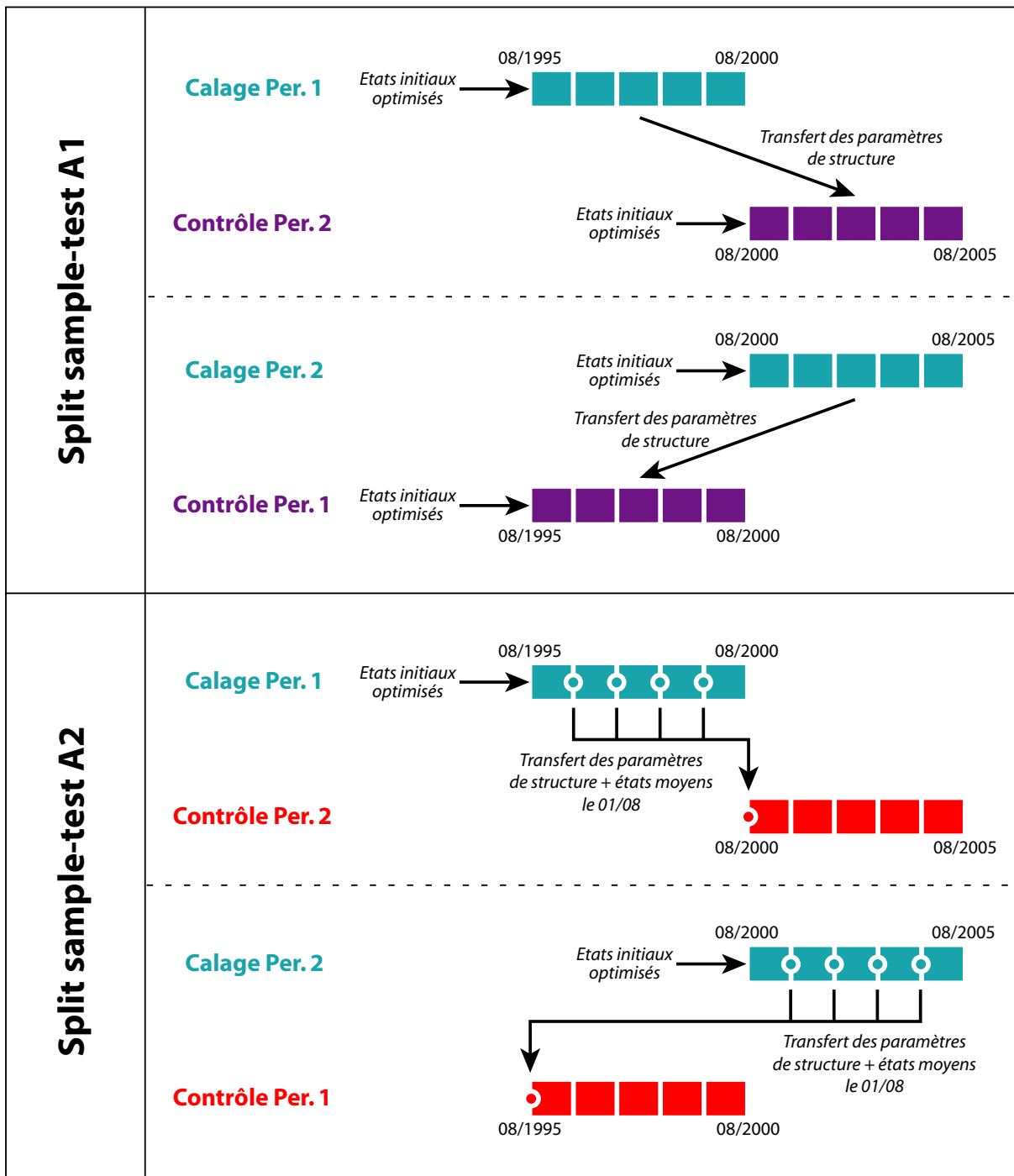


FIG. 10.1: Principe de l'étude de sensibilité aux conditions initiales. Nous procédons à deux split-sample tests : dans le premier (A1, haut), les états initiaux sont calés au même titre que les paramètres de structure puis, en contrôle, nous transférons les paramètres de structure mais nous autorisons le recalage des états initiaux (i.e., le niveau initial s_0 du réservoir sol et celui r_0 du réservoir de routage) ; dans le second test (A2, bas), nous calons de la même façon qu'en A1 mais nous ne recalons pas les états initiaux en contrôle : ceux-ci sont fixés à leur moyenne en début d'année hydrologique pendant la période de calage. Dans aucun des deux tests il n'y a de période de préchauffage : le calcul de la fonction objectif (en calage comme en contrôle) commence dès le début de chaque période (01/08/1995 et 01/08/2000 respectivement).

10.2.2 Résultats

La Figure 10.2 montre les distributions de performance obtenues dans les tests A1 et A2. La première boîte (à gauche) est commune aux tests A1 et A2. Fort logiquement, nous constatons une dégradation des performances en contrôle, qui est cependant moindre dans le test A1 que dans le test A2 (ce qui est également cohérent puisque l'on a autorisé le recalage des états initiaux dans le test A1 et non en A2).

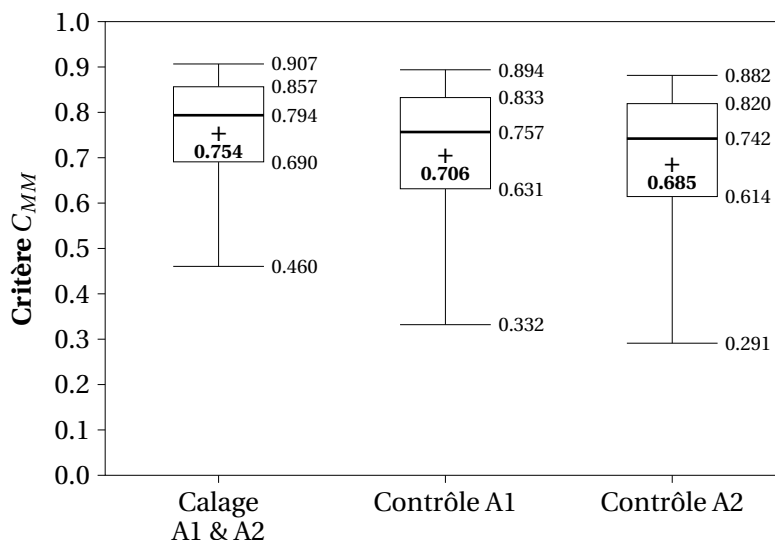


FIG. 10.2: Distribution des performances (critère C_{MM}) en calage (commun à A1 & A2) et en contrôle (A1 puis A2) sur les 2080 bassins-périodes de l'échantillon français.

La Figure 10.3 montre clairement que la chute de performance en contrôle peut être en partie due à un état initial inapproprié, pour peu que la plage de variation des états en début d'année hydrologique au calage ne reflète pas toute la distribution possible.

Tous ces résultats sont assez intuitifs et peu surprenants. Cependant, le plus intéressant reste de regarder *quels sont* les bassins qui sont affectés par les plus fortes baisses de performance entre les contrôles A1 et A2. Ces bassins sont mis en évidence sur la carte de la Figure 10.4. Nous constatons sans ambiguïté possible que la quasi-totalité de ces bassins est située dans le Bassin Parisien, et plus spécialement en Normandie et Picardie crayeuses. Un autre petit groupe est situé dans le secteur des vallées de Vienne, au sud de Lyon, sur des aquifères molassiques.

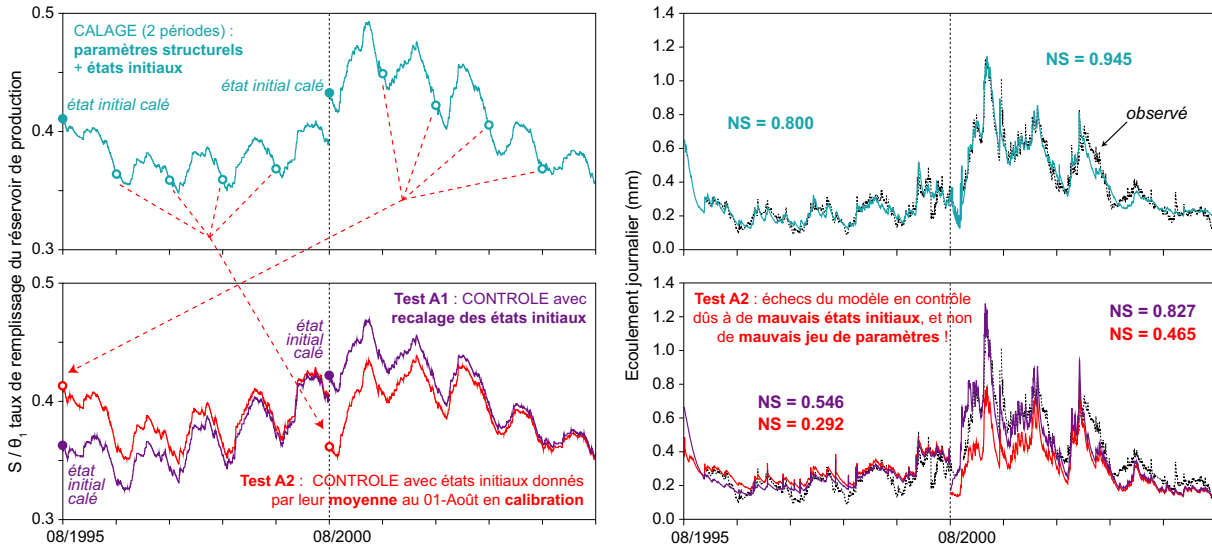


FIG. 10.3: Exemple de résultat obtenu pour le test de sensibilité aux conditions initiales : cas de l'Avre à Moreuil (E6406010). En calage (haut gauche et haut droit, courbes vertes), le débit est correctement simulé pendant les deux périodes (1995–2000 et 2000–2005). Nous voyons très nettement que le niveau optimisé en début de deuxième période (01/08/2000) est nettement plus élevé qu'à la même période de l'année entre 1995 et 1999. Lorsque nous autorisons le recalage des états initiaux (test A1), les paramètres obtenus sur la période 1995–2000 donnent des résultats acceptables en contrôle sur 2000–2005, et inversement ($NS = 0.827$ et 0.546 respectivement, en violet). En revanche, lorsque ces états initiaux en contrôle sont fixés à la moyenne obtenue pendant la période de calage, les résultats sont nettement dégradés.

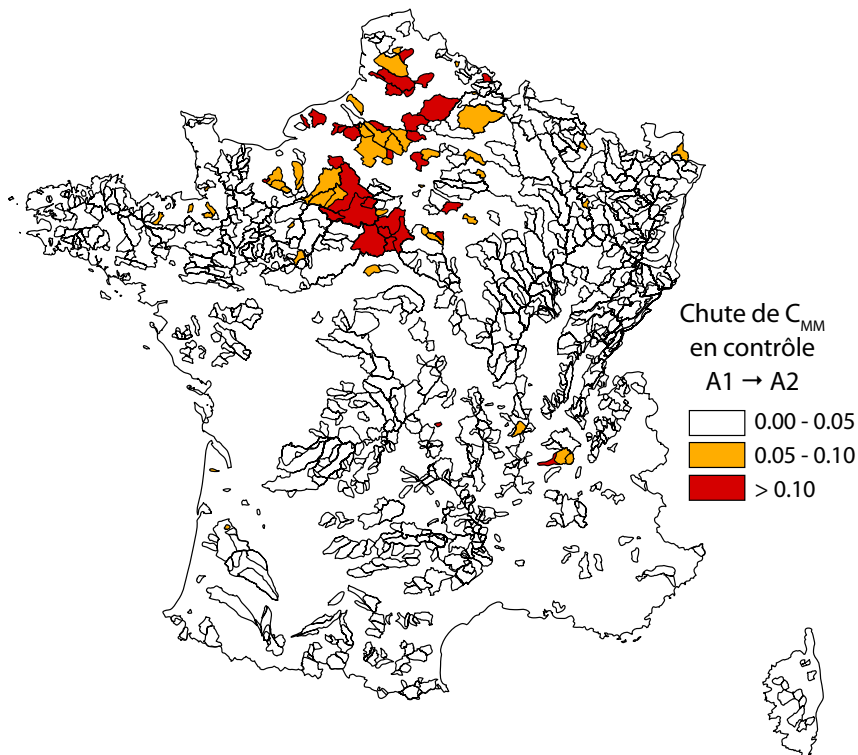


FIG. 10.4: Localisation des bassins versants les plus dépendants des conditions initiales d'après les tests A1 et A2. La plupart sont situés sur la Craie, dans le Bassin Parisien.

La Figure 10.4 permet donc de relier plus nettement le problème de l'initialisation à celui de la thèse. En effet, nous avons mis en évidence la sensibilité des performances des modèles aux conditions initiales, dans le cas où les valeurs de certains paramètres (capacités des réservoirs, et plus spécifiquement celle du réservoir de production) se situent plutôt en queue de distribution (valeurs « atypiques » mais pas nécessairement absurdes, qui peuvent effectivement traduire un fonctionnement longue mémoire). Les bassins versants concernés s'avèrent être principalement des bassins influencés par les eaux souterraines, ce qui n'est pas surprenant, puisque nous avons déjà mentionné ce type de comportement possible au chapitre 3. Le bassin de la Figure 10.3, l'Avre, est un affluent de la Somme : les rivières de cette région sont connues pour être en connexion étroite avec la nappe de la Craie, qui a elle-même une dynamique pluri-annuelle. Le contraste y est même particulièrement marqué entre les périodes 1995–2000 et 2000–2005, qui se trouvent être les deux périodes du *split-sample test* : les fameuses crues de 2001 sont situées en tout début de la seconde période, qui succède à une période 1995–2000 où les écoulements sont peu abondants.

Cependant, il ne faudrait pas conclure de ce test que le calage des états initiaux, au même titre que les paramètres de structure, constitue une solution à ces problèmes. En effet, le cas de la Figure 10.3 a été choisi *ad hoc* mais nous rencontrons également des cas où l'optimisation des états initiaux interfère avec l'identification des paramètres de structure, occasionnant une perte de robustesse (diminution des performances en contrôle par rapport à une simulation où le taux de remplissage serait systématiquement fixé à 0.5, par exemple).

10.3 Obtention du niveau d'équilibre du réservoir sol par une analyse de son comportement asymptotique

Les différentes solutions envisagées jusqu'ici présentent chacune des avantages et des inconvénients :

- le calcul d'un régime permanent suivi d'une courte période de forçage réel est une solution séduisante, mais très coûteuse à un pas de temps fin,
- le calage des états initiaux crée trop de degrés de liberté, ce qui peut nuire à la robustesse des jeux de paramètres de structure (et accessoirement au temps de calcul nécessaire à leur identification),
- l'utilisation de valeurs fixes (identiques pour tous les bassins, par exemple la médiane des valeurs obtenues en calage au test A, $s_0 = 0.41$ et $r_0 = 0.38$) n'en crée pas assez, et contraint excessivement les simulations.

L'origine du problème est que nous ne savons pas *a priori* si l'état initial a une importance ou pas. Dans certains cas, nous pouvons le suspecter, notamment lorsque les capacités des réservoirs deviennent grandes. Nous pourrions alors envisager de n'optimiser les états initiaux que lorsque la valeur de certains paramètres franchit un certain seuil, ou bien alors de restreindre la plage de variation de ces mêmes paramètres de façon à éviter les cas où l'état initial est trop important, etc. Ce dernier point nous amène à la question suivante : cette sensibilité aux conditions initiales est-elle un défaut du modèle (qu'il faudrait alors éliminer), ou traduit-elle la réalité du fonctionnement de certains systèmes hydrologiques ? Clairement, le test précédent et surtout la localisation géographique des bassins à problèmes exclut que cette sensibilité soit un pur artefact. Nous ne devons donc pas être surpris par ce fait ni l'exclure *a priori*, mais nous devons y être attentifs. En effet, si nous testons des modifications de structure des modèles, en prêtant tout spécialement attention aux bassins versants influencés par les nappes, nous pouvons ruiner nos efforts ou, pire encore, trouver des améliorations qui ne sont que des artefacts, si nous ne faisons pas attention à la façon d'initialiser les simulations.

Dans le paragraphe suivant, nous restreindrons le problème de l'initialisation au cas du réservoir de production, et nous nous concentrerons plus spécialement sur la détermination du niveau d'équilibre de ce réservoir dans le cas où sa capacité (paramètre θ_1 , en mm) devient grande. En effet, ce cas constitue l'essentiel des difficultés rencontrées.

10.3.1 Rappel sur la formulation des équations d'évolution du réservoir de production (*Edijatno, 1991*)

L'évolution du niveau du réservoir de production est décrite par un jeu de trois équations (cf. Figure 10.5) :

1. Les rendements respectif de la pluie et de l'ETP sont déterminés par les deux équations instantanées :

$$\begin{cases} dP_r = s^2 dP_n \quad \text{et} \quad dP_s = \theta_1 ds = (1 - s^2) dP_n & \text{si } P_n = P - E > 0, \\ \theta_1 ds = -s(2 - s) dE_n & \text{si } E_n = E - P > 0. \end{cases}$$

2. Dans la version proposée par *Perrin (2000)*, le réservoir est également soumis à une percolation dont le cumul entre (t, s_t) et $(t + 1, s_{t+1})$ est donné par l'intégration d'une loi de vidange d'exposant 5 et de paramètre $2.25 \theta_1$:

$$Q_t = \theta_1 q_t = \theta_1 s_t \left(1 - \left[1 + \left(\frac{s_t}{2.25} \right)^4 \right]^{-\frac{1}{4}} \right)$$

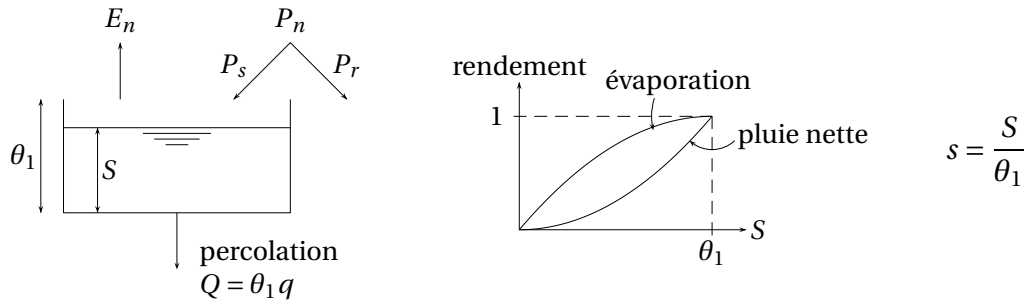


FIG. 10.5: Courbes de rendement de la pluie et de l'ETP dans les modèles GR. Cette fonction de production est commune à tous les modèles GR du pas de temps mensuel au pas de temps horaire.

10.3.2 Formulation intégrée sur un pas de temps (*Edijatno, 1991*)

Dans le cas d'une pluie nette, nous obtenons le cumul transféré vers la fonction de routage (P_r) et le cumul stocké (P_s) en intégrant l'équation d'évolution :

$$\frac{ds}{1-s^2} = d\left(\frac{P_n}{\theta_1}\right)$$

qui s'intègre en :

$$\begin{aligned} \operatorname{arctanh}(s_{t+1}) - \operatorname{arctanh}(s_t) &= \frac{P_n}{\theta_1} \\ \operatorname{arctanh}(s_{t+1}) + \operatorname{arctanh}(-s_t) &= \frac{P_n}{\theta_1} \end{aligned}$$

En utilisant la relation suivante¹

$$\operatorname{arctanh}(a) + \operatorname{arctanh}(b) = \operatorname{arctanh}\left(\frac{a+b}{1+ab}\right)$$

on obtient

$$\begin{aligned} \operatorname{arctanh}\left(\frac{s_{t+1} - s_t}{1 - s_t s_{t+1}}\right) &= \frac{P_n}{\theta_1} \\ s_{t+1} - s_t &= (1 - s_t s_{t+1}) \tanh\left(\frac{P_n}{\theta_1}\right) \\ s_{t+1} \left[1 + s_t \tanh\left(\frac{P_n}{\theta_1}\right)\right] &= s_t + \tanh\left(\frac{P_n}{\theta_1}\right) \\ s_{t+1} &= \frac{s_t + \tanh\left(\frac{P_n}{\theta_1}\right)}{1 + s_t \tanh\left(\frac{P_n}{\theta_1}\right)} \end{aligned} \quad (10.1)$$

¹démonstrable en utilisant le fait que $\operatorname{arctanh}(a) = \frac{1}{2} \ln\left(\frac{1+a}{1-a}\right)$

On déduit de cette expression le terme P_s :

$$\begin{aligned}
 P_s &= \theta_1 s_{t+1} - \theta_1 s_t = \frac{\theta_1 s_t + \theta_1 \tanh\left(\frac{P_n}{\theta_1}\right)}{1 + s_t \tanh\left(\frac{P_n}{\theta_1}\right)} - \frac{\theta_1 s_t + \theta_1 s_t^2 \tanh\left(\frac{P_n}{\theta_1}\right)}{1 + s_t \tanh\left(\frac{P_n}{\theta_1}\right)} \\
 P_s &= \frac{\theta_1 (1 - s_t^2) \tanh\left(\frac{P_n}{\theta_1}\right)}{1 + s_t \tanh\left(\frac{P_n}{\theta_1}\right)} \quad (10.2)
 \end{aligned}$$

Dans le cas d'une évaporation nette, l'équation d'évolution est cette fois :

$$\begin{aligned}
 ds &= -s(2-s)d\left(\frac{E_n}{\theta_1}\right) \\
 -\frac{ds}{s(2-s)} &= d\left(\frac{E_n}{\theta_1}\right) \\
 -\frac{ds}{1-(s-1)^2} &= d\left(\frac{E_n}{\theta_1}\right)
 \end{aligned}$$

Cette équation différentielle s'intègre en :

$$-\operatorname{arctanh}(s_{t+1} - 1) + \operatorname{arctanh}(s_t - 1) = \frac{E_n}{\theta_1}$$

Soit encore

$$\begin{aligned}
 \operatorname{arctanh}\left(\frac{(1 - s_{t+1}) + (s_t - 1)}{1 + (1 - s_{t+1})(s_t - 1)}\right) &= \frac{E_n}{\theta_1} \\
 \frac{s_t - s_{t+1}}{s_t + s_{t+1}(1 - s_t)} &= \tanh\left(\frac{E_n}{\theta_1}\right) \\
 s_{t+1}\left(- (1 - s_t) \tanh\left(\frac{E_n}{\theta_1}\right) - 1\right) &= s_t\left(\tanh\left(\frac{E_n}{\theta_1}\right) - 1\right) \\
 s_{t+1} &= s_t \frac{1 - \tanh\left(\frac{E_n}{\theta_1}\right)}{1 + (1 - s_t) \tanh\left(\frac{E_n}{\theta_1}\right)} \quad (10.3)
 \end{aligned}$$

On en déduit également l'évaporation soustraite au réservoir :

$$E_r = \theta_1 (s_t - s_{t+1}) = \frac{\theta_1 s_t (2 - s_t) \tanh\left(\frac{E_n}{\theta_1}\right)}{1 + (1 - s_t) \tanh\left(\frac{E_n}{\theta_1}\right)} \quad (10.4)$$

10.3.3 Niveau d'équilibre quand θ_1 devient très grand

Quand θ_1 devient très grand, l'initialisation du réservoir sol devient un problème crucial car son inertie est alors très forte. La Figure 10.6 montre l'évolution, sur dix ans, du niveau d'un réservoir de 15 000 mm (valeur rare mais pas improbable) et de sa production, avec la

même entrée de pluie et d'ETP et simplement des niveaux initiaux différents. Si la condition initiale s_0 s'écarte un peu trop du niveau d'équilibre, le retour vers ce niveau d'équilibre va l'emporter sur toute autre dynamique : cette non-stationnarité factice peut s'avérer très problématique.

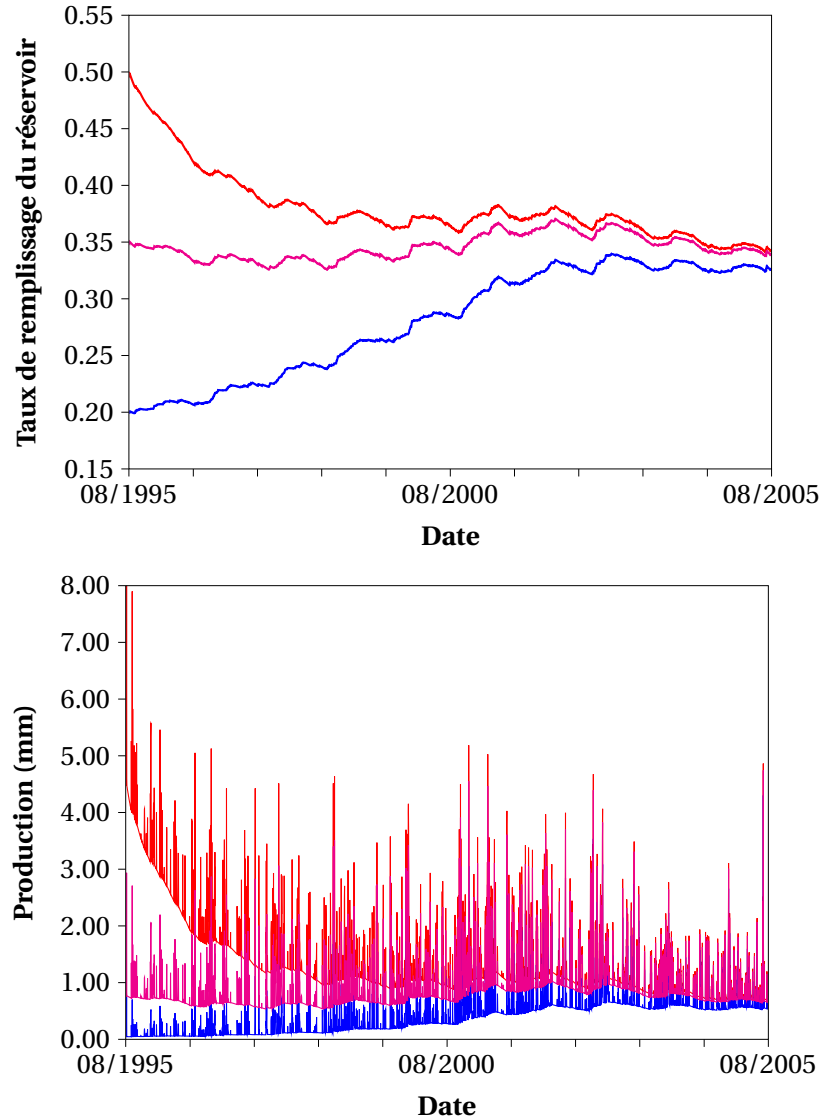


FIG. 10.6: Impact de la condition initiale sur l'évolution du niveau d'un réservoir de $\theta_1 = 15000$ mm (haut) et sur sa production (bas). En bleu : $s_0 = 0.2$; en magenta : $s_0 = 0.35$; en rouge : $s_0 = 0.5$.

Quand θ_1 devient grand comparé aux valeurs maximales de $|P - E|$, on peut utiliser les équations instantanées comme approximation, ou alors, ce qui revient exactement au même, écrire $\tanh\left(\frac{P-E}{\theta_1}\right) \simeq \frac{P-E}{\theta_1}$. Pour trouver le niveau d'équilibre \bar{s} , on utilise ensuite le fait que l'espérance des variations de niveau autour de \bar{s} doit être nulle :

$$\begin{aligned}\mathbb{E}[\Delta s^+ + \Delta s^- - q] &= 0 \\ \mathbb{E}[\Delta s^+] + \mathbb{E}[\Delta s^-] - \mathbb{E}[q] &= 0\end{aligned}$$

$$(1-\tilde{s}^2) \cdot \mathbb{P}(P>E) \cdot \mathbb{E} \left[\frac{P-E}{\theta_1} \middle| P>E \right] - \tilde{s}(2-\tilde{s}) \cdot \mathbb{P}(E>P) \cdot \mathbb{E} \left[\frac{E-P}{\theta_1} \middle| E>P \right] - \tilde{s} \left(1 - \left[1 + \left(\frac{\tilde{s}}{2.25} \right)^4 \right]^{-\frac{1}{4}} \right) = 0$$

$$(1-\tilde{s}^2) \cdot \underbrace{\mathbb{P}(P>E) \cdot \mathbb{E} \left[\frac{P-E}{\theta_1} \middle| P>E \right]}_{\pi_P} - \tilde{s}(2-\tilde{s}) \cdot \underbrace{\mathbb{P}(E>P) \cdot \mathbb{E} \left[\frac{E-P}{\theta_1} \middle| E>P \right]}_{\pi_E} - \theta_1 \tilde{s} \left(1 - \left[1 + \left(\frac{\tilde{s}}{2.25} \right)^4 \right]^{-\frac{1}{4}} \right) = 0$$

$$(1-\tilde{s}^2) \pi_P - \tilde{s}(2-\tilde{s}) \pi_E - \theta_1 \tilde{s} \left(1 - \left[1 + \left(\frac{\tilde{s}}{2.25} \right)^4 \right]^{-\frac{1}{4}} \right) = f(\tilde{s}) = 0 \quad (10.5)$$

Les termes π_P et π_E (espérances des pluies et ETP nettes) ne dépendent que des caractéristiques des forçages climatiques, ils sont donc connus et ne dépendent pas du jeu de paramètres testé pendant l'optimisation. L'Équation (10.5) est non-linéaire en \tilde{s} (f est approximativement un polynôme d'ordre 5) mais on voit que :

$$\begin{cases} f(0) = \pi_P > 0 \\ f(1) \simeq -\pi_E - 0.0095 \cdot X1 < 0 \\ \frac{df}{ds} < 0 \quad \forall s \in [0, 1] \end{cases}$$

D'un point de vue numérique, ces caractéristiques permettent de trouver le niveau \tilde{s} très rapidement par dichotomie (moins d'une dizaine d'évaluations de f suffisent pour déterminer \tilde{s} à 1% près). Dans le cas de la Figure 10.6, le calcul donne $\tilde{s} = 0.342$, ce qui est cohérent avec le comportement observé. Remarquons d'ailleurs que le capacité θ_1 du réservoir de production et son taux de remplissage moyen ne peuvent pas être invariants quand on change de pas de temps : la distribution temporelle (autant que spatiale) de la pluie étant dépendante de l'échelle d'observation, les espérances $\pi_P = \mathbb{P}(P_t > E_t) \cdot \mathbb{E}[P_t - E_t | P_t > E_t]$ et $\pi_E = \mathbb{P}(E_t > P_t) \cdot \mathbb{E}[E_t - P_t | E_t > P_t]$ sont également dépendantes de la résolution temporelle des données et du pas de temps de fonctionnement des modèles. Comme le signalent *Michel et al. (2006)*, cette dépendance est généralement méconnue : « *les modélisateurs présentent généralement leurs modèles comme si leur pas de temps de fonctionnement était juste une question de disponibilité des données, ou une question exclusivement calculatoire. En fait, l'influence du pas de temps sur la structure d'un modèle est bien plus importante que l'influence de la discrétisation spatiale* ».

Cette procédure permet donc, pour chaque jeu de paramètre testé (par exemple à chaque itération de la méthode « pas-à-pas »), d'obtenir très rapidement (i.e. sans faire tourner le modèle) le niveau d'équilibre que *Mouelhi (2003)* obtenait par l'établissement d'un régime permanent. La différence avec la solution de *Mouelhi* réside dans le fait que le calcul n'est

qu'approché, et qu'il sera très faux dans le cas où θ_1 est petit. Cependant, nous avons vu que dans ce cas, la condition initiale est très rapidement oubliée (la période de forçages réels suffisant à cette tâche).

10.4 Proposition et test d'une procédure d'initialisation alternative

10.4.1 Principe

Dans la section précédente nous avons exposé une méthode de calcul d'un état d'équilibre pour le réservoir sol (nous pourrions d'ailleurs effectuer le même raisonnement sur le réservoir de routage). Nous pouvons donc finalement proposer une procédure d'initialisation alternative qui, comme celle de *Mouelhi*, comporte deux étapes :

1. le calcul du niveau d'équilibre \tilde{s} avec les caractéristiques climatiques π_P et π_E de la période, ainsi que la valeur courante θ_1 de la capacité du réservoir sol,
2. l'utilisation d'une ou deux années maximum de forçages réels avant de commencer à calculer la fonction objectif. Si la valeur de θ_1 est grande, l'étape (1) aura déterminé de façon assez précise l'état d'équilibre, que cette période de forçages réels viendra « corriger » en fonction du contexte climatique du début de période ; inversement, si la valeur de θ_1 est faible, le niveau d'équilibre calculé en (1) ne correspondra pas à la valeur moyenne effective mais la mémoire du réservoir sera suffisamment faible pour « oublier » rapidement cet état erroné.

Pour la période de contrôle, deux possibilités s'offrent à nous : soit nous utilisons le paramètre \tilde{s} trouvé pendant la période de calage, soit nous le recalculons avec les valeurs climatiques π_P et π_E de la période de contrôle (mais bien sûr la valeur de θ_1 déterminée en calage).

10.4.2 Test de la procédure proposée

Dans cette section, nous évaluons l'apport de la méthode d'initialisation proposée. Nous testons deux stratégies :

- dans le premier cas (noté B1), nous prenons une valeur unique pour le taux de remplissage moyen des réservoirs en début d'année hydrologique. Nous choisissons la médiane des valeurs calées au test A précédent ($s_0 = 0.41$ et $r_0 = 0.38$ au pas de temps journalier, la médiane des r_0 étant de 0.19 au pas de temps horaire conformément au comportement décrit dans l'Annexe C) ; ensuite nous utilisons une période de n années de forçages réels (« préchauffage ») avant de commencer le calcul de la fonction objectif,

- dans le second cas (noté B2), nous calculons le niveau d'équilibre \bar{s} , que nous utilisons comme état initial pour le réservoir de production ($s_0 = \bar{s}$), et nous gardons $r_0 = 0.41$ comme précédemment ; ensuite nous utilisons comme précédemment une période de n années de forçages réels avant de commencer le calcul de la fonction objectif.

Le but du test est de comparer les performances des deux approches pour différentes valeurs de n (nous testerons 1, 2, 4, 8 et 16 ans), de façon à voir si la nouvelle procédure permet d'économiser des années de « préchauffage ». Quelle que soit la valeur de n , les périodes du split-sample test sont toujours les mêmes : 01/08/1995–31/07/2000 et 01/08/2000–31/07/2005. Ainsi, si on utilise $n = 8$ ans de préchauffage, les forçages réels de la première période vont du 01/08/1987 au 31/07/1995 et ceux de la deuxième période du 01/08/1992 au 31/07/2000, etc.

10.4.2.1 Résultats

Sur la Figure 10.7, nous avons tracé les résultats en contrôle avec la solution B2 ($s_0 = \bar{s}$) en fonction de ceux utilisant la solution B1 ($s_0 = 0.41$), pour différentes valeurs de n (nombre d'années de préchauffage). Sans surprise, lorsque n devient grand les deux approches sont équivalentes puisque le poids de la condition initiale devient nul. Par contre, pour les petites valeurs de n nous constatons qu'un bon nombre de simulations sont améliorées par l'utilisation du niveau initial \bar{s} plutôt que la valeur fixe 0.41.

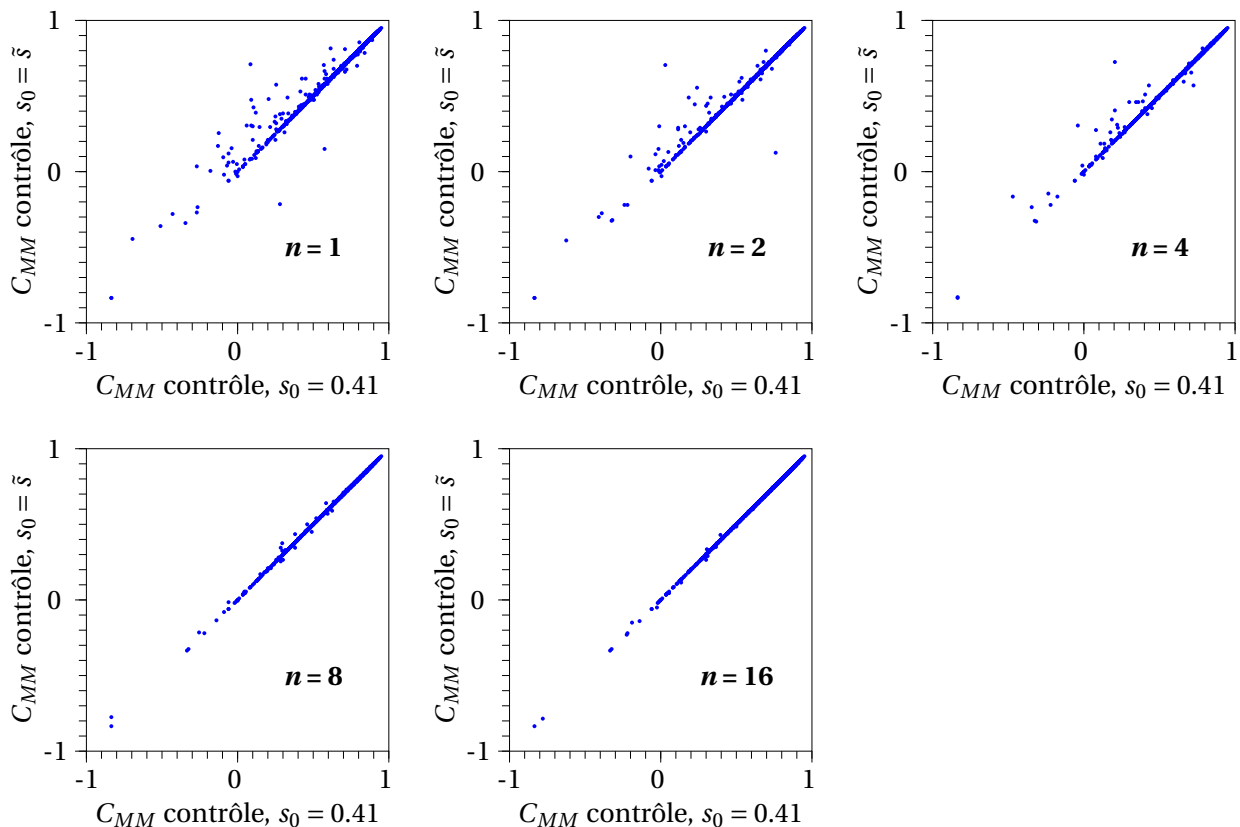


FIG. 10.7: .

10.5 Synthèse

Dans ce chapitre nous avons décrit les problèmes inhérents à l'utilisation de chroniques de données de longueur finie, et à la nécessité d'initialiser les états d'un modèle pluie-débit. Nous avons décrit une procédure permettant d'établir plus rapidement la valeur de régime permanent du réservoir de production, régime dont l'établissement est recommandé par *Mouelhi* (2002). Cette méthode convient tout à fait pour les structures GR4 et GR5, structures dans lesquelles il n'y a pas de composant plus lent que le réservoir sol (le réservoir de routage a en général une capacité nettement plus faible : typiquement quelques dizaines de millimètres, contre plusieurs centaines pour le réservoir de production, et sa loi de vidange est beaucoup plus « nerveuse »). Elle reste en revanche insuffisante lorsque l'on ajoute un réservoir lent, comme dans la structure M_{096} , ou que l'on utilise une autre structure. On peut imaginer le même genre de développements analytiques pour les réservoirs où cela est possible, mais ces développements impliquent de connaître leurs flux d'alimentation moyens. Cela n'est pas simple quand ces flux dépendent d'autres composants « à l'amont » dans le modèle (ce qui n'est pas le cas pour le réservoir sol de GR).

Dans ces cas de figure, l'alternative est d'obtenir, à l'issue d'une première simulation sur l'ensemble des bassins, la distribution des valeurs moyennes des taux de remplissage de chaque réservoir en début d'année. Ces valeurs moyennes peuvent ensuite être utilisées comme point de départ *a priori* de la période de préchauffage, sachant que rien ne remplacera jamais l'utilisation d'une longue période de forçages réels respectant la chronologie... avec ce que cela implique en coût de calcul et en « sacrifice » de données pluviométriques ! Une fois n'est pas coutume, c'est un type de contraintes qui disparaît généralement quand on passe du mode *simulation* au mode *prévision* : la totalité de l'historique est alors utilisable, notamment les précipitations antérieures sur une période de plusieurs années.

11

**Perspectives : Quel(s) critère(s) utiliser
pour sélectionner un modèle ?**

“

The river is as far as I can move from the world of numbers.”

Jim HARRISON — *The Theory & Practice of Rivers.*

11.1 Les stratégies de sélection d'un modèle et leurs limites

Au cours des chapitres précédents, nous avons usé¹ de critères d'évaluation des modèles. Nous avons mis en évidence un effet progressif de « saturation » des performances des modèles en réponse à l'accroissement de la complexité des paramétrisations (par exemple, avec l'ajout de composantes lentes). Ce constat et la question qui en découle (quelle est l'origine de cette saturation ?) suggèrent trois réponses possibles :

1. la première consiste à dire que cette saturation est intrinsèque à la modélisation de la relation pluie-débit : celle-ci ne tolérerait pas plus d'une poignée de paramètres (voir par ex. *Jakeman et Hornberger*, 1993), la quantité d'information « extractible » des séries temporelles pour leur estimation étant très limitée ;
2. les limites reconstruites peuvent aussi tenir à la méthode d'évaluation (*split-sample test*), qui n'est peut-être pas assez discriminante. Les autres volets du test de *Klemeš* (1986) pourraient peut-être engendrer des différences plus marquées entre modèles notamment sur leur capacité d'extrapolation, avec le *differential split-sample test* ;
3. une dernière possibilité est de questionner les critères utilisés (i.e., les « notes » que nous mettons aux modèles) : permettent-ils de détecter les contrastes qui font la différence entre nos modèles ? En effet, même si la quantité d'information contenue dans les chroniques pluie-débit pour l'estimation des paramètres est vraisemblablement limitée, il est également indéniable que les critères couramment utilisés (Nash-Sutcliffe et variantes) sont loin d'épuiser cette information : rien ne garantit donc qu'ils permettent de départager avec pertinence des hypothèses concurrentes de plus ou moins grande complexité. Ainsi, l'examen de la surface de réponse (*via* l'utilisation de matrices d'information par exemple, comme dans *Jakeman et Hornberger*, 1993) n'a de sens que dans la mesure où la fonction objectif est adaptée.

11.1.1 Rôles respectifs de la fonction objectif et du/des critère(s) de contrôle

Rappelons quelques idées générales concernant le développement et la sélection d'un modèle :

- Tous les modèles, qu'ils soient conceptuels ou basés physiquement, doivent être évalués par un critère de contrôle (C.C.) qui peut être adapté à l'objectif de la modélisation. Dans le cas d'un modèle opérationnel, l'utilisateur a la liberté de choix de ce critère. Mais dans notre cas, dans la mesure où le modèle que nous souhaitons sélectionner a l'ambition d'être généraliste, la question du choix du critère de contrôle se pose,

¹Certains diront, peut-être non sans raison, abusé

- Certains modèles (entre autres, les modèles conceptuels, mais en pratique cela touche pratiquement tous les modèles) nécessitent également le choix d'une fonction objectif (F.O.) à minimiser pour identifier le/les jeu(x) de paramètres adéquat(s).

La fonction objectif et le critère de contrôle n'ont pas nécessairement à être identiques : par exemple, si l'on s'intéresse à la prédétermination des crues, on peut tout à fait choisir de caler le vecteur de paramètres θ d'un modèle pluie-débit avec une F.O. continue sur toute la période de calage de longueur N , par exemple

$$FO(\mathbf{q}, \theta) = \sum_{t=1}^N \left(\sqrt{\hat{q}_t(\theta)} - \sqrt{q_t} \right)^2$$

puis, sur une période de contrôle, calculer la moyenne des facteurs d'erreur sur les pics des dix plus fortes crues (repérés à t_1, \dots, t_{10}) :

$$CC(\hat{\theta}) = \frac{1}{10} \sum_{i=1}^{10} \left| \ln(\hat{q}_{t_i}(\hat{\theta})) - \ln(q_{t_i}) \right|$$

où $\hat{\theta} = \underset{\theta}{\operatorname{argmin}}\{FO(\mathbf{q}, \theta)\}$, c'est-à-dire le vecteur de paramètres minimisant la fonction objectif.

Naturellement, les modèles étant très imparfaits, nous avons tout intérêt, à *modèle donné*, à adapter *dans une certaine mesure* la F.O. au critère de contrôle : si nous nous intéressons aux crues, il est probablement plus sage d'estimer les paramètres en calant sur les débits ou la racine carrée des débits, plutôt que sur le logarithme. Cela revient à exploiter inégalement l'information, voire à ignorer une partie, des séries d'erreurs.

Cette adaptation de la F.O. au critère de contrôle (i.e. à l'objectif final de modélisation) peut s'avérer profitable, mais dans certaines limites uniquement. Ainsi, dans le cas ci-dessus, si nous choisissons précisément, comme fonction objectif pour l'identification des paramètres, l'erreur relative moyenne sur les dix plus fortes crues de la période de calage, le résultat serait probablement très mauvais (à cause d'un sur-apprentissage entraînant une perte de robustesse en contrôle).

Cet exemple est bien sûr caricatural. Pourtant, rien ne garantit que la saturation observée sur les performances ne vienne pas d'un effet semblable, quoique moins flagrant. Dès lors, nous pouvons formuler notre problématique : quelle(s) fonction(s) objectif et quel(s) critère(s) de contrôle choisir pour l'amélioration et le choix entre différentes *structures*, sans présager des objectifs potentiels de modélisation sur des gammes particulières de débits (crues, étiages, etc.) ?

11.1.2 Que signifie classer des hypothèses ?

Cette problématique est particulièrement sensible pour la modélisation des influences souterraines sur les débits de surface. En effet, nous avons expliqué au Chapitre 3 que les flux entre hydrosystèmes superficiels et souterrains pouvaient affecter toute la gamme des débits, aussi bien les crues que les étiages. Partant de ce constat, il n'est souhaitable de privilégier aucune gamme de débit particulière pour classer nos hypothèses : sans abandonner toute vision utilitariste, il semble raisonnable de vouloir conserver une certaine polyvalence au modèle, ambition d'ailleurs implicite dans le choix du mode continu (d'autres approches sont envisageables, telle que le mode évènementiel).

Nous avons donc besoin d'un critère opérationnel permettant de « noter » des hypothèses en prenant en compte la totalité des observations. Examinons par exemple la proposition de *Popper* (1954 ; 1957 ; 1958 ; 1983) pour caractériser le degré auquel des observations e (*evidence*) corroborent une hypothèse h , étant donnée une connaissance de base b (*background knowledge*) :

$$C_P(h, e, b) = \frac{\mathbb{P}(e|hb) - \mathbb{P}(e|b)}{\mathbb{P}(e|hb) - \mathbb{P}(eh|b) + \mathbb{P}(e|b)} \quad (11.1)$$

Nous voyons intervenir plusieurs termes dans ce critère :

- $\mathbb{P}(e|b)$, la probabilité d'observer e étant donnée la connaissance de base b ,
- $\mathbb{P}(e|hb)$, la probabilité d'observer e étant donnée la conjonction de la connaissance de base b et de l'hypothèse h ,
- enfin, un terme $\mathbb{P}(eh|b)$, qui est la probabilité de la conjonction de e et h conditionnellement à b .

Le dénominateur de l'Équation 11.1 est choisi pour normaliser le critère, de sorte que :

- $-1 \leq C_P(h, e, b) \leq 1$,
- $C_P(h, e, b)$ approche de la valeur maximale 1 si $\mathbb{P}(e|hb) \simeq 1$ et $\mathbb{P}(e|b) \simeq 0$: les observations ne peuvent corroborer fortement l'hypothèse h que si elles sont très improbables sous la connaissance de base seule,
- si e sont incompatibles avec l'hypothèse h , c'est à-dire si h implique \bar{e} , alors l'observation de e ruine totalement l'hypothèse h : $C_P(h, e, b) = -1$,
- si e n'est pas discriminant pour l'hypothèse testée (c'est-à-dire si le test n'est pas exigeant) ou que sa probabilité absolue est élevée, alors e ne fait ni corroborer ni discréditer h : $C_P(h, e, b) = 0$.

Dans le cas de variables continues, nous pouvons garder l'esprit de ce critère en remplaçant les probabilités conditionnelles $\mathbb{P}(e|.)$ par les vraisemblances $\mathcal{L}(e|.)$. D'autre part, si l'hypothèse h prétend au statut de loi universelle (i.e., h est une hypothèse anticipatrice

excessivement téméraire), nous pouvons à coup sûr écrire que sa probabilité logique est égale à zéro et que $\mathbb{P}(eh|b) = \mathbb{P}(h|b) = \mathbb{P}(h) = 0$. Le critère de corroboration devient alors :

$$C_P(h, e, b) = \frac{\mathcal{L}(e|hb) - \mathcal{L}(e|b)}{\mathcal{L}(e|hb) + \mathcal{L}(e|b)} \quad (11.2)$$

Dans ses travaux, Popper ne prétend pas construire un critère opératoire : son but est simplement de montrer qu'un critère de corroboration ne peut pas correspondre à la probabilité logique de l'hypothèse, et il est très sceptique sur l'utilisation pratique de sa définition. Pourtant, nous pouvons montrer que le critère de Nash (ou plus exactement une « anamorphose » de ce critère) peut tout à fait être vu comme un critère de corroboration dans lequel la connaissance de base b est le modèle naïf $Q = \bar{Q}$ (noté $b = m_0$), l'hypothèse h le modèle pluie-débit à tester ($bh = m_1$), et dans lequel nous supposons que les erreurs sont indépendantes et normalement distribuées (i.n.d.). Avec ces nouvelles notations, le critère de corroboration s'écrit :

$$C_P(m_1, e, m_0) = \frac{\mathcal{L}(e|m_1) - \mathcal{L}(e|m_0)}{\mathcal{L}(e|m_1) + \mathcal{L}(e|m_0)} \quad (11.3)$$

$$C_P(m_1, e, m_0) = \frac{1 - \frac{\mathcal{L}(e|m_0)}{\mathcal{L}(e|m_1)}}{1 + \frac{\mathcal{L}(e|m_0)}{\mathcal{L}(e|m_1)}} \quad (11.4)$$

Remplaçons les vraisemblances par leur expression dans le cas d'erreurs i.n.d. :

$$\mathcal{L}(e|m_0) = \left(\frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_0} \right)^N \exp\left(-\frac{N}{2}\right) \quad (11.5)$$

$$\mathcal{L}(e|m_1) = \left(\frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_1} \right)^N \exp\left(-\frac{N}{2}\right) \quad (11.6)$$

où σ_0^2 et σ_1^2 désignent les variances respectives des erreurs des modèles m_0 et du modèle m_1 .

Le critère devient alors :

$$C_P(m_1, e, m_0) = \frac{1 - \left(\frac{\sigma_1}{\sigma_0}\right)^N}{1 + \left(\frac{\sigma_1}{\sigma_0}\right)^N} \quad (11.7)$$

L'expression $1 - \left(\frac{\sigma_1}{\sigma_0}\right)^N$ au numérateur de l'Équation 11.7 rappelle typiquement le critère de Nash, à la différence que le rapport $\frac{\sigma_1}{\sigma_0}$ est élevé à la puissance N du nombre d'observations au lieu de 2. L'Équation 11.7 complète est quant à elle similaire au critère borné C_{MM} introduit par *Mathevet et al.* (2006), que nous avons présenté au Chapitre 1 (Équation 1.2) et abondamment utilisé tout au long de cette thèse.

11.1.3 Statistique suffisante et vraisemblance formelle

Le critère de Nash (ou toute anamorphose) représenterait donc un critère opérationnel tout à fait satisfaisant pour sélectionner un modèle si la statistique σ_i^2 , variance des erreurs du modèle (i.e., de l'hypothèse) m_i , était effectivement suffisante pour résumer la totalité de l'information apportée par la série d'erreurs. Si les erreurs étaient effectivement indépendantes, identiquement et normalement distribuées, alors la moyenne des erreurs (le biais) et leur variance seraient conjointement suffisantes (Fisher, 1922) : aucune autre statistique calculée d'après la série d'erreurs n'apporterait davantage d'information sur cette série et donc sur la valeur optimale des paramètres.

Le problème est que cette hypothèse concernant la structure des erreurs n'est absolument pas vérifiée, comme le montre l'exemple de la Figure 11.1 : les erreurs sont notablement hétéroscédastiques et autocorrélées, ce qui invalide les Équations 11.5 et 11.6. Cette violation du formalisme n'est pas gênante en elle-même (sauf si l'on veut associer des incertitudes aux simulations, ce qui n'est pas notre souci ici), mais c'est ce qu'elle implique qui pose problème : le critère calculé ne satisfait pas du tout la condition de suffisance, et il devient en toute rigueur nécessaire de lui adjoindre d'autres statistiques pour avoir un résumé correct de la qualité du modèle.

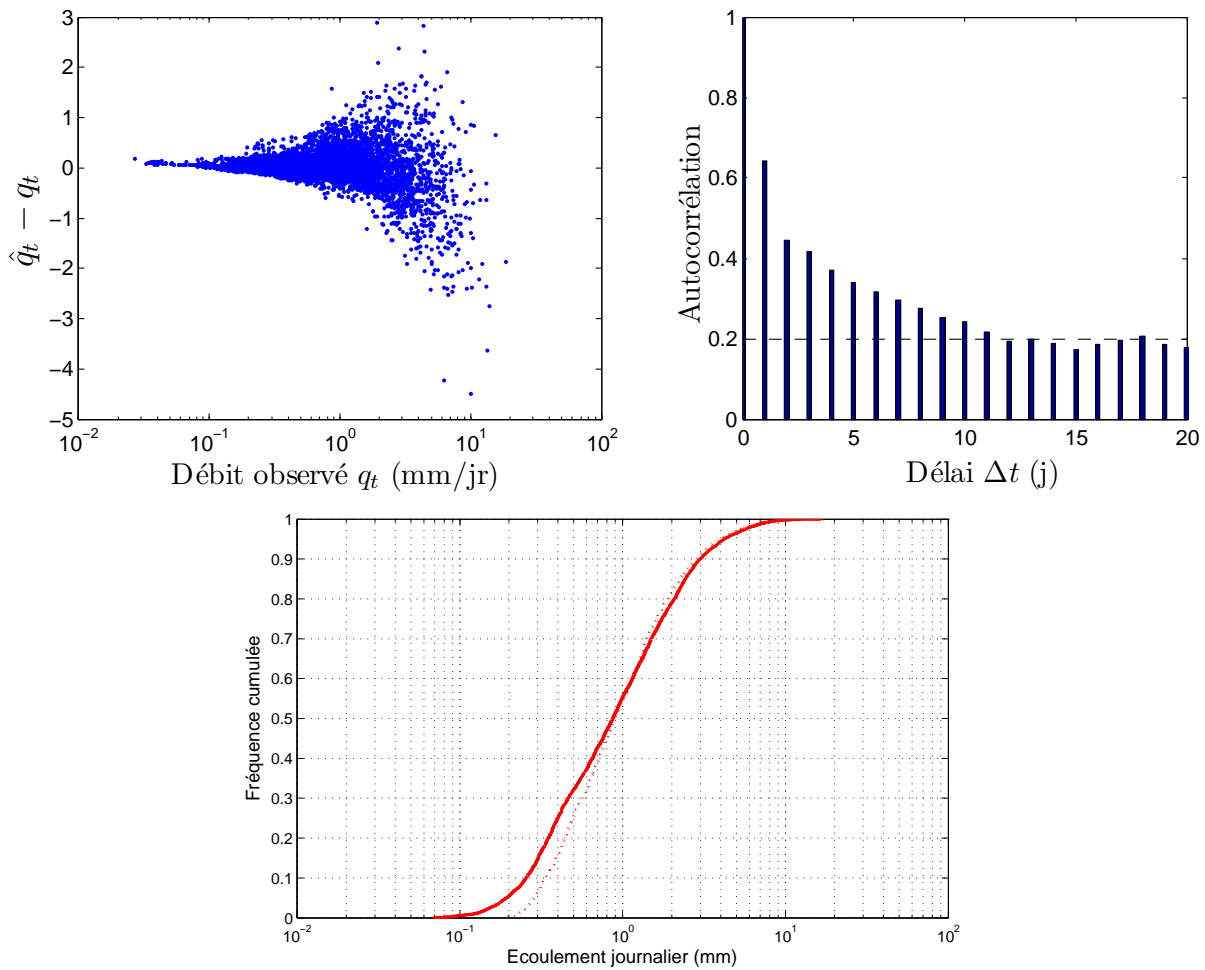


FIG. 11.1: Exemple de structure des erreurs obtenues en minimisant l'erreur quadratique moyenne sur les débits. En haut à gauche : mise en évidence de l'hétéroscédasticité des erreurs (variance nettement plus grande pour les forts débits que pour les faibles) ; En haut à droite : autocorrélogramme des erreurs (autocorrélation non-négligeable jusqu'à $\Delta t = 10$ jours) ; En bas : distribution des débits observés (trait plein) et simulés (pointillé). Le critère de Nash-Sutcliffe correspondant à ces résidus est de 82.6 (bassin versant de la Tardoire à Maisonnais-sur-Tardoire, R1132510).

11.1.3.1 Une famille de critères de contrôle

L'insuffisance du critère basé sur l'écart-type des erreurs ouvre la voie à un risque de sur-apprentissage important. Ce sur-apprentissage peut précisément être illustré par les performances en contrôle des jeux de paramètres identifiés avec une telle fonction objectif. Nous proposons de calculer, pour une fonction objectif donnée, une famille de critères de

contrôle notée $\{C_{MM}(\lambda)\}_{\lambda \in]-\infty, +\infty[}$ et définie par :

$$C_{MM}(\lambda) = \frac{1 - \frac{F(\lambda)}{F_0(\lambda)}}{1 + \frac{F(\lambda)}{F_0(\lambda)}} \quad (11.8)$$

$$\frac{F(\lambda)}{F_0(\lambda)} = \frac{\sum \left((\hat{q}_t(\boldsymbol{\theta}) + \delta)^\lambda - (q_t + \delta)^\lambda \right)^2}{\sum \left(\mathbb{E} \left[(Q_t + \delta)^\lambda \right] - (q_t + \delta)^\lambda \right)^2} \quad (11.9)$$

Il s'agit donc simplement des critères C_{MM} calculés sur une transformation puissance (Box-Cox) des séries de débits observé et simulé (auxquels on ajoute un petit incrément δ , égal par exemple à la précision numérique, de façon à pouvoir prendre en compte les valeurs nulles pour $\lambda \leq 0$). Comme $x^\lambda = \exp\{\lambda \cdot \ln(x)\} \underset{\lambda \rightarrow 0}{\sim} 1 + \lambda \cdot \ln(x)$, les séries transformées sont équivalentes aux logarithmes en $\lambda \rightarrow 0$:

$$\frac{F(\lambda)}{F_0(\lambda)} \underset{\lambda \rightarrow 0}{\sim} \frac{\sum \left(\ln(\hat{q}_t(\boldsymbol{\theta}) + \delta) - \ln(q_t + \delta) \right)^2}{\sum \left(\mathbb{E}[\ln(Q_t + \delta)] - \ln(q_t + \delta) \right)^2} \quad (11.10)$$

L'intérêt d'une telle famille de transformations réside également dans les équivalences suivantes lorsque $\lambda \rightarrow +\infty$ ou $-\infty$. Dans ces cas, un seul terme domine alors chacune des sommes, aux dates notées respectivement t_M et t_m , de sorte que :

$$\left\{ \begin{array}{l} \frac{F(\lambda)}{F_0(\lambda)} \underset{\lambda \rightarrow +\infty}{\sim} \frac{\left((\hat{q}_{t_M}(\boldsymbol{\theta}) + \delta)^\lambda - (q_{t_M} + \delta)^\lambda \right)^2}{\frac{1}{N} (q_{t_M} + \delta)^{2\lambda}} = N \left[\left(\frac{\hat{q}_{t_M}(\boldsymbol{\theta}) + \delta}{q_{t_M} + \delta} \right)^\lambda - 1 \right]^2 \\ \frac{F(\lambda)}{F_0(\lambda)} \underset{\lambda \rightarrow -\infty}{\sim} \frac{\left((\hat{q}_{t_m}(\boldsymbol{\theta}) + \delta)^\lambda - (q_{t_m} + \delta)^\lambda \right)^2}{\frac{1}{N} (q_{t_m} + \delta)^{2\lambda}} = N \left[\left(\frac{\hat{q}_{t_m}(\boldsymbol{\theta}) + \delta}{q_{t_m} + \delta} \right)^\lambda - 1 \right]^2 \end{array} \right.$$

Ainsi, lorsque $\lambda \rightarrow -\infty$, $C_{MM}(\lambda)$ ne reflète plus que les erreurs sur les étiages extrêmes; symétriquement, lorsque $\lambda \rightarrow +\infty$, le critère ne reflète plus que les erreurs sur les plus fortes crues. Ce comportement est illustré sur la Figure 11.2 : pour différentes valeurs de λ , nous avons repéré par des points l'ensemble des dates contribuant pour 80% à la somme $F(\lambda)$. Quelle que soit la valeur de λ , le critère conserve donc un sens hydrologique (i.e., il reflète les erreurs sur une gamme particulière de débits, même si nous ne pouvons définir simplement cette gamme que pour les valeurs extrêmes de λ). Les valeurs de λ faibles et fortes renseignent donc probablement sur les capacités d'extrapolation des modèles, en crue et en étiage.

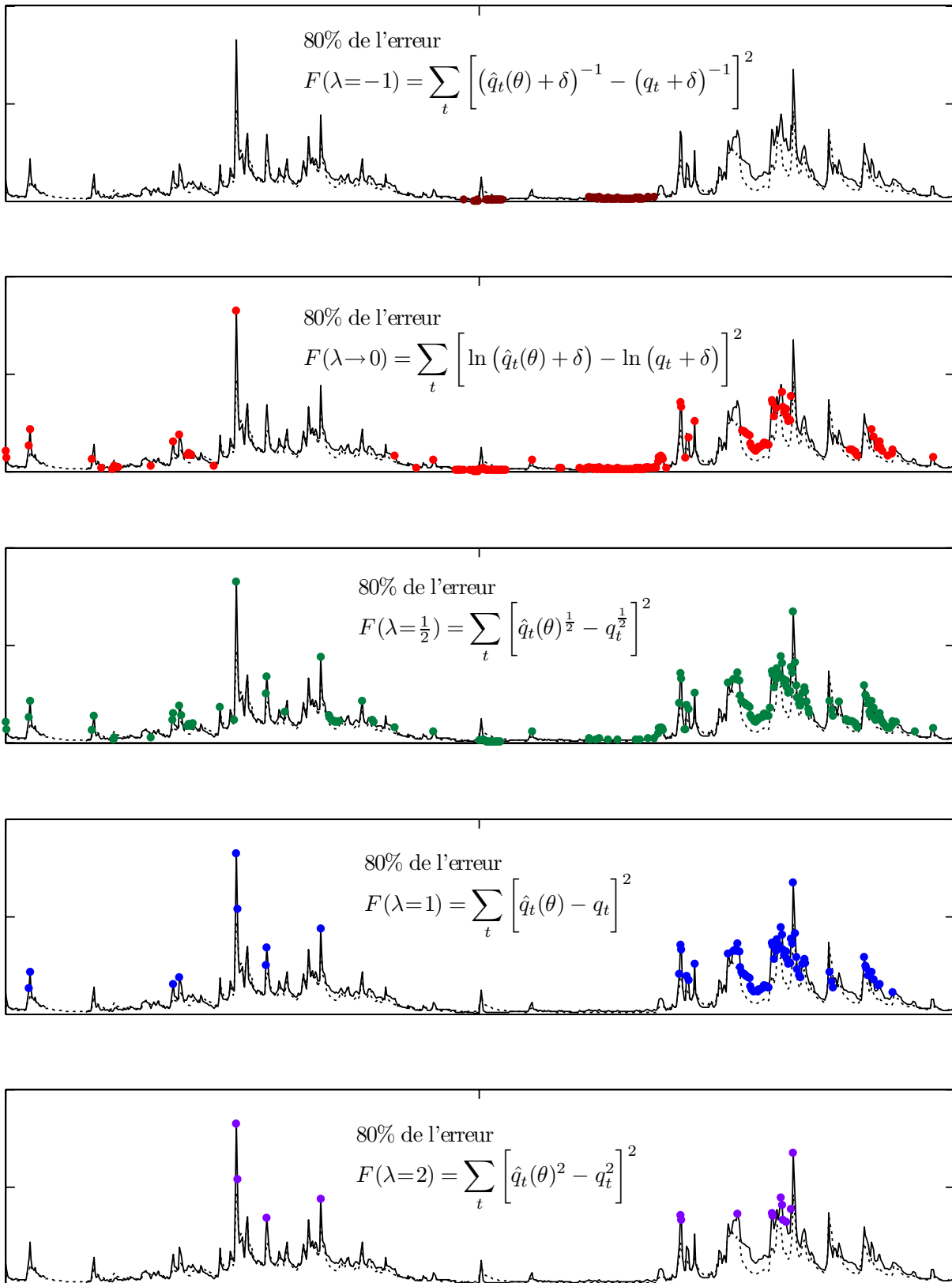


FIG. 11.2: Exemple de localisation des points représentant 80% du terme d'erreur $F(\lambda) = \sum \left[(\hat{q}_t(\theta) + \delta)^\lambda - (q_t + \delta)^\lambda \right]^2$ sur une période de contrôle de deux ans, pour différentes valeurs de λ ($-1, 0, \frac{1}{2}, 1$ et 2). Le débit observé est en trait plein, le débit simulé en pointillé (bassin versant R1132510).

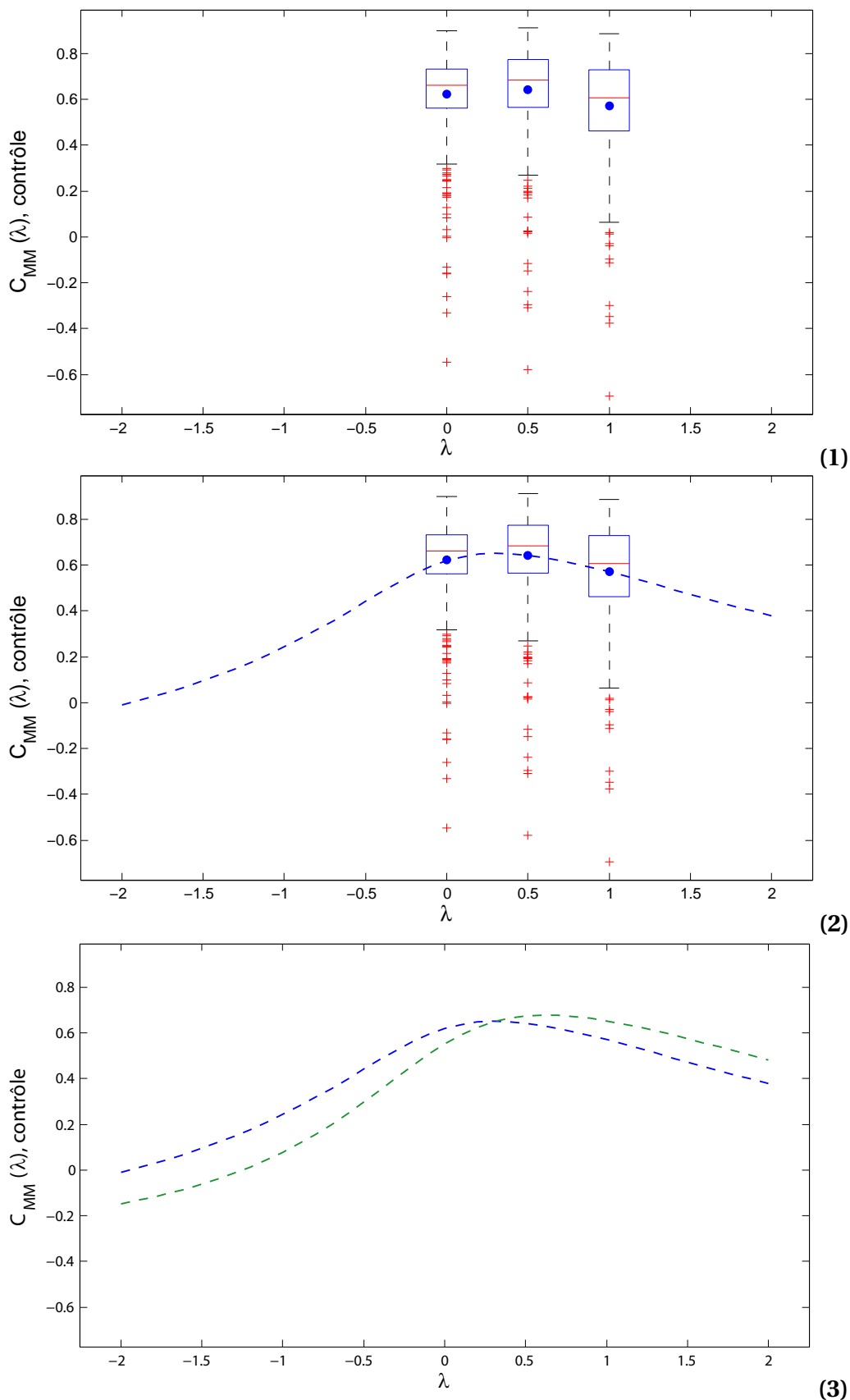


FIG. 11.3: Résumé des performances en contrôle d'un couple modèle+fonction objectif. (1) Calcul de la distribution des critères $C_{MM}(\lambda)$ pour différentes valeurs de λ ; (2) Résumé de chaque distribution par sa moyenne, i.e. construction d'une courbe $\lambda \mapsto \overline{C_{MM}(\lambda)}$; (3) Comparaison de ces courbes lorsque l'on change de fonction objectif ou de modèle.

11.1.3.2 Représentation graphique

La Figure 11.3 illustre la façon de résumer les performances d'un modèle en contrôle. Pour différentes valeurs de λ , nous calculons la distribution (représentée sous forme de boîte à moustaches) des critères $C_{MM}(\lambda)$ en contrôle sur les bassins de l'échantillon. Si on choisit de résumer chacune de ces boîtes par sa moyenne, on obtient une courbe $\lambda \mapsto \overline{C_{MM}}(\lambda)$. Cette courbe illustre la performance d'un modèle vue sous tous les angles hydrologiques possibles, depuis le point de vue « *seuls les étiages extrêmes m'importent* » (pour $\lambda \rightarrow -\infty$) jusqu'au point de vue « *seules les crues extrêmes m'importent* » (pour $\lambda \rightarrow +\infty$). Nous pouvons alors observer comment se déforme cette courbe lorsque l'on change de modèle ou de fonction objectif pour identifier les paramètres au calage. Cette représentation est donc une visualisation « continue » d'une sorte d'évaluation multi-critères.

11.1.3.3 Illustrations pour quatre fonctions objectif

Dans ce paragraphe nous illustrons la méthode précédente pour quatre fonctions objectif :

- l'erreur quadratique moyenne sur le logarithme des débits (i.e., $FO = C_{MM}(\lambda = 0)$),
- l'erreur quadratique moyenne sur la racine carrée des débits (i.e., $FO = C_{MM}(\lambda = 0.5)$),
- l'erreur quadratique moyenne sur le débit (i.e., $FO = C_{MM}(\lambda = 1)$),
- l'erreur quadratique moyenne sur le carré des débits (i.e., $FO = C_{MM}(\lambda = 2)$)

Nous avons utilisé successivement ces fonctions objectif pour optimiser les paramètres du modèle GR4J, et tracé les courbes $\overline{C_{MM}}(\lambda)$ correspondantes sur la Figure 11.4, en calage et en contrôle. Plusieurs phénomènes intéressants sont à noter :

- en calage (graphique du haut), il existe toujours une plage de valeurs de λ sur laquelle chaque fonction objectif est supérieure aux autres, ce qui est logique :
 - Pour la fonction objectif « erreur quadratique sur $\ln(Q)$ », cette plage est approximativement $]-\infty; 0.3]$,
 - Pour la fonction objectif « erreur quadratique sur \sqrt{Q} » : $[0.3; 0.8]$
 - Pour la fonction objectif « erreur quadratique sur Q » : $[0.8; 1.5]$
 - Pour la fonction objectif « erreur quadratique sur Q^2 » : $[1.5; +\infty[$
- en contrôle (graphique du bas), cela n'est plus vérifié : notamment, les jeux de paramètres sélectionnés par la F.O. « erreur quadratique sur Q^2 » donnent, en contrôle, des performances inférieures ou égales à celles des jeux sélectionnés par la F.O. « erreur quadratique sur Q », quelque soit la valeur de λ (en particulier, à $\lambda = 2$!).

Ces résultats n'ont rien de surprenant. En effet, les propriétés du $C_{MM}(\lambda)$ en calage (comme F.O.) sont les mêmes qu'en contrôle : plus $|\lambda|$ est grand, plus est faible le nombre de points ayant une contribution significative à la somme des erreurs, ce qui entraîne une perte de robustesse très importante. Sur la Figure 11.5, nous avons tracé les courbes $\overline{C_{MM}}(\lambda)$ en

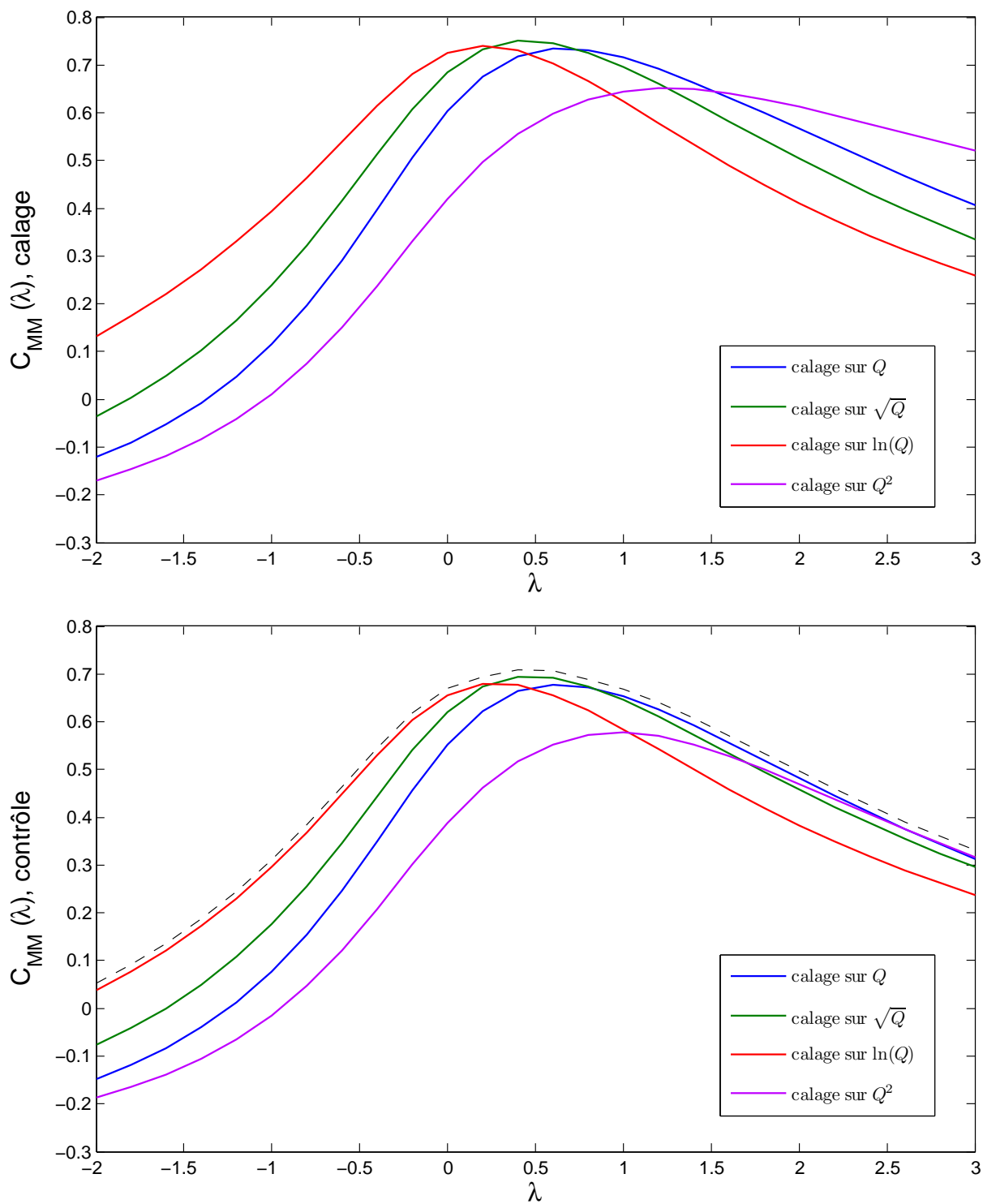


FIG. 11.4: Dépendance des performances en contrôle à la fonction objectif utilisée en calage : illustration pour le modèle GR4J.

contrôle pour deux modèles (GR4J et GR5J) et deux fonctions objectif. Nous constatons que le modèle à cinq paramètres donne des résultats en contrôle égaux ou moins bons que le modèle à quatre paramètres, sur toutes les valeurs de λ , lorsque l'on utilise la fonction objectif « erreur quadratique sur Q^2 ». Au contraire, les performances sont améliorées (à la fois en valeur absolue et par rapport au modèle à quatre paramètres) sur toutes les valeurs de λ lorsque l'on utilise la F.O. « erreur quadratique sur \sqrt{Q} ». En se basant sur la vision parcelaire des erreurs du modèle que nous donne la F.O. « erreur quadratique sur Q^2 », nous pourrions donc conclure que le modèle à cinq paramètres est déjà sur-paramétrisé. Se pourrait-il que nous courrions le même type de risque en utilisant d'autres critères, apparemment plus raisonnables (le cas Q^2 étant un peu extrême), mais ne vérifiant pas non plus (ou pas assez) la condition de suffisance? Cette possibilité n'est absolument pas exclue : **cet exemple démontre clairement que le classement des structures en contrôle peut dépendre, dans une mesure que nous ignorons, de la fonction objectif utilisée pour identifier les paramètres au calage.**

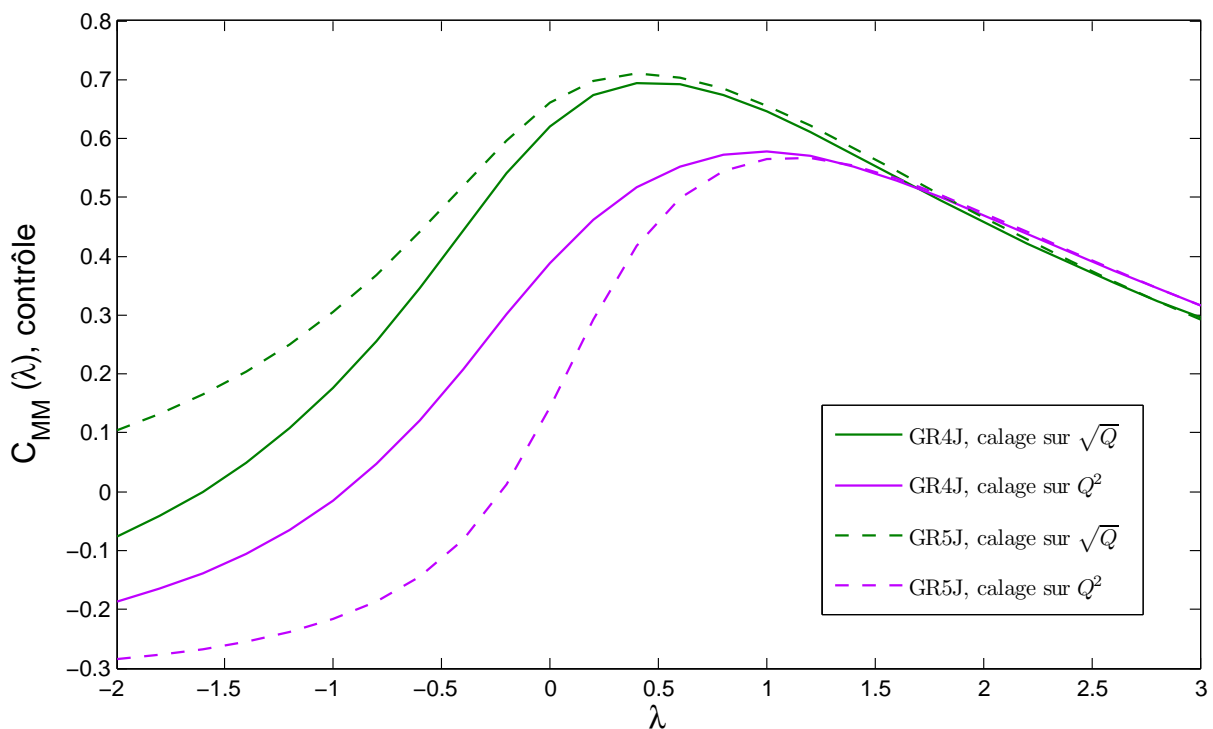


FIG. 11.5: Effets d'une F.O. inadaptée sur l'amélioration ou la dégradation apparente des performances : illustration pour les modèles GR4J et GR5J, calés avec les F.O. « erreur quadratique sur \sqrt{Q} » ou « erreur quadratique sur Q^2 ».

Idéalement, notre but serait d'identifier une fonction objectif produisant, en contrôle, la courbe pointillée de la Figure 11.4, c'est-à-dire une fonction objectif extrayant plus d'information de la série d'erreurs — toutes gammes de débit confondues — que n'importe quelle autre F.O. testée : la représentation graphique proposée est une façon empirique

d'évaluer la suffisance de la statistique utilisée comme fonction objectif. Il est important de tracer ces courbes en contrôle, car c'est uniquement par une validation croisée que l'on peut faire la part entre l'information *effectivement* extraite, et la « modélisation du bruit » (surcalage) : cette vérification ne peut être qu'empirique.

11.2 Tests empiriques de vraisemblances formelles

11.2.1 Principe général

Nous avons vu au paragraphe 11.1.2 qu'une définition raisonnable de la corroboration d'une hypothèse par les observations faisait intervenir une vraisemblance formelle. Le problème — clairement mentionné par Popper comme un obstacle opératoire majeur — est que nous ne connaissons pas la structure des erreurs : nous ne pouvons donc ni déduire a priori d'expression formelle de la vraisemblance, ni par conséquent identifier les statistiques conjointement suffisantes pour son estimation. Il nous faut donc avoir recours à une procédure empirique, qui se résume aux points suivants :

- La définition d'une série temporelle $g_t(\mathbf{q}, \boldsymbol{\theta}, \mathbf{x})$, fonction des observations \mathbf{q} , des paramètres $\boldsymbol{\theta}$ et éventuellement de variables exogènes \mathbf{x} . Cette fonction est choisie de telle sorte que $\mathbf{g} = \{g_t\}$ forme une série de valeurs *indépendantes et identiquement distribuées*, de densité f exprimable analytiquement,
- L'expression de la vraisemblance à partir de la densité conjointe des g_t et des termes jacobiens $\left| \frac{\partial g_t}{\partial q_t} \Big|_{\mathbf{q}, \boldsymbol{\theta}, \mathbf{x}} \right|$:

$$f_G(g_1, \dots, g_N) = \prod_t f(g_t) \quad (11.11)$$

$$\mathcal{L}(\mathbf{q}|\boldsymbol{\theta}, \mathbf{x}) = f_Q(q_1, \dots, q_N) = \prod_t \left| \frac{\partial g_t}{\partial q_t} \Big|_{\mathbf{q}, \boldsymbol{\theta}, \mathbf{x}} \right| \cdot f(g_t) \quad (11.12)$$

11.2.2 Exemple issu de Yang et al., 2007

Nous illustrons la démarche précédente avec la formulation proposée par Yang et al. (2007). Ces auteurs proposent l'hypothèse d'erreurs multiplicatives sur les débits (donc additives sur le logarithme) et autocorrélées, modélisables par un processus AR(1) u_t :

$$\begin{cases} \ln(\hat{q}_t(\boldsymbol{\theta})) = \ln(q_t) + u_t \\ u_t = \rho u_{t-1} + \epsilon_t, \quad \epsilon_t \rightsquigarrow \mathcal{N}(0, \sigma_\epsilon) \end{cases}$$

La série \mathbf{g} choisie est alors simplement la série $\boldsymbol{\epsilon}$ des innovations :

$$\begin{aligned} g_t(\mathbf{q}, \boldsymbol{\theta}) &= \epsilon_t \\ &= \ln(\hat{q}_t(\boldsymbol{\theta})) - \ln(q_t) - \rho \left\{ \ln(\hat{q}_{t-1}(\boldsymbol{\theta})) - \ln(q_{t-1}) \right\} \end{aligned}$$

Les ϵ_t sont supposés indépendants et identiquement distribués, de densité :

$$f(\epsilon_t) = \frac{1}{\sqrt{2\pi}\sigma_\epsilon} \exp\left(-\frac{\epsilon_t^2}{2\sigma_\epsilon^2}\right)$$

Le jacobien de la transformation se réduit, pour tout t , à :

$$\left| \frac{\partial g_t}{\partial q_t} \Big|_{\mathbf{q}, \boldsymbol{\theta}} \right| = \frac{1}{q_t}$$

L'expression de la vraisemblance se réduit donc à :

$$\mathcal{L}(\mathbf{q}|\boldsymbol{\theta}) = \prod_t \frac{1}{q_t} \cdot \frac{1}{\sqrt{2\pi}\sigma_\epsilon} \exp\left(-\frac{\epsilon_t^2}{2\sigma_\epsilon^2}\right)$$

Nous en déduisons l'expression de la log-vraisemblance :

$$\ell(\mathbf{q}|\boldsymbol{\theta}) = \sum_t \left(-\ln(q_t) - \frac{1}{2} \ln(2\pi) - \ln(\sigma_\epsilon) - \frac{\epsilon_t^2}{2\sigma_\epsilon^2} \right)$$

En posant $(\sum_{t=1}^N \epsilon_t^2) = N\sigma_\epsilon^2$, nous obtenons finalement l'expression de la log-vraisemblance concentrée, dans laquelle nous regroupons tous les termes ne dépendant pas du modèle et de ses paramètres :

$$\ell(\mathbf{q}|\boldsymbol{\theta}) = C(\mathbf{q}) - \frac{N}{2} \ln\left(\sum_t \epsilon_t^2\right)$$

D'après ce modèle, la statistique suffisante (et donc la meilleure fonction objectif) est la somme des carrés des innovations : $FO = \sum_{t=1}^N \epsilon_t^2$. Nous avons donc testé cette fonction objectif comme premier exemple de vraisemblance formelle. En pratique, nous devons estimer le coefficient d'autocorrélation ρ des erreurs u_t , ce que nous avons effectué par un processus itératif :

- Optimisation du vecteur de paramètres $\hat{\boldsymbol{\theta}}$ à $\hat{\rho}$ fixé,
- Calcul de la nouvelle valeur de $\hat{\rho}$ d'après la série des $u_t = \ln(\hat{q}_t(\hat{\boldsymbol{\theta}})) - \ln(q_t)$, par

$$\hat{\rho} = \frac{\mathbb{E}[u_t u_{t+1}]}{\mathbb{E}[u_t^2]}$$

- Nouvelle optimisation de $\hat{\theta}$, et ainsi de suite. On arrête la procédure lorsque $\hat{\rho}$ ne varie plus (à un seuil près).

Les performances en contrôle (courbe $\overline{C_{MM}}(\lambda)$) du modèle GR4J calibré avec cette fonction objectif sont illustrées sur la Figure 11.6, où nous avons également fait figurer les courbes des fonctions objectif sur Q , \sqrt{Q} et $\ln(Q)$.

Nous constatons donc que les jeux de paramètres obtenus par la vraisemblance formelle de *Yang et al.* ne conduisent pas à une amélioration des performances en contrôle. Comme les auteurs le signalent (bien qu'ils n'utilisent qu'un seul bassin versant pour leur étude et qu'ils se placent dans un cadre bayésien), c'est encore une fois la vérification des hypothèses distributionnelles qui pose problème. L'hypothèse d'erreurs multiplicatives semble relativement bien vérifiée (les $\{u_t\}$ sont asymptotiquement normalement distribués) ; en revanche, l'hypothèse d'un coefficient d'autocorrélation de ces $\{u_t\}$ invariant dans le temps ne tient pas (p. 175) : les erreurs sont bien plus autocorrélées en récession et en étiage (périodes sèches) que lors des montées de crues (périodes pluvieuses). Cet effet n'est pas flagrant sur l'autocorrélogramme des innovations calculées avec un ρ unique : le coefficient à d'ordre 1 est effectivement très proche de zéro mais cela masque très certainement une sur-différentiation des erreurs en période de crue, et une sous-différentiation en récession. Il paraît donc prudent, lors des tests de vraisemblance formelle, de 1) certes vérifier les hypothèses distributionnelles, mais plus encore de vérifier que les performances en contrôle sont 2) améliorées sur certaines gammes de débits (i.e. certaines valeurs de λ), et 3) pas ou pas trop dégradées sur les autres gammes.

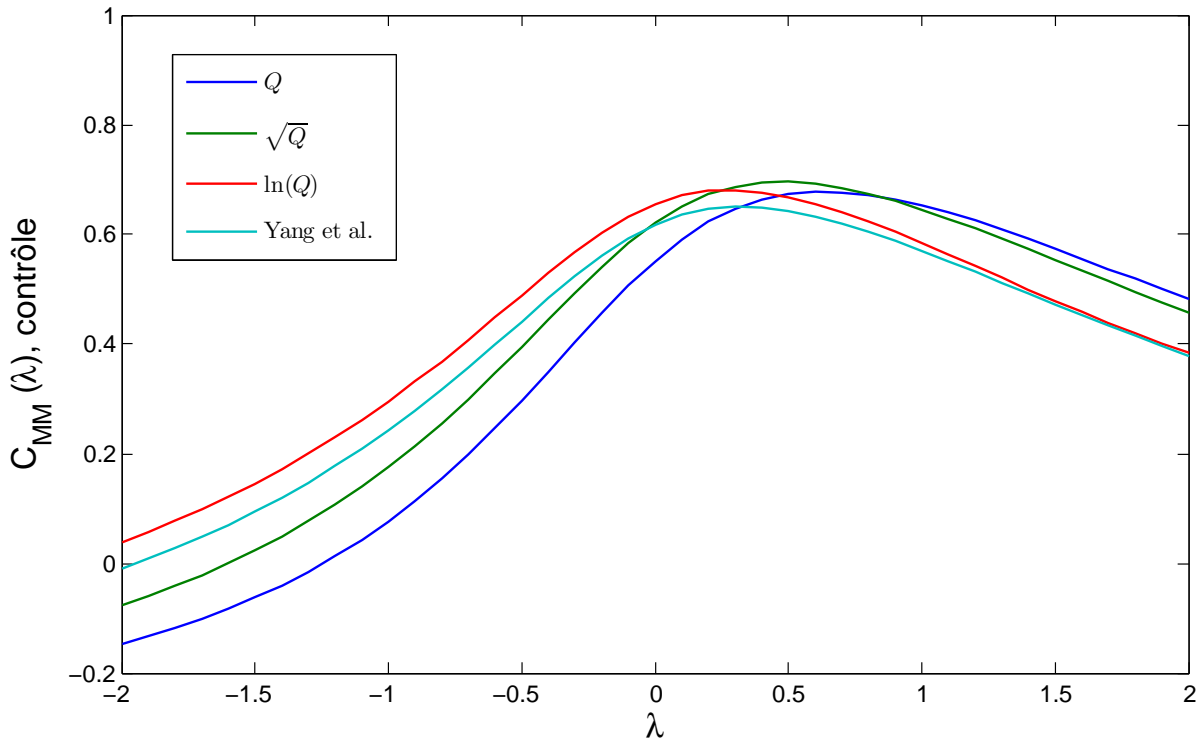


FIG. 11.6: Courbes $\overline{C_{MM}}(\lambda)$ obtenues en contrôle en utilisant quatre fonctions objectif différentes en calage : erreur quadratique moyenne sur Q , \sqrt{Q} ou $\ln Q$, et modèle d'erreur de *Yang et al.* (2007).

11.2.3 Formulations testées

Ayant accepté ce principe empirique, nous pouvons alors tester différentes hypothèses concernant la forme de la vraisemblance. Étant donnée l'importance que semble prendre la pluie p_t (ou la pluie nette $p_t - e_t$) dans l'autocorrélation des erreurs, nous proposons de chercher g_t sous la forme :

$$g_t(\mathbf{q}, \boldsymbol{\theta}, \mathbf{p}, \mathbf{e}) = \epsilon_t = T(\hat{q}_t(\boldsymbol{\theta})) - T(q_t) - \rho(q_{t-1}, p_t, e_t) \cdot \left\{ T(\hat{q}_{t-1}(\boldsymbol{\theta})) - T(q_{t-1}) \right\}$$

On peut voir une telle fonction comme une variante de la fonction HMLE (*Heteroscedastic Maximum Likelihood Estimator*, voir e.g. *Sorooshian and Dracup*, 1980). La fonction T est choisie parmi l'une des deux suivantes :

- logarithme
- transformation par quantiles normaux (NQT, *Kelly et Krzysztofowicz*, 1997), qui permet de mieux traiter les débits d'étiage (voir l'Annexe E pour une description complète de cette transformation)

Pour la fonction $\rho(q_{t-1}, p_t, e_t)$, nous testons des expressions de la forme :

$$\left\{ \begin{array}{l} \rho(q_{t-1}, p_t) = \left(1 - F_Q(q_{t-1})\right)^\alpha \left(1 - F_P(p_t)\right)^\beta \\ \text{ou} \\ \rho(q_{t-1}, p_t, e_t) = \left(1 - F_Q(q_{t-1})\right)^\alpha \left(1 - F_{P-E}(p_t - e_t)\right)^\beta \end{array} \right.$$

où $F_X(x)$ est la fonction de répartition de la variable X . $\left(1 - F_X(x)\right)$ est donc la probabilité au non-dépassement de x , d'où le comportement supposé de la fonction d'autocorrélation :

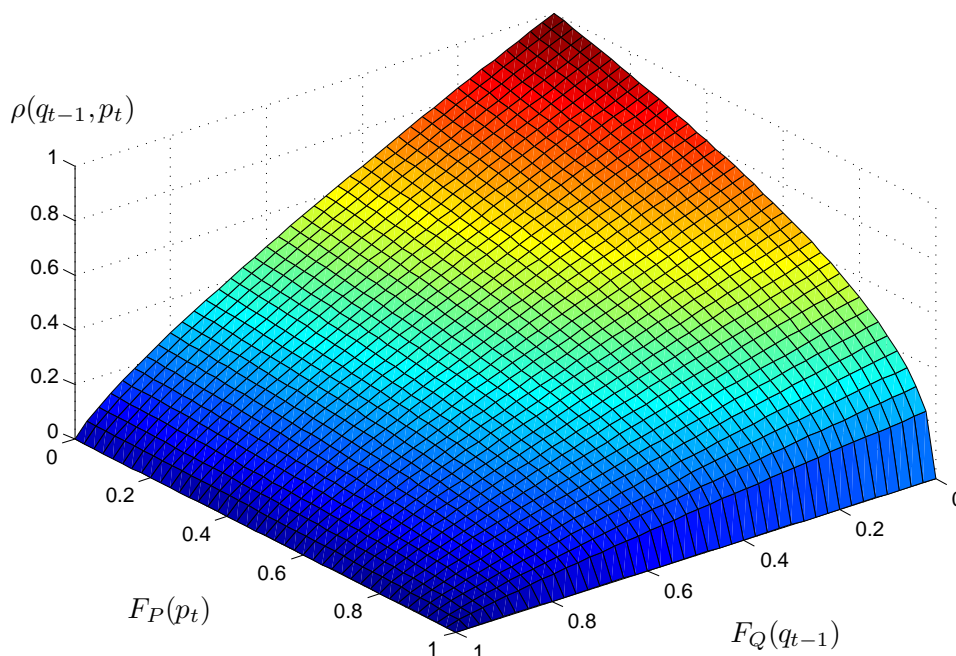


FIG. 11.7: Hypothèse sur la forme du coefficient d'autocorrélation en fonction de la pluie p_t au pas de temps t , et du débit précédent q_{t-1} : exemple pour $\alpha = 0.8$, $\beta = 0.4$.

Plus le cumul de pluie p_t (ou de pluie nette $p_t - e_t$) est rare, plus le coefficient d'autocorrélation est supposé faible ; de même, plus le débit précédent q_{t-1} est rare, plus on réduit le coefficient d'autocorrélation. Avec cette forme, l'autocorrélation est minimale pendant les montées de crues (fortes pluies et forts débits), et maximale pendant les fins de récession (pas de pluie, débit faible). Elle permet d'augmenter le poids des erreurs commises lorsque le débit est très faible, mais qu'il pleut (cf. Figure 11.8), et inversement de réduire le poids des erreurs commises lorsque le débit est fort mais que la pluie a cessé (i.e., les décrues). En résumé, nous essayons de définir (toujours empiriquement) les innovations en fonction de la sollicitation du bassin : intuitivement, une erreur apporte davantage d'information pour

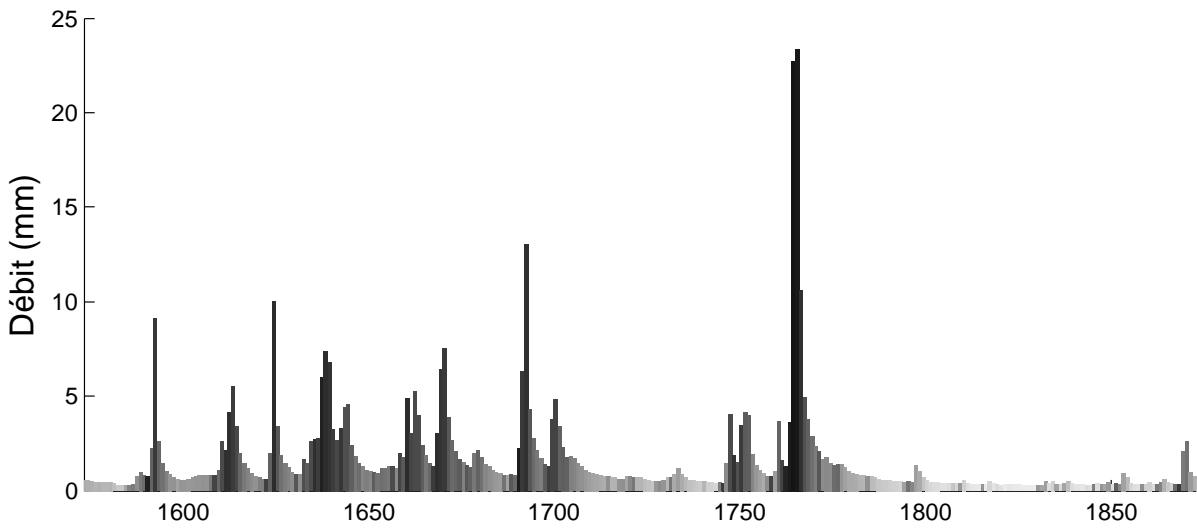


FIG. 11.8: Exemple de variation au cours du temps de la fonction $\rho(q_{t-1}, p_t, e_t)$. Plus l'intensité de gris est forte et plus on suppose l'autocorrélation faible (et donc un poids fort pour l'erreur).

la détermination des paramètres lorsque le bassin est sollicité que lorsqu'il ne l'est pas.

Comme dans la formulation de *Yang et al.* (2007), les termes jacobiens, qui valent

$$\left| \frac{\partial g_t}{\partial q_t} \Big|_{q, \theta} \right| = \left| \frac{\partial T}{\partial q_t} \Big|_q \right|$$

sont indépendants du modèle et du jeu de paramètres θ (que T soit le logarithme ou la transformation par quantiles normaux). Ils peuvent donc être « sortis » de la log-vraisemblance concentrée, de sorte que la fonction objectif se réduit, comme précédemment, à la somme $\sum_{t=1}^N e_t^2$.

11.2.4 Résultats des tests effectués

11.2.4.1 Résultats au pas de temps journalier

Dans cette section nous présentons les résultats des tests de 39 fonctions objectif de la forme définie précédemment.

La Figure 11.9 montre les courbes $\overline{C_{MM}}(\lambda)$ obtenues en contrôle en calant le modèle journalier GR4J avec chacune de ces fonctions objectif. Pour comparaison, nous rappelons les courbes obtenues en utilisant les trois F.O. « classiques » sur les débits (bleu), la racine carrée des débits (vert), et le logarithme des débits (rouge).

Nous constatons que la famille de F.O. définies empiriquement donne des résultats intéressants :

- Sur les forts débits ($\lambda \gg 0$), le faisceau de courbes obtenu se situe légèrement en-dessous de la courbe verte (calage sur \sqrt{Q}), mais assez nettement au dessus de la courbe rouge (calage sur $\ln Q$)
- Sur les faibles débits ($\lambda \ll 0$), le faisceau se situe nettement au-dessus de de la courbe verte, et nombre de courbes se situent même au-dessus de la courbe rouge (calage sur $\ln Q$)

Cette famille de fonctions objectif semble donc permettre d'exploiter davantage d'information des erreurs sur les bas débits, sans que les paramètres identifiés diminuent notablement les performances sur les forts débits.

La même procédure appliquée au modèle à cinq paramètres (échanges linéaires avec seuil) donnent des résultats similaires même si l'amélioration de la simulation des débits d'étiage est moindre (cf. Figure 11.10). Toutefois, il faut reconnaître qu'aucune de ces fonctions n'est vraiment satisfaisante, aucune ne se distinguant réellement. Une procédure de choix possible serait, pour chaque valeur de p , de classer les fonctions objectif et de prendre celle qui a le plus grand rang minimum sur toute la gamme de p . Cette procédure amène à choisir la F.O. 17 (la même pour GR4J et GR5J, ce qui est encourageant) et, en deuxième choix la F.O. 26. Les transformations correspondantes sont :

$$\begin{cases} g_{17}[t] = \text{NQT}(\hat{q}_t) - \text{NQT}(q_t) - \left(1 - F_Q(q_{t-1})\right)^{0.4} \left(1 - F_P(p_t)\right)^{0.4} \left\{ \text{NQT}(\hat{q}_{t-1}) - \text{NQT}(q_{t-1}) \right\} \\ g_{26}[t] = \text{NQT}(\hat{q}_t) - \text{NQT}(q_t) - \left(1 - F_Q(q_{t-1})\right)^{0.25} \left(1 - F_{P-E}(p_t - e_t)\right)^{0.25} \left\{ \text{NQT}(\hat{q}_{t-1}) - \text{NQT}(q_{t-1}) \right\} \end{cases}$$

où NQT désigne la transformation en quantiles normaux.

Sur la Figure 11.11, on a représenté les améliorations « apparentes » que l'on observe quand on passe du modèle GR4J au modèle GR5J. Le changement de modèle apparaît raisonnablement profitable que l'on ait calé avec la fonction objectif sur \sqrt{Q} (graphique du haut) ou avec la F.O. 17 (graphique du bas). Cependant au Chapitre 8, nous nous sommes basés exclusivement sur la différence de +1.3 pt observée à $\lambda = 0.5$, critère dont la subjectivité est manifeste à la lecture de ce graphique. Au contraire, la F.O. 17 donne des améliorations plus uniformes : +3.0 pt en moyenne sur toutes les valeurs de λ (et tous les bassins, puisque chaque courbe est une moyenne sur l'échantillon), ce qui donne peut-être davantage de sens à la phrase « *les gains de performances sont de +3.0 pt en moyenne* ». Cependant il faut reconnaître que la fonction objectif sur \sqrt{Q} reste en l'état actuel le choix le plus simple et le plus polyvalent, exception faite du calcul d'incertitudes. En effet, la Figure 11.12 montre la structure des résidus pour différentes fonctions objectif (erreur quadratique moyenne

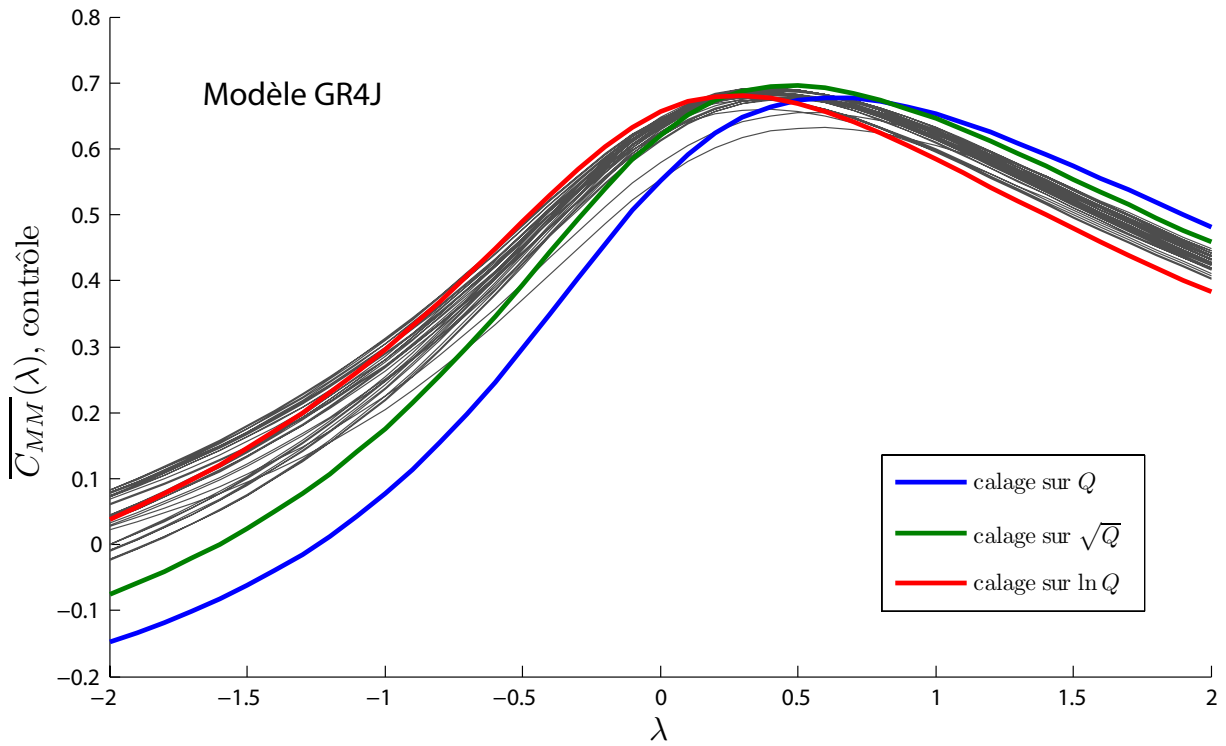


FIG. 11.9: Courbes $\overline{C_{MM}}(\lambda)$ en contrôle sur l'échantillon de bassins versants, obtenues avec le modèle journalier à quatre paramètres.

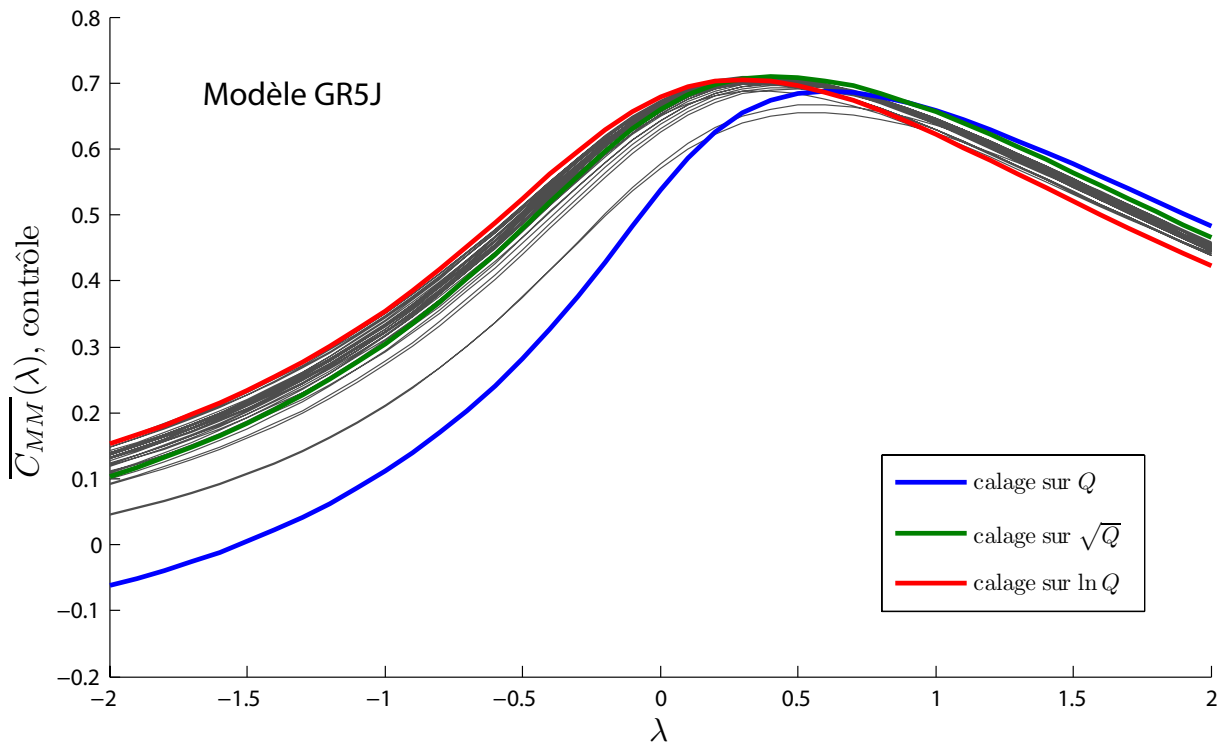


FIG. 11.10: Courbes $\overline{C_{MM}}(\lambda)$ en contrôle sur l'échantillon de bassins versants, obtenues avec le modèle journalier à cinq paramètres.

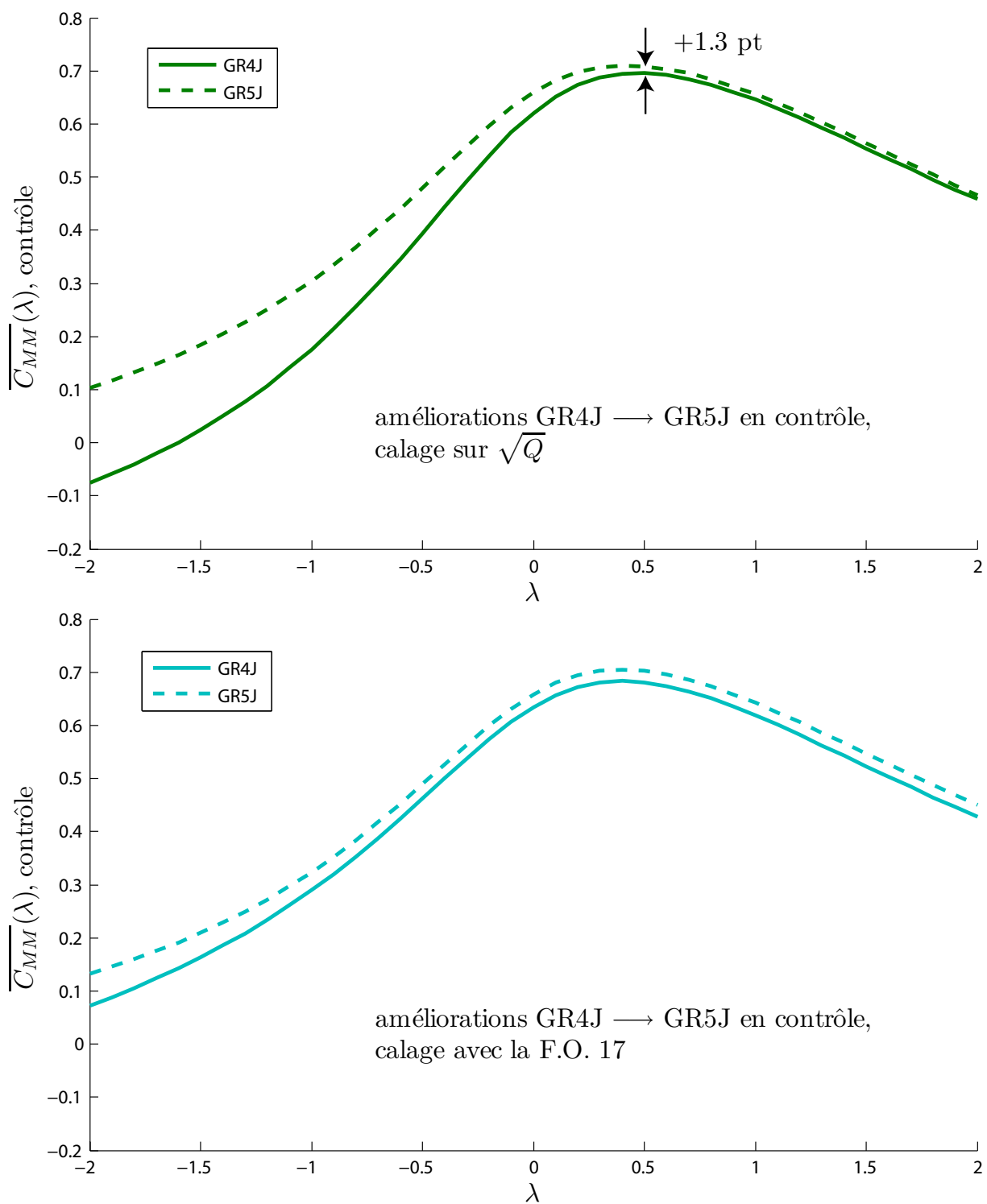


FIG. 11.11: Gains de performance apparents entre les deux modèles GR4J et GR5J, dépendant de la fonction objectif utilisée.

sur Q , \sqrt{Q} , $\ln Q$ et la F.O. 17) sur un bassin versant. Nous constatons en particulier que l'homoscédasticité et l'indépendance des termes transformés est bien mieux vérifiée pour la F.O. 17 que pour les résidus sur les racines.

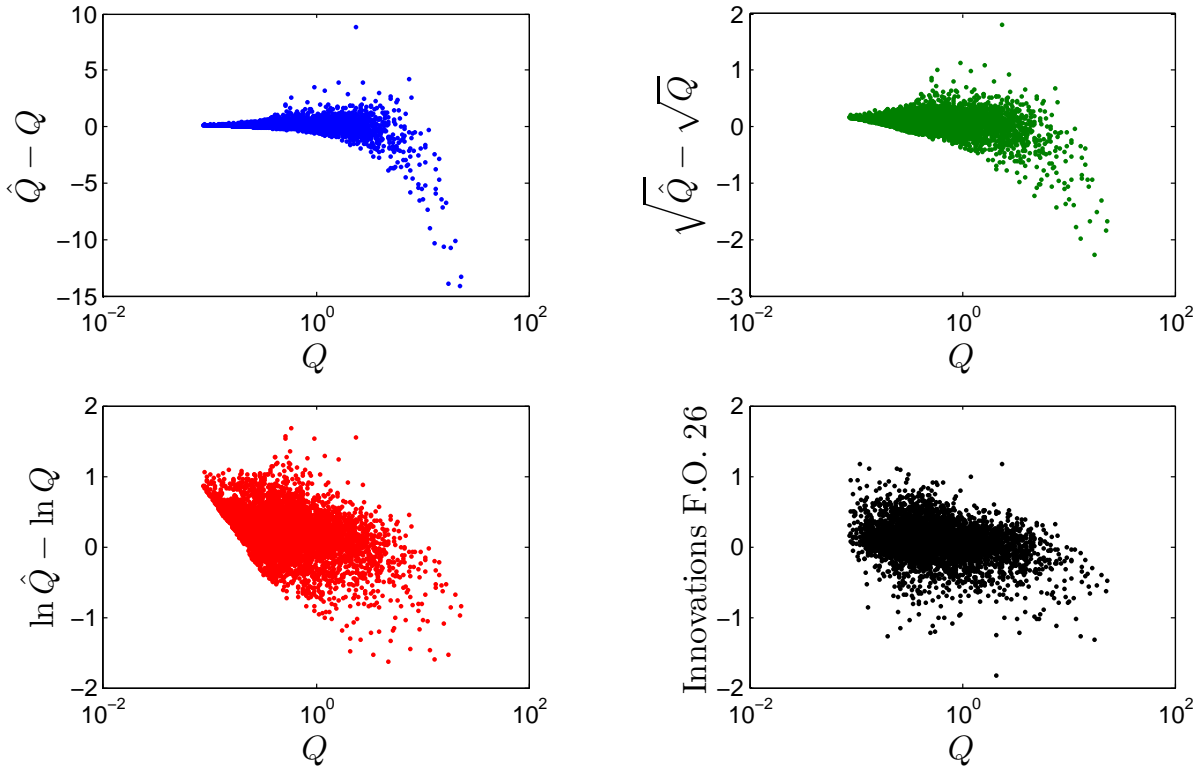


FIG. 11.12: Structure des résidus pour quatre fonctions objectif, au pas de temps journalier (bassin versant de l'Ill à Altkirch, A1050310, modèle GR5).

11.2.4.2 Résultats au pas de temps horaire

Étant donnée la durée des calculs au pas de temps horaire, nous n'avons pas testé les 39 fonctions objectif à ce pas de temps. Nous avons testé celles qui semblaient les plus intéressantes d'après les résultats obtenus au pas de temps journalier, à savoir les F.O. 17 et 26 (en plus bien sûr des fonctions sur Q , \sqrt{Q} et $\ln Q$). Les Figures 11.13 et 11.14 présentent les résultats obtenus, pour le modèle à quatre paramètres (Figure 11.13) et pour le modèle à cinq paramètres (Figure 11.14). Comme pour le pas de temps journalier, aucune de ces fonctions ne donne pleinement satisfaction. Comme dans le cas journalier également, les améliorations sont plus uniforme en utilisant l'une de ces deux F.O. plutôt que l'erreur quadratique moyenne sur \sqrt{Q} (Figure 11.15). Notons qu'il faudrait adapter un peu la forme de la fonction objectif pour le pas de temps horaire : en effet, si l'on fait varier l'autocorrélation des erreurs en fonction de la fréquence de l'évènement de pluie, on aurait peut-être intérêt à tenir compte d'un délai de propagation, en utilisant par exemple une

fonction de transfert h rudimentaire et en calculant la fréquence de $(h \star p)[t]$ plutôt que p_t . Mais tant de complication n'en vaut peut-être pas la peine !

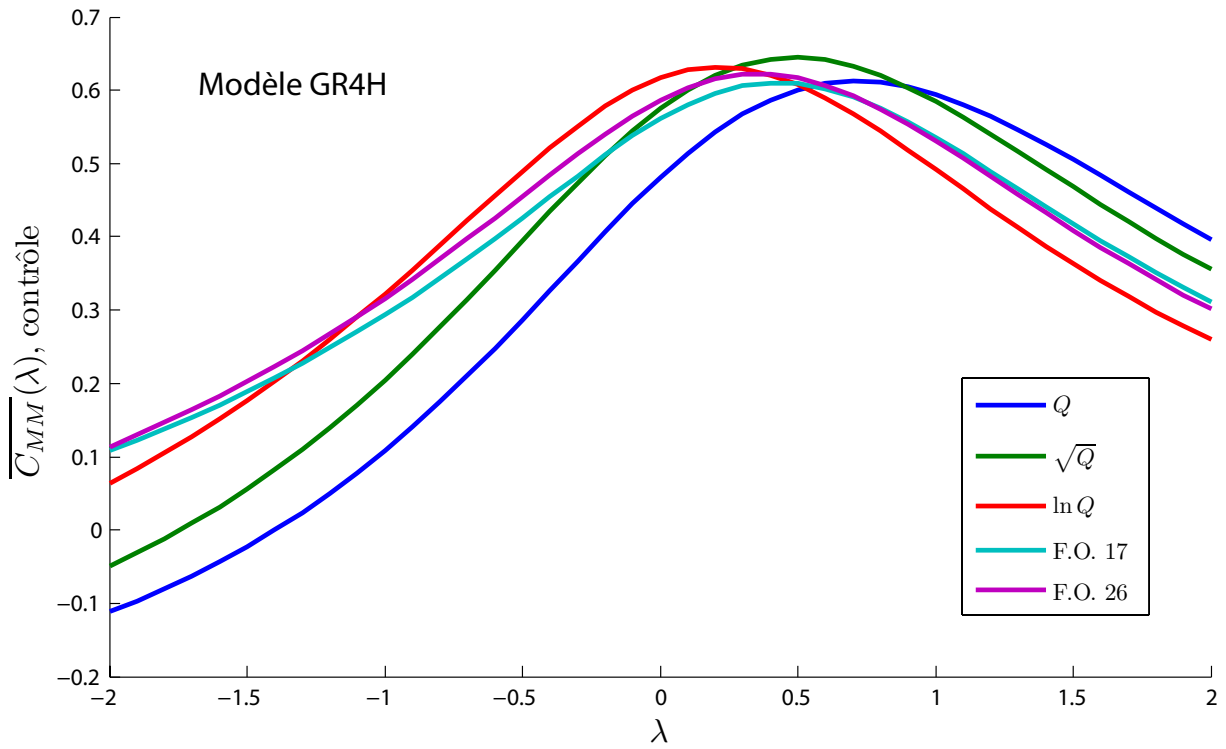


FIG. 11.13: Courbes $\overline{C_{MM}}(\lambda)$ en contrôle sur l'échantillon de bassins versants, obtenues avec le modèle horaire à quatre paramètres.

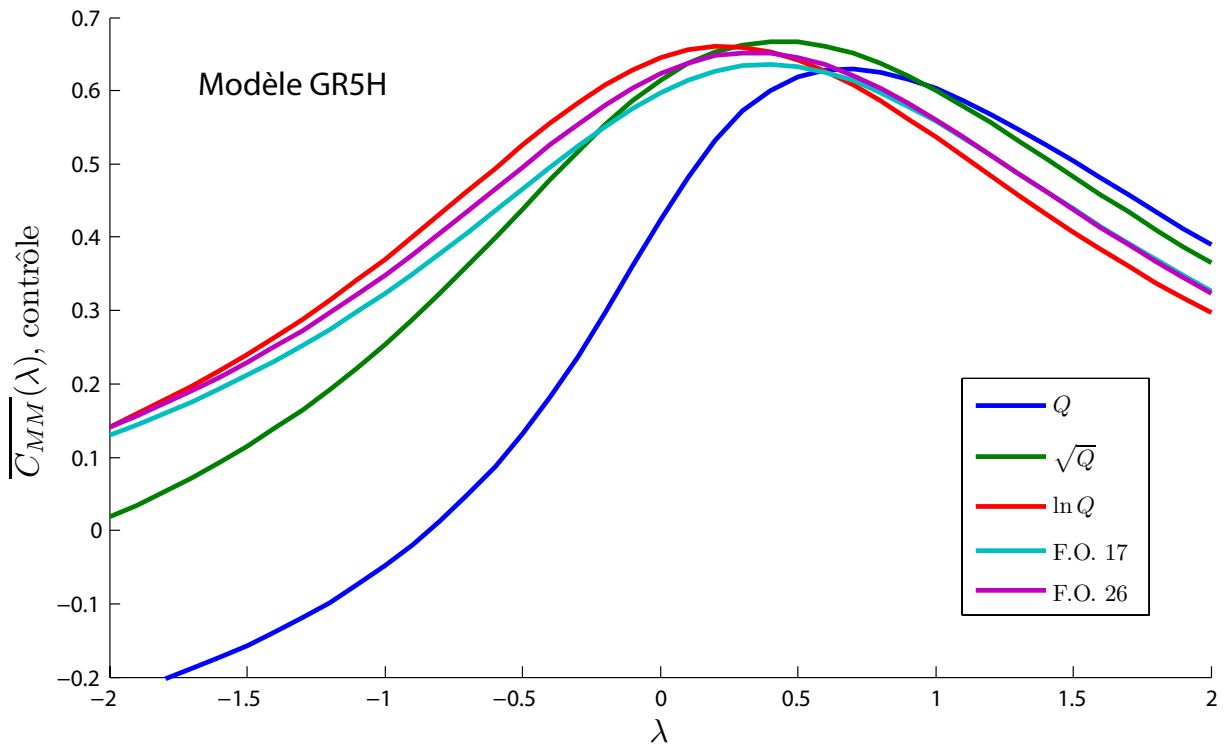


FIG. 11.14: Courbes $\overline{C_{MM}}(\lambda)$ en contrôle sur l'échantillon de bassins versants, obtenues avec le modèle horaire à cinq paramètres.

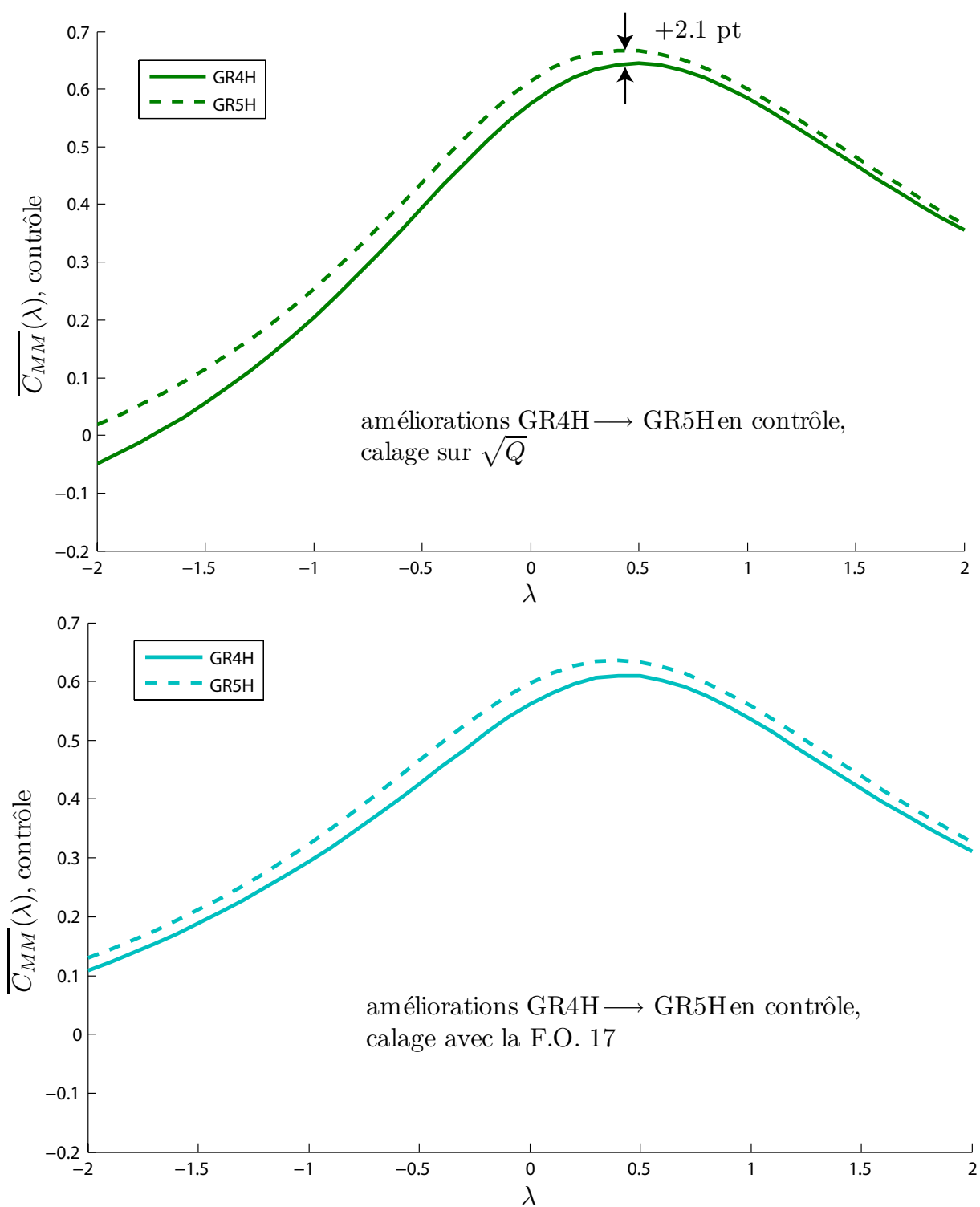


FIG. 11.15: Gains de performance apparents entre les deux modèles GR4H et GR5H, dépendant de la fonction objectif utilisée.

11.3 Synthèse

Dans ce dernier chapitre, nous avons souhaité ouvrir une discussion sur les méthodes d'évaluation des modèles. En effet, si l'approche conceptuelle recherche le pouvoir prédictif plutôt que la profondeur explicative (*Beven, 2001*), alors la mesure de la qualité des modèles (et donc leur qualité relative) est conditionnée au fait de disposer d'une mesure adéquate de ce pouvoir prédictif.

Ce problème est souvent abordé du point de vue du calcul d'incertitudes, et plus rarement du point de vue de la sélection des modèles les uns par rapport aux autres. Dans les deux cas, la recherche d'une fonction de vraisemblance formelle semble théoriquement être la solution, mais cette recherche se heurte à d'immenses difficultés et soulève plus de problèmes qu'elle n'en résout.

S'il est une conclusion ferme que l'on puisse tirer de ces tests, c'est une incitation à la prudence dans l'interprétation des résultats, dans un sens comme dans l'autre. Les tests empiriques sur de grands échantillons de bassins versants indiquent incontestablement que certaines structures sont à abandonner (*Michel et al, 2006*) ; en revanche, il apparaît extrêmement difficile de trancher dans le cas typique d'une « stagnation » des performances : les modifications et complexifications apportées sont-elles vraiment inutiles, ou bien ne savons-nous simplement pas les *mesurer*? Progresser dans ce domaine est, à notre avis, l'une des priorités de la modélisation pluie-débit. En effet, ne pas disposer d'un modèle très satisfaisant n'a rien d'inquiétant tant que l'on dispose des outils et des méthodes pour le faire progresser... Mais est-on seulement sûr de disposer de tels outils, et de telles méthodes ?

12

Conclusion générale

“ Ring the bell that still can ring
Forget your perfect offering
There is a crack in everything
That’s how the light gets in ”

Leonard COHEN — *Anthem*.

Ces dernières années ont vu naître, dans la communauté des hydrologues « opérationnels », une prise de conscience du lien extrêmement étroit pouvant exister entre les eaux souterraines et les eaux de surface. Cette thèse se proposait de jeter un pont entre ces deux rives de l'hydrologie qui, comme le remarquait *Llamas* (1978), ont trop longtemps coexisté sans véritablement dialoguer. Il ne s'agit bien évidemment pas de la première étude sur ce sujet. Cependant, nous avons choisi de l'appréhender au niveau conceptuel du bassin versant.

Dans la première partie, nous avons rappelé les principes définissant cette approche conceptuelle. Nous avons insisté sur la nécessité de construire des représentations de la relation pluie-débit susceptibles d'être testées — et réfutées — dans un grand nombre de cas. À la base de cette thèse, il y avait donc une nécessité méthodologique liée à l'adoption de cette approche empirique : celle de se placer dans un cadre comparatif similaire à celui initié par *Edijatno* (1991) et systématisé par *Perrin* (2000). Nous avons donc constitué un large échantillon de bassins versants permettant la comparaison de modèles aux pas de temps journalier et horaire. À l'intérieur de ce cadre, nous avons cherché à expliciter et questionner les hypothèses souvent tacites utilisées en modélisation pluie-débit.

La première de ces hypothèses est que la majeure partie des flux entrant et sortant du système « bassin versant » sont échangés avec l'atmosphère : les précipitations, l'évapotranspiration et le débit suffisant à fermer les bilans de masse et d'énergie à l'échelle de ce système. Faire cette hypothèse, ce n'est pas forcément nier l'existence d'écoulements souterrains, mais c'est supposer qu'ils se limitent à des systèmes d'écoulement locaux, dont les limites coïncident plus ou moins avec les bassins topographiques. Au Chapitre 3, nous avons montré qu'il n'y a avait aucune raison théorique pour qu'il en soit systématiquement ainsi. Nous avons alors redéfini plus précisément notre système d'étude, le bassin versant topographique considéré comme un système ouvert sur l'atmosphère, mais aussi sur d'autres hydrosystèmes souterrains.

Peut-on évaluer le réalisme des flux simulés ?

La deuxième partie a tenté de répondre à la question suivante : l'introduction d'une telle fonction d'échange est-elle *réaliste* ? Nous avons dû convenir d'une définition de ce que nous appelions *réalisme* : il ne s'agit pas d'un réalisme naïf, qui chercherait par exemple à associer chacun des réservoirs du modèle à un compartiment du bassin (sol, zone saturée, etc.). Le réalisme que nous visons est bien plus distant et bien moins ambitieux : il s'agit simplement de savoir, aussi quantitativement que possible, si les millimètres échangés par le modèle ont bien une contrepartie quelque part sur le bassin réel (peu importe où) : cette vérification ne

va pas de soi, étant donné le caractère non-mesurable de ces flux.

Au Chapitre 4, nous avons montré que diverses solutions alternatives, consistant à corriger les forçages, produisaient des paramètres manifestement incompatibles avec l'idée d'une correction d'erreurs aléatoires, et que de tels paramètres possédaient un déterminisme géologique.

Le Chapitre 5 a été l'occasion de confronter des résultats de mesure (les débits d'une résurgence collectant les pertes d'un karst binaire) avec des simulations (les pertes des bassins topographiques constituant la zone de recharge de l'aquifère karstique). Le test s'est avéré positif, mais reste ce qu'il est : un cas unique, ne permettant aucune « vérification ».

Les tests de semi-distribution présentés au Chapitre 6 ont constitué une tentative plus systématique pour sortir de l'impasse entre un raisonnement général par l'absurde (sans quantification possible), et une collection de cas particuliers. Nous avons utilisé le découpage des bassins versants en sous-bassins non comme une fin en soi (nous avons vu qu'un tel découpage n'est pas une voie directe d'amélioration des simulation du débit à l'exutoire aval), mais comme un moyen de comparer les bilans en les calculant sur différentes partitions d'un même système. Quoique limités à deux modèles, les résultats sont encourageants : le critère de cohérence que nous avons défini s'est avéré sensible à la formulation des échanges, et peut être utilisé comme un complément aux critères classiques. Ces conclusions restent très liées aux choix de modélisation des bassins intermédiaires, mais les travaux de *Lerat* (thèse en cours) semblent offrir de nombreuses perspectives.

Nous avons malgré tout conclu cette première partie en exprimant notre scepticisme sur la possibilité de « valider » des flux non-observables, et réaffirmé la priorité des tests de *pouvoir prédictif* (i.e., de calage-contrôle sur les débits) par rapport à l'évaluation de la cohérence interne des modèles.

Vers une meilleure représentation des étiages ?

Dans la deuxième partie, nous nous sommes donc concentrés sur l'amélioration de ce pouvoir prédictif en testant systématiquement différentes structures empiriques de modèles. Les résultats obtenus ont tout d'abord confirmé la logique régionale de la sensibilité à la formulation des échanges. Dans la continuité des travaux de *Mathevet* (2005), ils nous ont permis de retenir une structure à cinq paramètres, l'adoption de cette structure (baptisée GR5J ou GR5H) résultant d'un arbitrage entre :

- des avantages objectifs :
 - [+] l'amélioration de la simulation des étiages : à la fois pour le *standard split sample test* et le *differential split sample test* en journalier (Chapitre 8), et pour le *standard split sample test* en horaire (le seul utilisé au Chapitre 9) ;
 - [+] de meilleurs critères de cohérence du bilan (semi-distribution, Chapitre 6) ;
 - [+] une continuité entre le pas de temps journalier et le pas de temps horaire, ne nécessitant que l'adaptation de la constante de percolation (d'un facteur conforme à celui prévu par un calcul analytique) et l'exposant de forme de l'hydrogramme unitaire. Cette continuité existait cependant déjà pour la structure GR4H (Mathevet, 2005) ;

- des inconvénients ou, au minimum, des réserves :
 - [-] des performances qui progressent peu en crue, en contrepartie de l'introduction d'un paramètre supplémentaire ;
 - [-] une supériorité non démontrée dans le *proxy basin (differential) split sample test*, même pour les étiages.

Tous ces éléments autorisent raisonnablement à préférer la structure à cinq paramètres. Cependant, pour être déclarée satisfaisante, elle devra encore passer par l'épreuve d'applications opérationnelles telles que la prévision d'étiage. Les résultats du Chapitre 11 nous incitent en effet à penser que la sélection d'un modèle est en fin de compte indissociable d'un objectif de modélisation. Des tests exigeants sur de grands échantillons de bassins versants climatiquement variés sont une étape indispensable à la sélection de modèles robustes ; malgré tout, la sanction d'années de mise en oeuvre pratique avec une grande variété d'objectifs de gestion reste le verdict ultime.

Quelques perspectives

Que reste-t-il à faire ? Formule presque comique tant est longue la liste des questions en suspens...

L'une des questions les plus pressantes est peut-être celle du choix des critères utilisés pour identifier les paramètres des modèles (fonction objectif) et pour contrôler leurs performances. Le Chapitre 11 a ouvert quelques pistes de réflexion sur cette question, liée à celle des incertitudes, mais tout reste à faire pour écrire l' ϵ des modèles hydrologiques ! De la réponse à cette question dépend, en un sens, notre capacité à les faire progresser, et en particulier à accepter des paramétrisations éventuellement un peu plus complexes.

Dans un tout autre registre, nous avons insisté à plusieurs reprises sur l'importance que pouvait avoir le contexte géologique pour expliquer le comportement des bassins versants. Cette insistance a pu susciter une interrogation chez le lecteur : pourquoi n'avoir pas présenté, dans ce cas, les résultats de l'utilisation de l'information géologique pour régionaliser les paramètres des modèles hydrologiques (comme par exemple dans *Oudin et al.*, 2008) ? Pour être honnête, ce n'est pas tant parce qu'ils sont décevants (même s'ils le sont effectivement), mais plutôt parce que nous manquons pour l'instant de descripteurs « synthétisant » cette information. Que faut-il choisir ? La fraction du bassin versant constituée de roches cristallines ? Celle constituée de carbonates ? Mais la situation est complètement différente selon que la moitié « imperméable » est à l'amont ou à l'aval, pour prendre ce seul exemple... Nous avons reporté en Annexe F un projet d'article lié à ces travaux : il y est question de la définition d'un indice de densité de drainage, tentative pour l'heure la plus aboutie pour lier la géologie et l'hydrologie de surface. La méthode proposée, basée sur l'identification d'un ordre de grandeur de la taille des bassins « sources » du réseau hydrographique, semble aller dans le sens des hypothèses de *Winter* (2001), pour qui la compétition entre écoulement souterrain et écoulement de surface est essentiellement contrôlée par la géologie et le climat. Nous montrons également la pertinence d'un tel indice pour la prédétermination de certains débits caractéristiques des bassins versants de notre échantillon : il ne s'agit pas encore de la prédétermination des paramètres de modèles hydrologiques, mais nous ne désespérons pas que cette information géologique puisse *in fine* y contribuer.

Le Chapitre 10, consacré à la définition de l'état initial d'une simulation pluie-débit, a laissé entrevoir l'existence de variabilités interannuelles qui n'apparaissent pas clairement dans les forçages climatiques. Les bassins versants de la Craie sont emblématiques de ces problèmes : on se heurte ici clairement à une limite de l'approche pluie-débit, limite qui n'étonnera certainement pas les hydrogéologues. De façon plus générale, les résultats des tests sur l'échantillon français, croisés à la richesse de l'information géographique disponible, ont permis de « *Penser / Classer* »¹ un peu mieux les gammes de comportements hydrologiques auxquelles un modèle peut être confronté. Pour reprendre l'exemple précédent, les bassins de la Craie présentent un taux d'échec que l'on ne retrouve dans aucune autre catégorie de bassins pourtant eux aussi influencés par le souterrain : ni parmi les bassins karstiques, ni parmi les grands bassins aux étiages soutenus par des aquifères alluviaux. Être capable d'anticiper dans quel(s) cas le modèle risque d'être mis en défaut est, après tout, un double progrès. Le premier est à usage externe : il permet la transparence envers les utilisateurs ; le second est à usage interne : il permet de savoir où consacrer ses efforts !

¹Hommage à Georges Perec !

Bibliographie

- [1] Allen, R. G., M. Smith, A. Perrier, et L. S. Pereira (1998), Crop evaporation — Guidelines for computing crop water requirements, *Irrigation and Drainage Paper*, 56, FAO, Rome, Italie.
- [2] Amiot, M. (1989), *Les ressources en eau du Châtillonnais*, Conseil Général de la Côte-d'Or, Dijon.
- [3] Andersen, O. B., et J. Hinderer (2005), Global inter-annual gravity changes from GRACE : Early results, *Geophys. Res. Lett.*, 32, 101402, doi :10.1029/2004GL020948.
- [4] Andréassian, V., C. Loumagne, T. Mathevet, C. Michel, L. Oudin, et C. Perrin (2007), What is really undermining hydrologic science today?, *Hydrological Processes*, 21, 2819–2822.
- [5] Asch, K. (2005), *IGME 5000 : 1: 5 Million International Geological Map of Europe and Adjacent Areas — final version for the internet*, BGR (Hannover).
- [6] Atkinson, T. C. (1977), Diffuse flow and conduit flow in limestone terrain in Mendip-Hills, Somerset (Great Britain), *J. Hydrol.*, 35, 93–110.
- [7] Barrett, M. E., et R. J. Charbeneau (1997), A parsimonious model for simulating flow in karst aquifer, *J. Hydrol.*, 196, 47–65.
- [8] Baudez, J. C. (1997), Déterminants hydrologiques régionaux pour la gestion et la prévision des ressources en eau, Rapport de stage, ENGEES.
- [9] Bergström, S. (1991), Principles and confidence in hydrological modelling, *Nordic Hydrology*, 22, 123–136.
- [10] Besbès, M. (2006), Recharge des aquifères par les crues d'oueds, in *Frontiers in Flood Research*, pp. 43–72, edited by I. Tchiguirinskaia, K.N.N. Thein, and P. Hubert, IAHS Publ. n° 305.
- [11] Beven, K. J. (2001), On hypothesis testing in hydrology, *Hydrol. Processes*, 15, 1655–1657.

- [12] Beven, K. J. (2001), On explanatory depth and predictive power, *Hydrol. Processes*, 15, 3069–3072.
- [13] Beven, K. J. (2006), Searching for the Holy Grail of scientific hydrology : $Q_t = H(\underline{S}, \underline{R}, \Delta t)A$ as closure, *Hydrol. Earth Syst. Sci.*, 10, 609–618.
- [14] Bonacci, O. (2001), Analysis of the maximum discharge of karst springs, *Hydrogeology Journal*, 9, 328–338.
- [15] Booij, M. J. (2002), Appropriate modelling of climate change impact on river flooding, Thèse de doctorat, Technical University of Twente, the Netherlands, 206 pp.
- [16] Box, G. E. P., et G. M. Jenkins (1976), *Time Series Analysis : Forecasting and Control*, rev. ed., Holden-Day, San Francisco, CA.
- [17] BRGM (1975), Motivations extraéconomiques en matière d'utilisation des eaux souterraines : l'hydroschizophrénie, *Rapport 75 SGN 206 AME*, traduction de Llamas (1974) et Llamas (1975) par J. Margat.
- [18] Carillo-Rivera, J. J. (2000), Application of the groundwater-balance equation to indicate interbasin and vertical flow in two semi-arid drainage basins, Mexico, *Hydrogeology Journal*, 8, 503–520.
- [19] Carillo-Rivera, J. J., A. Cardona, et D. Moss (1996), Importance of the vertical component of groundwater flow : a hydrogeochemical approach in the valley of San Luis Potosi, Mexico, *J. Hydrol.*, 185, 23–44.
- [20] Carrera, J., A. Alcolea, A. Medina, J. Hidalgo, et L. J. Slooten (2005), Inverse problem in hydrogeology, *Hydrogeology Journal*, 13, 206–222.
- [21] Chiew, F. H. S., et T. A. McMahon (1990), Estimating groundwater recharge using a surface watershed modelling approach, *J. Hydrol.*, 114, 285–304.
- [22] Chorley, R. J., A. J. Dunn, et R. P. Beckinsale (1964), *The History of the Study of Landforms, Vol. 1 : Geomorphology before Davis*, Methuen, London, England, 678 pp.
- [23] Cornet, J., et F. Lelong (2006), Bourgogne, in *Aquifères et eaux souterraines en France*, pp. 309–314, BRGM-AIH.
- [24] Costelloe, J. F., R. B. Grayson, R. M. Argent, et T. A. McMahon (2003), Modelling the flow regime of an arid zone floodplain river, Diamantina River, Australia, *Environmental Modelling and Software*, 18 (8–9), 693–703.
- [25] Croke, B. F. W. (2005), Land use impacts on hydrologic response in the Mae Chaem catchment, Northern Thailand, in *Proceedings of the 2005 International Conference on*

- Simulation and Modelling : Integrating Sciences and Technology for Effective Resource Management*, pp. 434–439, edited by V. Kachitvichyanukul, U. Purintrapiban, and P. Utayopas.
- [26] de Grissac, B., S. Rénié, et J.-C. Roux (2006), Sources de la Touvre et karst de la Rochefoucauld, in *Aquifères et eaux souterraines en France*, pp. 458–460, BRGM-AIH.
- [27] de Marsily, G. (1994), Quelques réflexions sur l'utilisation des modèles en hydrologie [Tribune libre], *Rev. Sci. Eau*, 7 (3), 219–234.
- [28] de Marsily, G., J.-P. Delhomme, F. Delay, et A. Buoro (1999), Regards sur 40 ans de problèmes inverses en hydrogéologie, "Le Point sur", *C. R. Acad. Sci. Paris*, 329, 73–87.
- [29] de Marsily, G., F. Delay, J. Gonçalves, P. Renard, V. Teles, et S. Violette (2005), Dealing with spatial heterogeneity, *Hydrogeology Journal*, 13, 161–183.
- [30] Delu, R., J. P. Hole, et J. C. Roux (1978), Vulnérabilité de la nappe de la Craie aux pollutions dans le bassin hydrogéologique d'Yport (Seine-Maritime), in *Hydrogéologie de la craie du bassin de Paris, tome 1 : Communications, Colloque Régional, Rouen, 25–26 mai 1978*, pp. 187–199, BRGM.
- [31] Diermanse, F. L. M. (2001), Physically-based modelling of rainfall-runoff processes, Thèse de doctorat, Technical University of Delft, the Netherlands, 234 pp.
- [32] Donnelly-Makowecki, L. M., et R. D. Moore (1999), Hierarchical testing of three rainfall-runoff models in small forested catchments, *J. Hydrol.*, 219 (3-4), 136–152.
- [33] Durand, Y. (1995), Analyse des principaux paramètres météorologiques de surface sur les massifs alpins et d'autres zones climatologiquement cohérentes, in *Notice scientifique et technique SAFRAN, volumes 1 et 2*, Centre d'Étude de la Neige, Grenoble. Météo France CNRM.
- [34] Durand, Y., E. Brun, L. Mérindol, G. Guyomarc'h, B. Lesassre, et E. Martin (1993), A meteorological estimation of relevant parameters for snow models, *Annals of Glaciology*, 18, 65–71.
- [35] Eakin, T. E. (1966), A regional interbasin groundwater system in the White River area, southeastern Nevada, *Water Resour. Res.*, 2(2), 251–271.
- [36] Edijatno (1991), Mise au point d'un modèle élémentaire pluie-débit au pas de temps journalier, Thèse de doctorat, Cemagref Antony, Université Louis Pasteur / ENGEES, Strasbourg, 242 pp.

- [37] Eisenlohr, L., L. Kiraly, M. Bouzelboudjen, et Y. Rossier (1997a), A numerical simulation as a tool for checking the interpretation of karst springs hydrographs, *J. Hydrol.*, 193, 306–315.
- [38] Eisenlohr, L., M. Bouzelboudjen, L. Kiraly, et Y. Rossier (1997b), Numerical versus statistical modeling of natural response of a karst hydrogeological system, *J. Hydrol.*, 202, 244–262.
- [39] Fisher, R. A. (1922), On the mathematical foundation of theoretical statistics, *Philos. Trans. Royal Soc. London, A* 222, 309–368.
- [40] Gan, T. Y., et G. F. Biftu (1996), Automatic calibration of conceptual rainfall-runoff models : optimization algorithms, catchment conditions, and model structure, *Water Resour. Res.*, 32(12), 3513–3524.
- [41] Garçon, R. (1999), Modèle global Pluie-Débit pour la prévision et la prédétermination des crues, *La Houille Blanche*, 54 (7–8), 88–95.
- [42] Genereux, D. P., et M. T. Jordan (2005), Interbasin groundwater flow and groundwater interaction with surface water in a lowland rainforest, Costa Rica : A review, *J. Hydrol.*, 320, 385–399.
- [43] Genereux, D. P., M. T. Jordan, et D. Carbonell (2005), A paired-watershed budget study to quantify interbasin groundwater flow in a lowland rain forest, Costa Rica, *Water Resour. Res.*, 41.
- [44] Gibrat, R. (1936), Aménagement hydroélectrique des cours d'eau : statistique mathématique et probabilités, *Revue Générale de l'Hydraulique*, Sep.–Oct.
- [45] Gleeson, T., et A. H. Manning (2008), Regional groundwater flow in mountainous terrain : Three-dimensional simulations of topographic and hydrogeologic controls, *Water Resour. Res.*, 44, W10403, doi :10.1029/2008WR006848.
- [46] Goswami, M. (2006), Rainfall-runoff simulation and forecasting using hydrological and data-driven Models & Techniques, Thèse de doctorat, National University of Ireland, Galway.
- [47] Goswami, M., et K. M. O'Connor (2005), Application of a conceptual rainfall-runoff simulation model to three European catchments characterised by non-conservative system behaviour, *International Conference on Hydrological Perspectives for Sustainable Development (HYPESD-'05)*, Indian Institute of Technology Roorkee, India.
- [48] Gröblacher, S., T. Paterek, R. Kaltenbaek, Č. Brukner, M. Żukowski, M. Aspelmeyer, et A. Zeilinger (2007), An experimental test of non-local realism, *Nature*, 446, 871–875.

- [49] Guillot, P. (1964), Une extension des lois A de Halphen comprenant comme cas limite la loi de Galton-Gibrat, *Revue de statistique appliquée*, 12 (1), 63–73.
- [50] Halihan, T., C. M. Wicks, et J. F. Engeln (1998), Physical response of a karst drainage basin to flood pulses : Example of the Devil's Icebox cave system (Missouri, USA), *J. Hydrol.*, 204, 24–36.
- [51] Hayami, S. (1951), On the propagation of flood waves, *Bulletin of the Disaster Prevention Research Institute, Kyoto Univ., Japan*, 1, 1–16.
- [52] Hudson, M. R., et D. N. Mott (1997), Faulting and coincident interbasin groundwater flow in a karst aquifer, Buffalo National River region, northwestern Arkansas, *Geological Society of America Abstracts with Programs*, 29 (6), 181–182.
- [53] Hughes, D. A. (2004), Incorporating groundwater recharge and discharge functions into an existing monthly rainfall-runoff model, *Hydrological Sciences Journal*, 49(2), 297–311.
- [54] Ivkovic, K. M., B. F. W. Croke, R. Letcher, et W. R. Evans (2005), The development of a simple model to investigate the impact of groundwater extraction on river flows in the Namoi catchment, NSW Australia, 'Where Waters Meet' NZHS-IAH-NSSSS Conference, 28 November – 2 December 2005, Auckland, New Zealand.
- [55] Jackson, D., et K. R. Rushton (1987), Assessment of recharge components for a chalk aquifer unit, *J. Hydrol.*, 92, 1–15.
- [56] Jakeman, A. J., et G. M. Hornberger (1993), How much complexity is warranted in a rainfall-runoff model?, *Water Resour. Res.*, 29 (8), 2637–2649.
- [57] Jakeman, A. J., I. G. Littlewood, et P. G. Whitehead (1990), Computation of the instantaneous unit hydrograph and identifiable component flow with application to two small upland catchments, *J. Hydrol.*, 117, 275–300.
- [58] Johannet, A., A. Mangin, et D. D'Hulst (1994), Subterranean water infiltration modelling by neural networks : use of water source flow, in *ICANN '94 - Proceedings of the International Conference on artificial neural networks, Sorrento, Italy, 26–29 May 1994*, pp. 1033–1036, Springer-Verlag, London.
- [59] Jorgensen, D., D. Signor, et J. Imes (1989a), Accounting for intracell flow in models with emphasis on water-table recharge and stream-aquifer interaction — part 1. Problems and concepts, *Water Resour. Res.*, 25 (4), doi : 10.1029/88WR04249.
- [60] Jorgensen, D., D. Signor, et J. Imes (1989b), Accounting for intracell flow in models with emphasis on water-table recharge and stream-aquifer interaction — part 2. A procedure, *Water Resour. Res.*, 25 (4), doi : 10.1029/88WR04248.

- [61] Jukić, D., et V. Denić-Jukić (2006), Nonlinear kernel functions for karst aquifers, *J. Hydrol.*, 328(1–2), 360–374.
- [62] Katz, B. G., J. S. Catches, T. D. Bullen, et R. L. Michel (1998), Changes in isotopic and chemical composition of ground water resulting from a recharge pulse from a sinking stream, *J. Hydrol.*, 211, 178–207.
- [63] Kauark-Leite, L. A., et N. Nascimento (1993), Développement, utilisation et incertitudes des modèles conceptuels en hydrologie, in *Modélisation du Comportement des Polluants dans les Hydrosystèmes*, pp. 191–219, Ministère de l'Environnement.
- [64] Kelly, K. S., et R. Krzysztofowicz (1997), A bivariate meta-Gaussian density for use in hydrology, *Stochastic Hydrol. Hydraul.*, 11, 17–31.
- [65] Kendall, C., et J. J. McDonnell (1998), *Isotope Tracers in Catchment Hydrology*, C. Kendall and J. J. McDonnell (Eds.), Elsevier Science, Amsterdam, 839 pp.
- [66] Kirchner, J. W. (2006), Getting the right answers for the right reasons : Linking measurements, analyses, and models to advance the science of hydrology, *Water Resour. Res.*, 42, W03S04, doi :10.1029/2005WR004362.
- [67] Klemeš, V. (1986), Dilettantism in hydrology : transition or destiny?, *Water Resour. Res.*, 22, 177–188.
- [68] Klemeš, V. (1986), Operational testing of hydrological simulation models, *Hydrological Sciences Journal*, 31(1), 13–24.
- [69] Konikow, L. F., et J. D. Bredehoeft (1992), Ground-water models cannot be validated, *Advances in Water Resources*, 15 (1), 75–83.
- [70] Konrad, C. P. (2006), Location and timing of river-aquifer exchanges in six tributaries to the Columbia River in the Pacific Northwest of the United States, *J. Hydrol.*, 329, 444–470.
- [71] Labat, D., R. Ababou, et A. Mangin (1999), Linear and nonlinear input/output models for karstic springflow and flood prediction at different time scales, *Stochastic Environ. Res. Risk Assessment*, 13(5), 337–364.
- [72] Labat, D., R. Ababou, et A. Mangin (2000), Rainfall-runoff relations for karstic springs. Part I : convolution and spectral analyses, *J. Hydrol.*, 238(3–4), 123–148.
- [73] Labat, D., R. Ababou, et A. Mangin (2000), Rainfall-runoff relations for karstic springs. Part II : continuous wavelet and discrete orthogonal multiresolution analyses, *J. Hydrol.*, 238(3–4), 149–178.

- [74] Lachassagne, P., et R. Wyns (2006), Aquifères de socle, in *Aquifères et eaux souterraines en France*, pp. 43–51, BRGM-AIH.
- [75] Laroque, M., et M. Razack (1998), Hydrogéologie du karst de La Rochefoucauld (Charente) : Bilan des nouvelles connaissances, *Hydrogéologie*, 3, 35–45.
- [76] Laroque, M., A. Mangin, M. Razack, et O. Banton (1998), Contribution of correlation and spectral analyses to the regional study of a large karst aquifer (Charente, France), *J. Hydrol.*, 205, 217–231.
- [77] Le Moigne, P. (2002), Description de l'analyse des champs de surface sur la France par le système SAFRAN, in *Note de centre n° 77*, Météo France CNRM/GAME/MC2, Toulouse.
- [78] Le Moine, N. (2005), Déterminants du bilan en eau à l'échelle du bassin versant : Vers une estimation réaliste des pertes des bassins versants de surface, Rapport de DEA / Master, UPMC / ENGREF.
- [79] Le Moine, N., V. Andréassian, C. Perrin, et C. Michel (2007), "Outlier" catchments : what can we learn from them in terms of prediction uncertainty in rainfall-runoff modelling?, in *Quantification and Reduction of Predictive Uncertainty for Sustainable Water Resources Management (Proceedings of Symposium HS2004 at IUGG2007, Perugia, July 2007)*. IAHS Publ. n° 313, pp. 195–203, edited by E. Boegh, H. Kunstmann, T. Wagener, A. Hall, L. Bastidas, S. Franks, H. Gupta, D. Rosbjerg & J. Schaake.
- [80] Lerat, J., C. Perrin, et V. Andréassian (soumis), Value of upstream measured discharges in hydrological modeling : a model for intermediary areas.
- [81] Linsley, R. K. (1982), Rainfall-runoff models — An overview, in *Proceedings of the international symposium on rainfall-runoff modelling*, pp. 3–22, Edited by V. P. Singh, Water Resources Publications, Littleton, CO.
- [82] Linsley, R. K., M. A. Kohler, et J. L. H. Paulus (1949), *Applied Hydrology*, McGraw-Hill, Inc., New York, NY.
- [83] Llamas, M. R. (1974), Motivaciones extraeconómicas en la utilización de las aguas subterráneas : la hidroesquizofrenia, *Agua*, 82, 29–36.
- [84] Llamas, M. R. (1974), Hacia una política hidráulica sin hidroesquizofrenia, in *Actas de la I^a Conferencia Nacional sobre Hidrología Aplicada*, Zaragoza.
- [85] Longuevergne, L. (2008), Contribution à l'Hydrogéodésie, Thèse de doctorat, Université Pierre et Marie Curie, 292+xxviii pp.

- [86] Mangin, A. (1981), Apport des analyses corrélatoire et spectrale dans l'approche des systèmes hydrologiques, *C.R. Acad. Sci. Paris, Série II* 293, 401–404.
- [87] Mangin, A. (1984), Pour une meilleure connaissance des systèmes hydrologiques à partir des analyses corrélatoires et spectrales, *J. Hydrol.*, 67, 25–43.
- [88] Mann, H. B., et D. R. Whitney (1947), On a test of whether one of two random variables is stochastically larger than the other, *Annals of Mathematical Statistics*, 18, 50–60.
- [89] Margat, J. (2008), *Les eaux souterraines dans le monde*, BRGM Éditions, Orléans / UNESCO, Paris, 187 pp.
- [90] Matheron, G. (1978), *Estimer et choisir — Essai sur la pratique des Probabilités*, École Nationale des Mines de Paris, 175 pp.
- [91] Matheron, G. (1984), L'émergence de la loi de Darcy, *Annales des Mines, Mai–Juin 1984*, 11–16.
- [92] Mathevet, T. (2005), Quels modèles pluie-débit globaux pour le pas de temps horaire ? Développements empiriques et comparaison de modèles sur un large échantillon de bassins versants, Thèse de doctorat, ENGREF, Paris.
- [93] Mathevet, T., C. Michel, V. Andréassian, et C. Perrin (2006), A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins, in *Large sample basin experiments for hydrological model parameterization : Results of the MModel Parameter EXperiment — MOPEX*, pp. 117–125, Edited by V. Andréassian, A. Hall, N. Chahinian, and J. Schaake, IAHS Publ. n° 307.
- [94] MEDD / BRGM (2003), Mise en oeuvre de la DCE — Identification et délimitation des masses d'eau souterraine, guide méthodologique.
- [95] Medina, A., et J. Carrera (1996), Coupled estimation of flow and solute transport parameters, *Water Resour. Res.*, 32, 3063–3076.
- [96] Michel, C., V. Andréassian, et C. Perrin (2006), Has rainfall-runoff modeling gone beyond empiricism?, in *Large sample basin experiments for hydrological model parameterization : Results of the MModel Parameter EXperiment — MOPEX*, pp. 117–125, Edited by V. Andréassian, A. Hall, N. Chahinian, and J. Schaake, IAHS Publ. n° 307.
- [97] Mouelhi, S. (2003), Vers une chaîne cohérente de modèles pluie-débit conceptuels globaux aux pas de temps pluriannuel, annuel, mensuel et journalier, Thèse de doctorat, ENGREF, Paris.
- [98] Mouelhi, S., C. Michel, C. Perrin, et V. Andréassian (2006), Stepwise development of a two-parameter monthly water balance model, *J. Hydrol.*, 318(1-4), 200–214.

- [99] Moulin, L. (2007), Utilisation de modèles hydrologiques globaux pour la prévision des crues rapides, Thèse de doctorat, ENGREF, Paris, 653 pp.
- [100] Nace, R. L. (1973), Communication AWRA, *Ground Water*, 11 (2), 48–49.
- [101] Nandakumar, N., et R. G. Mein (1997), Uncertainty in rainfall-runoff model simulations and the implications for predicting the hydrologic effects of land-use change, *J. Hydrol.*, 192, 211–232.
- [102] Nascimento, N. O. (1995), Appréciation à l'aide d'un modèle empirique des effets d'action anthropiques sur la relation pluie-débit à l'échelle du bassin versant, Thèse de doctorat, ENPC, Paris.
- [103] O'Connell, P. E., J. E. Nash, et J. P. Farrell (1970), River flow forecasting through conceptual models. Part II - The Brosna catchment at Ferbane, *J. Hydrol.*, 10, 317–329.
- [104] Oudin, L. (2004), Recherche d'un modèle d'évapotranspiration potentielle comme entrée d'un modèle pluie-débit global, Thèse de doctorat, ENGREF, Paris.
- [105] Oudin, L., C. Michel, et F. Anctil (2005a), Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 1 — Can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs?, *J. Hydrol.*, 303(1–4), 275–289.
- [106] Oudin, L., F. Hervieu, C. Michel, C. Perrin, V. Andréassian, F. Anctil, et C. Loumagne (2005b), Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2 — Towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling, *J. Hydrol.*, 303(1–4), 290–306.
- [107] Oudin, L., V. Andréassian, T. Mathevet, C. Perrin, et C. Michel (2006), Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations, *Water Resour. Res.*, 42(7), W07410, doi :10.1029/2005WR004636.
- [108] Oudin, L., V. Andréassian, C. Perrin, C. Michel, et N. Le Moine (2008), Spatial proximity, physical similarity, regression and ungaged catchments : a comparison of regionalization approaches based on 913 French catchments, *Water Resour. Res.*, 44, W03413, doi :10.1029/2007WR006240.
- [109] Padilla, A., et A. Pulido-Bosch (1995), Study of hydrographs of karstic aquifers by means of correlation and cross-spectral analysis, *J. Hydrol.*, 168(1–4), 73–89.
- [110] Paquet, E. (2004), Évolution du modèle hydrologique MORDOR : modélisation du stock nival à différentes altitudes, *La Houille Blanche*, 2, 75–82.

- [111] Paquet, E., J. Gailhard, et R. Garçon (2006), évolution de la méthode du gradex : approche par type de temps et modélisation hydrologique, *La Houille Blanche*, 5, 80–90.
- [112] Payan, J.-L. (2007), Prise en compte de barrages-réservoirs dans un modèle global pluie-débit, Thèse de doctorat, ENGREF / CEMAGREF, 256 pp.
- [113] Peach, D., P. Shand, D. Gooddy, C. Abesser, J. P. Bloomfield, et A. Gallagher (2006), The Hydrogeology and Hydrogeochemistry of the Pang and Lambourn Catchments, Berkshire, *British Geological Survey Commissioned Report*, CR/06/050N.
- [114] Perrin, C. (2000), Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative, Thèse de doctorat, INPG.
- [115] Perrin, C., C. Michel, et V. Andréassian (2001), Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments, *J. Hydrol.*, 242(3–4), 275–301.
- [116] Perrin, C., C. Michel, et V. Andréassian (2003), Improvement of a parsimonious model for streamflow simulation, *J. Hydrol.*, 279, 275–289.
- [117] Pike, J. G. (1964), The estimation of annual runoff from metrological data in tropical climate, *J. Hydrol.*, 2, 116–123.
- [118] Pitman, W. V. (1973), A mathematical model for generating river flows from meteorological data in South Africa., *Report n° 2/73, Hydrological Research Unit, University of the Witwatersrand, Johannesburg, South Africa.*
- [119] Popper, K. R. (1934), *Logik der Forschung*, Springer, Vienna.
- [120] Popper, K. R. (1954), Degree of Confirmation, *British Journal for the Philosophy of Science*, 5 (18), 143–149.
- [121] Popper, K. R. (1957), A Second Note on Degree of Confirmation, *British Journal for the Philosophy of Science*, 7 (28), 350–353.
- [122] Popper, K. R. (1958), A Third Note on Degree of Confirmation, *British Journal for the Philosophy of Science*, 8 (32), 294–302.
- [123] Popper, K. R. (1959), *The Logic of Scientific Discovery*, Hutchinson, London, English translation, with new preface, footnotes and appendices of *Logik der Forschung*, 1934.,.
- [124] Popper, K. R. (1983), *Postscript to the Logic of Scientific Discovery — Volume I: Realism and the Aim of Science*, Routledge.

- [125] Porter, J. W., et T. A. McMahon (1976), The Monash model : user manual for daily program HYDROLOG, Res. Rep. 2/76, Department of Civil Engineering, Monash University, Victoria, 41 pp.
- [126] Portniaguine, O., et D. K. Solomon (1998), Parameter estimation using groundwater age and head data, Cape Cod, Massachussets, *Water Resour. Res.*, 34, 637–645.
- [127] Quelennec, R. E., J. C. Sauret, M. Seguin, et J. Vouve (1971), Les résurgences de la Touvre — Étude préliminaire, in *Actes du Colloque d'Hydrologie en Pays Calcaire, Besançon*, Annales Scientifiques de l'Université de Besançon, fascicule 15, 3ème série, Géologie, 1971.
- [128] Refsgaard, J. C., et J. Knudsen (1996), Operational validation and intercomparison of different types of hydrological models, *Water Resour. Res.*, 32 (7), 2189–2202.
- [129] Rimmer, A., et Y. Salingar (2006), Modelling precipitation-streamflow processes in karst basin : The case of the Jordan River sources, Israel, *J. Hydrol.*, 331, 524–542.
- [130] Rousset, F., F. Habets, E. Gomez, P. Le Moigne, S. Morel, J. Noilhan, et E. Ledoux (2004), Hydrometeorological modeling of the Seine basin using the SAFRAN-ISBA-MODCOU system, *Journal of Geophysical Research*, 109, doi :10.1029/2003JD004403.
- [131] Sanford, W. (2002), Recharge and groundwater models : an overview, *Hydrogeology Journal*, 10, 110–120.
- [132] Schmidt, R., P. Schwintzer, F. Flechtner, C. Reigber, A. Guntner, P. Doll, G. Ramillien, A. Cazenave, S. Petrovic, H. Jochmann, et J. Wunsch (2006), GRACE observations of changes in continental water storage, *Global and Planetary Change*, 50 (1–2), 112–126.
- [133] Sorooshian, S., et J. A. Dracup (1980), Stochastic parameter estimation procedures for hydrologic rainfall-runoff models : Correlated and heteroscedastic error cases, *Water Resour. Res.*, 16, 430–442.
- [134] Tabbagh, A., H. Bendjoudi, et Y. Benderitter (1999), Determination of recharge in unsaturated soils using temperature monitoring, *Water Resour. Res.*, 35, 2439–2446.
- [135] Tangara, M. (2005), Nouvelle méthode de prévision de crue utilisant un modèle pluie-débit global, Thèse de doctorat, CEMAGREF, EPHE, Paris, 374 pp.
- [136] Therrien, R., et E. A. Sudicky (1996), Three-dimensional analysis of variably-saturated flow and solute transport in discretely-fractured porous media, *Journal of Contaminant Hydrology*, 23(1–2), 1–44.
- [137] Thyer, M. A., G. Kuczera, et Q. J. Wang (2002), Quantifying parameter uncertainty in stochastic models using the Box-Cox transformation, *J. Hydrol.*, 265, 246–257.

- [138] Thyne, G. D., J. M. Gillespie, et J. R. Ostdick (1999), Evidence for interbasin flow through bedrock in the southeastern Sierra Nevada, *Geological Society of America Bulletin*, 111(11), 1600–1616.
- [139] Tóth, J. A. (1963), A theoretical analysis of groundwater flow in small drainage basins, *Journal of Geophysical Research*, 68, 4795–4812.
- [140] Tóth, J. A. (1995), Hydraulic continuity in large sedimentary basins, *Hydrogeology Journal*, 3(4), 4–16.
- [141] Turc, L. (1954), Le bilan d'eau des sols : relation entre les précipitations, l'évapotranspiration et l'écoulement, *Annales agronomiques, Série A*, 491–595.
- [142] Weizenbaum, J. (1976), *Computer Power and Human Reason : From Judgment to Calculation*, W. H. Freeman, San Francisco.
- [143] Wheater, H. S., C. Neal, et D. Peach (2006), Hydro-ecological functioning of the Pang and Lambourn catchments, UK : An introduction to the special issue, *J. Hydrol.*, 330, 1–9, doi :10.1016/j.jhydrol.2006.04.035.
- [144] Wilcoxon, F. (1945), Individual comparisons by ranking methods, *Biometrics Bulletin*, 1, 80–83.
- [145] Winter, T. C. (2001), The concept of hydrologic landscapes, *J. Am. Water Resour. Assoc.*, 37, 335–349.
- [146] Yang, J., P. Reichert, K. C. Abbaspour, et H. Yang (2007), Hydrological modelling of the Chaohe Basin in China : Statistical model formulation and Bayesian inference, *J. Hydrol.*, 340, 167–182.
- [147] Yevjevich, V. (1972), *Stochastic Processes in Hydrology*, Water Research Publications, Fort Collins, CO.

Annexes

A

Table de correspondance Critère de Nash – Critère C_{MM}

C_{NS}	$\rightarrow C_{MM} = \frac{C_{NS}}{2 - C_{NS}}$	C_{MM}	$\rightarrow C_{NS} = \frac{2 \cdot C_{MM}}{1 + C_{MM}}$
-∞	-1.000	-1.000	-∞
-10.000	-0.833	-0.900	-18.000
-8.000	-0.800	-0.800	-8.000
-6.000	-0.750	-0.700	-4.667
-5.000	-0.714	-0.600	-3.000
-4.000	-0.667	-0.500	-2.000
-3.000	-0.600	-0.400	-1.333
-2.000	-0.500	-0.300	-0.857
-1.000	-0.333	-0.200	-0.500
-0.500	-0.200	-0.100	-0.222
0.000	0.000	0.000	0.000
0.100	0.053	0.100	0.182
0.200	0.111	0.200	0.333
0.300	0.176	0.300	0.462
0.400	0.250	0.400	0.571
0.500	0.333	0.500	0.667
0.600	0.429	0.600	0.750
0.700	0.538	0.700	0.824
0.800	0.667	0.800	0.889
0.900	0.818	0.900	0.947
1.000	1.000	1.000	1.000

B

Estimation de la significativité des écarts entre structures à l'aide d'un test non-paramétrique

B.1 Problématique

Aux Chapitres 8 et 9, nous avons testé un grand nombre de structures de modèles sur un large échantillon de bassins versants. Cependant, nous avons déjà mentionné que la sélection de *l'une* de ces structures parmi l'ensemble est un choix *in fine* subjectif : la différence de performances entre deux structures M_i et M_j est-elle statistiquement significative *pour une taille donnée d'échantillon*¹ ?

Mathevet (2005) propose de tester l'alternative suivante :

$$\left\{ \begin{array}{l} H_0 : \text{ Les deux séries sont issues de la même distribution, i.e. } M_i \text{ et } M_j \text{ sont } \textit{ex-aequo}, \\ H_1 : \text{ Les deux séries ne sont pas issues de la même distribution,} \\ \text{ i.e. l'un des modèles est supérieur l'autre.} \end{array} \right.$$

Dans cette annexe, nous reprenons la méthode de sous-échantillonnage proposée, en l'adaptant sur trois points :

- plutôt que le test de Student, nous choisissons le test non-paramétrique de Wilcoxon–Mann–Whitney (Wilcoxon, 1945 ; Mann et Whitney, 1947) pour tester si deux échantillons sont issus de la même distribution (la normalité des distributions de critères C_{MM} n'étant pas assurée),

¹et sous réserve, bien entendu, que le critère utilisé soit une bonne mesure de la qualité du modèle, ce qui est encore un autre problème...

- nous ne considérons pas le résultat du test sur l'échantillon complet comme donnant la réponse à l'alternative (même au risque α) : l'échantillon complet peut lui-même être vu comme un sous-échantillon d'autres échantillons potentiels,
- enfin, une différence pratique est que nous utilisons des modèles peu différents les uns des autres, voire emboîtés, de façon à estimer les régions de rejet dans le cas où l'on doit choisir un niveau de complexité parmi plusieurs modèles d'une même « famille ».

B.2 Description du test

1. Dans ce test, nous utilisons comme ensemble de modèles, les structures M_{001} à M_{096} au pas de temps journalier, présentées au Chapitre 8. Ces structures possèdent entre trois et sept paramètres, leur performance moyenne en terme de $C_{MM}(\sqrt{Q})$ variant entre 0.472 et 0.642. Nous disposons donc de $C_2^{96} = \frac{96 \times 95}{2} = 4560$ couples de modèles (M_i, M_j) , dont les écarts de performance se situent entre 0 et 17 points, sur lesquels nous pouvons estimer le taux de rejet *empirique* de l'hypothèse nulle H_0 pour différentes tailles de sous-échantillon,
2. à chaque fois, les deux échantillons de départ sont les distributions de critères $C_{MM}(\sqrt{Q})$ en contrôle, soit 4070 valeurs¹. Nous testons trois tailles de sous-échantillons : $N = 10, 30, 100, 300$ et 1000 bassins-périodes, en tirant à chaque fois 1000 sous-échantillons de taille N dans les deux distributions de départ,
3. Pour chaque paire (M_i, M_j) et chaque taille N , nous relevons la fréquence de rejet de l'hypothèse H_0 sur les 1000 essais, ainsi que l'espérance des différences de performance sur ces essais.

B.3 Résultats

Sur la Figure B.1 nous avons tracé, pour chaque paire (M_i, M_j) et chaque taille d'échantillon, le taux de rejet empirique en fonction de l'espérance de la différence de performance des structures sur les sous-échantillons.

¹La réunion des échantillons français et australiens totalise 1389 bassins, mais nous avons en général utilisé un *split sample-test* à deux périodes de cinq ans pour les bassins français et à trois périodes de cinq ans pour les bassins australiens de façon à rentabiliser la longueur des séries de débits australiennes, plus longues que les séries françaises.

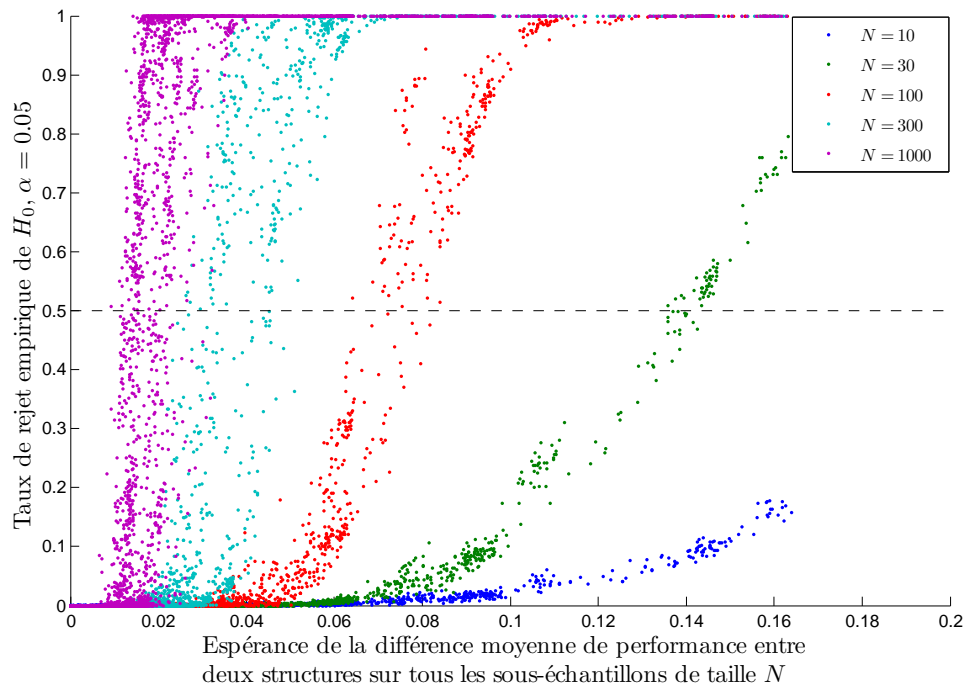


FIG. B.1: Taux de rejet de l'hypothèse nulle (identité des distributions de performance pour un couple de modèles donné) au risque de première espèce α , en fonction de l'espérance de la différence moyenne de performance sur tous les sous-échantillons de taille N .

Nous cherchons alors à définir une zone « raisonnable » de rejet de H_0 en fonction de la taille de l'échantillon et de la différence moyenne de performance entre les structures (puisque c'est l'indicateur synthétique que nous avons choisi de prendre au Chapitre 1). Pour cela, nous devons fixer un second seuil, subjectif, de fréquence de rejet empirique à partir duquel nous acceptons effectivement de rejeter H_0 . Comme dans *Mathevet* (2005), nous prenons la valeur 0.5 (si nous voulions être plus conservatifs nous pourrions prendre 0.9 ou 0.95!). Cela revient à dire que si le test de Wilcoxon, effectué avec un risque de première espèce $\alpha = 0.05$, conduit à rejeter l'hypothèse nulle pour un sous-échantillon sur deux de taille N , alors nous acceptons de considérer que ce taux de rejet empirique élevé est effectivement dû à une différence dans les distributions d'origine.

Ce seuil fixé, nous pouvons sélectionner, pour chaque valeur de N , l'espérance de la différence moyenne de performances qui donne un taux de rejet correspondant (à un intervalle près, par exemple en sélectionnant les points situés entre les ordonnées 0.45 et 0.55). Les points sélectionnés sont placés sur la Figure B.2.

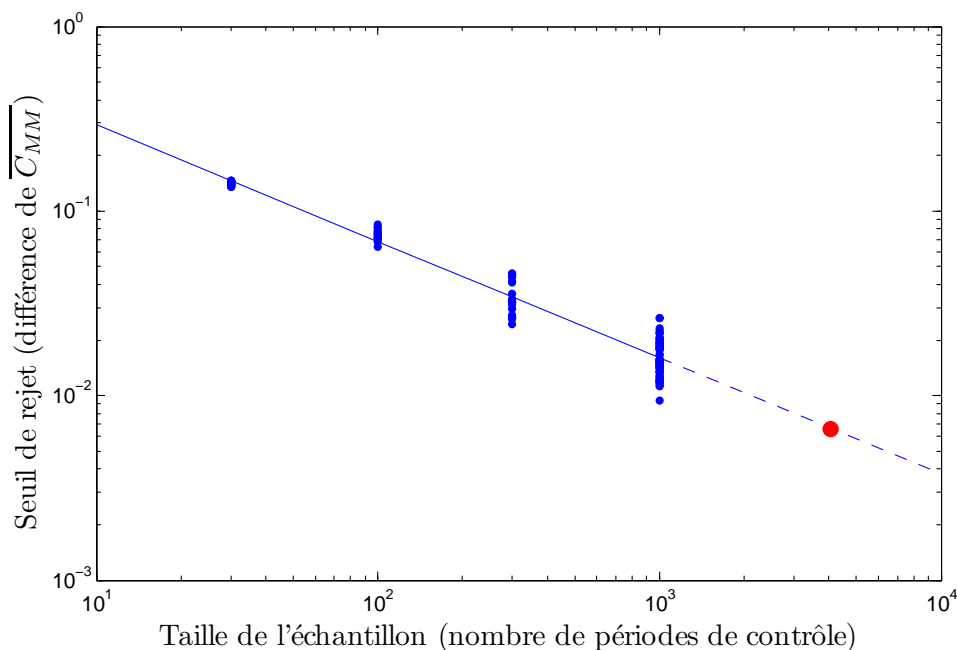


FIG. B.2: Relation entre la taille de l'échantillon et la différence moyenne de performance permettant d'atteindre un taux de rejet empirique de 0.5 au risque de première espèce $\alpha = 0.05$.

Nous pouvons alors nous risquer à une extrapolation de cette courbe empirique jusqu'à la taille complète de l'échantillon (4070 périodes de contrôle), en considérant cet échantillon (et les différences moyennes de performances qui lui sont associées) comme une réalisation d'un échantillon de taille $N = 4070$. Cela nous donne une valeur d'environ 0.007, soit 0.7 point de C_{MM} , avec naturellement une variance que nous ne chercherons toutefois pas à estimer : celle-ci dépend des performances absolues des modèles, de leur écart-type (i.e., de la robustesse des modèles). En étant conservatifs, nous pourrions aisément rejeter tout écart inférieur à 2 points.

B.4 Synthèse

La méthode que nous avons présentée est bien sûr extrêmement grossière mais nous pouvons retenir l'ordre de grandeur identifié : même sur notre très large échantillon, une différence de performance moyenne n'a de chance d'être significative qu'à partir d'**1 point en contrôle**. Le test mis en oeuvre donne d'ailleurs des ordres de grandeur comparables à ceux obtenus par *Mathevet* (2005), qui avance la valeur de 5 points sur un échantillon de 313 bassins versants et 2093 périodes de contrôle (plus courtes que celles de notre échantillon : généralement un an, occasionnant des performances globalement plus faibles et des écarts-type plus importants). C'est pourquoi nous reprenons sa conclusion : « *D'un point de vue pratique, ce test montre que, quand la différence de performance entre deux modèles est faible [...], il est généralement impossible de conclure significativement qu'un modèle est meilleur*

qu'un autre avec un petit échantillon de bassins versants, comme c'est fait généralement [...] Bien que très exigeant, ce test indique que, quand les différences de performances sont inférieures à 5 points, elles sont à peine suffisantes pour différencier deux modèles, même sur un échantillon assez grand. »

Dépendance des paramètres de capacités au pas de temps de fonctionnement

C.1 Loi de vidange instantanée

Dans cette annexe nous considérons un réservoir dont la vidange suit une loi puissance, d'exposant $\alpha > 0$, de forme instantanée :

$$\begin{aligned}\frac{dR}{dt} &= -h \cdot R^\alpha \\ \frac{dR}{R^\alpha} &= -h \cdot dt\end{aligned}$$

Le modèle fonctionnant avec un pas de temps Δt , nous cherchons l'expression du volume écoulé par le réservoir entre t et $t + \Delta t$ en fonction des paramètres h , α et du contenu R_t du réservoir à la date t .

C.2 Cas où $\alpha > 1$

L'équation instantanée s'écrit :

$$d\left(\frac{1}{1-\alpha} R^{1-\alpha}\right) = -h \cdot dt$$

Le niveau $R_{t+\Delta t}$ après un intervalle de durée Δt vérifie donc :

$$\begin{aligned} R_{t+\Delta t}^{1-\alpha} - R_t^{1-\alpha} &= (\alpha - 1)h\Delta t \\ \left(\frac{R_{t+\Delta t}}{R_t}\right)^{1-\alpha} &= 1 + (\alpha - 1)R_t^{\alpha-1}h\Delta t \\ R_{t+\Delta t} &= R_t \left[1 + (\alpha - 1)R_t^{\alpha-1}h\Delta t\right]^{\frac{1}{1-\alpha}} \end{aligned}$$

En posant $C = \left(\frac{1}{(\alpha-1)h\Delta t}\right)^{\frac{1}{\alpha-1}}$, on obtient :

$$R_{t+\Delta t} = R_t \left[1 + \left(\frac{R_t}{C}\right)^{\alpha-1}\right]^{-\frac{1}{\alpha-1}} \quad (\text{C.1})$$

et donc :

$$Q_{\Delta t} = R_t - R_{t+\Delta t} = R_t \left(1 - \left[1 + \left(\frac{R_t}{C}\right)^{\alpha-1}\right]^{-\frac{1}{\alpha-1}}\right) \quad (\text{C.2})$$

L'Équation C.1 permet de justifier l'appellation de « capacité à un pas de temps » pour C . En effet, en notant $r_t = \frac{R_t}{C}$, cette équation s'écrit :

$$r_{t+\Delta t} = r_t \left[1 + r_t^{\alpha-1}\right]^{-\frac{1}{\alpha-1}}$$

On peut vérifier que la fonction qui à r_t associe $r_{t+\Delta t}$ est strictement croissante, et de limite 1. Quelque soit la valeur de départ R_t , le stock $R_{t+\Delta t}$ ne peut donc pas dépasser la valeur C et tout excédent est transféré à $Q_{\Delta t}$.

On en déduit également la formule permettant de transformer le paramètre C en C' lorsque l'on change le pas de temps Δt en $\Delta t'$:

$$\frac{C'}{C} = \left(\frac{\Delta t}{\Delta t'}\right)^{\frac{1}{\alpha-1}} \quad (\text{C.3})$$

Si l'on passe par exemple du pas de temps journalier au pas de temps horaire, la capacité est multipliée par $24^{\frac{1}{\alpha-1}}$, et par conséquent le taux de remplissage moyen est divisé par une valeur du même ordre (les stocks R_t ne changeant pas, eux, d'ordre de grandeur).

C.3 Cas où $\alpha < 1$

Dans ce cas, on a encore :

$$R_{t+\Delta t}^{1-\alpha} - R_t^{1-\alpha} = (\alpha - 1)h\Delta t$$

Cependant on ne peut pas intégrer sur tout le pas de temps quelque soit la valeur de R_t : en effet, si R_t est tel que $R_t^{1-\alpha} < (1 - \alpha)h\Delta t$, alors $R_{t+\Delta t}$ est nul.

En posant cette fois $C = \left(\frac{1}{(1-\alpha)h\Delta t}\right)^{\frac{1}{\alpha-1}}$, on obtient :

$$R_{t+\Delta t} = \begin{cases} R_t \left[1 - \left(\frac{R_t}{C}\right)^{\alpha-1}\right]^{-\frac{1}{\alpha-1}} & \text{si } R_t > C \\ 0 & \text{si } R_t \leq C \end{cases}$$

et donc :

$$Q_{\Delta t} = \begin{cases} R_t \left(1 - \left[1 - \left(\frac{R_t}{C}\right)^{\alpha-1}\right]^{-\frac{1}{\alpha-1}}\right) & \text{si } R_t > C \\ R_t & \text{si } R_t \leq C \end{cases}$$

En d'autre terme, si le réservoir contient moins que la valeur C (qui n'est plus une « capacité à un pas de temps ») au pas de temps t , il se vide entièrement. L'expression du rapport entre C et C' à deux pas de temps différents reste en revanche la même que dans le cas $\alpha > 1$.

C.4 Cas où $\alpha = 1$

C'est le cas classique du réservoir linéaire :

$$\begin{aligned} \frac{dR}{R} &= -h \cdot dt \\ d(\ln R) &= -h \cdot dt \\ \ln \frac{R_{t+\Delta t}}{R_t} &= -h\Delta t \\ R_{t+\Delta t} &= R_t e^{-h\Delta t} \end{aligned}$$

En posant $f = 1 - e^{-h\Delta t}$ ($f < 1$), on écrira :

$$\begin{aligned}R_{t+\Delta t} &= (1 - f) \cdot R_t \\ Q_{\Delta t} &= f \cdot R_t\end{aligned}$$

où f est la fraction du réservoir écoulee à chaque pas de temps. La transformation de f en changeant de pas de temps s'effectue par :

$$\begin{aligned}\frac{\ln(1 - f')}{\ln(1 - f)} &= \frac{\Delta t'}{\Delta t} \\ f' &= 1 - (1 - f)^{\frac{\Delta t'}{\Delta t}}\end{aligned}\tag{C.4}$$

Méthode alternative d'identification de l'hydrogramme unitaire horaire

D.1 Objectifs

La question de la forme optimale de l'hydrogramme unitaire au pas de temps horaire a été l'un des points importants développés par *Mathevet* (2005). Nous avons donc poursuivi ces tests avec la structure à cinq paramètres sélectionnée au Chapitre 9, en testant différentes formes paramétriques. Dans cette annexe, nous présentons une méthode non-paramétrique suggérée par *Croke* (2005) pour l'identification de la fonction de transfert du modèle IHACRES (e.g., *Jakeman et al.*, 1990) basée sur la déconvolution du signal de débit et du signal de pluie nette. Nous expliquons le principe général, puis la façon dont nous avons dû l'adapter à la structure des modèles GR (dans laquelle l'hydrogramme unitaire est « imbriqué » entre d'autres composants non linéaires). Enfin nous présenterons les résultats du test au pas de temps horaire, avec la structure M_{046} (GR5H).

D.2 Principe (d'après *Croke*, 2005)

Si l'on fait l'hypothèse que la relation entre l'entrée $X(t)$ et la sortie $Y(t)$ d'un système est linéaire et invariante dans le temps, on peut écrire la sortie sous la forme d'une convolution entre l'entrée et la fonction de transfert H du système (noyau de convolution invariant dans le temps) :

$$Y = X * H \quad (\text{D.1})$$

En passant dans le domaine fréquentiel, la transformée de Fourier de cette convolution est le produit des transformées de Fourier :

$$\mathcal{F}[Y] = \mathcal{F}[X] \cdot \mathcal{F}[H] \quad (\text{D.2})$$

On en déduit donc la transformée de Fourier $\mathcal{F}[H]$, puis le noyau H lui-même par une transformée inverse :

$$\begin{aligned} \mathcal{F}[H] &= \frac{\mathcal{F}[Y]}{\mathcal{F}[X]} \\ H &= \mathcal{F}^{-1} \left[\frac{\mathcal{F}[Y]}{\mathcal{F}[X]} \right] \end{aligned} \quad (\text{D.3})$$

Cependant, en procédant ainsi on obtient souvent une estimation de H assez bruitée et peu robuste. L'alternative est de transformer l'Équation D.1 en prenant la corrélation croisée¹ des deux membres avec la série X :

$$X \star Y = X \star (X * H) \quad (\text{D.4})$$

On en déduit alors, par associativité (convolution et corrélation croisée) :

$$\begin{aligned} X \star Y &= (X \star X) * H \\ \mathcal{F}[X \star Y] &= \mathcal{F}[X \star X] \cdot \mathcal{F}[H] \\ \mathcal{F}[H] &= \frac{\mathcal{F}[X \star Y]}{\mathcal{F}[X \star X]} \\ H &= \mathcal{F}^{-1} \left[\frac{\mathcal{F}[X \star Y]}{\mathcal{F}[X \star X]} \right] \end{aligned} \quad (\text{D.5})$$

On constate que l'Équation D.5 est tout à fait similaire à l'Équation D.3 : le signal d'entrée X est simplement remplacé par sa fonction d'autocorrélation $X \star X$, et le signal de sortie Y par la fonction de corrélation croisée $X \star Y$. Ces deux fonctions sont, comme X et Y , des fonctions d'une variable homogène à un temps (un délai τ). Leur substitution aux signaux d'origine a simplement pour effet d'améliorer le rapport signal/bruit de la fonction de transfert H déconvoluée, l'opérateur \star étant construit en pratique par une série de moyennes.

D.3 Adaptation à l'hydrogramme unitaire du modèle horaire

On peut appliquer cette méthode de déconvolution avec, par exemple, la pluie en entrée et le débit en sortie. Dans le cas présent, la fonction de transfert à identifier est l'hydrogramme

¹La définition statistique s'exprime au moyen d'une espérance : $(X \star Y)(\tau) = E \left[\frac{(X(t) - \mu_X) \cdot (Y(t+\tau) - \mu_Y)}{\sigma_X \sigma_Y} \right]$

unitaire du modèle pluie-débit, qui est situé entre le module de production et le module de routage non-linéaire. L'entrée est donc la production Pr , tandis que la sortie n'est pas directement connue : c'est le terme I qui alimente le routage (cf. Figure D.1). Pour effectuer la déconvolution, il faut nécessairement avoir une estimation de cette série I , estimation que l'on va donc définir comme la *série idéale qui, alimentant le module de routage, donnerait le débit observé*. On la calculera donc en partant du débit observé et en inversant la fonction de routage non-linéaire, procédure décrite plus bas.

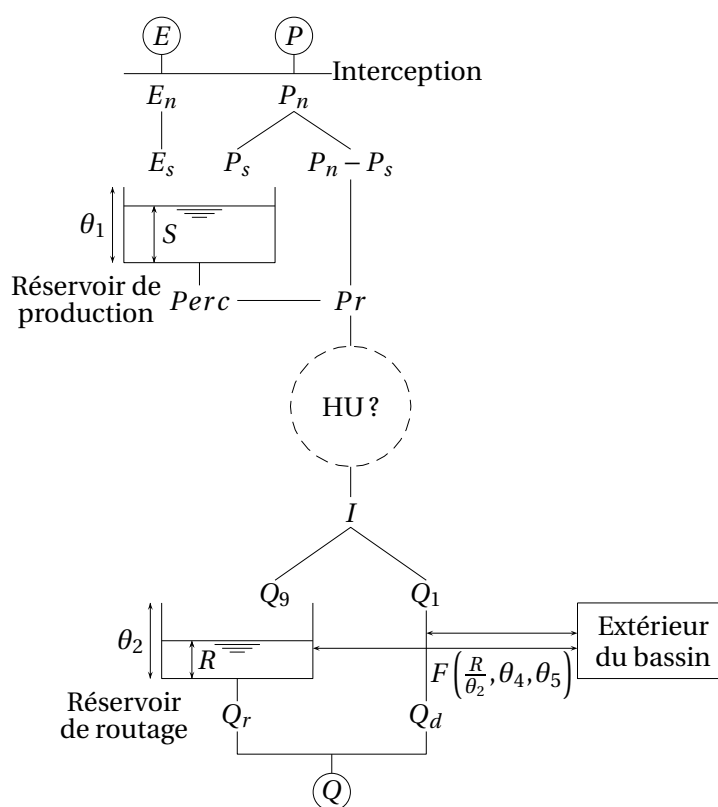


FIG. D.1: Variante de modèle dans lequel aucune forme n'est imposée à l'hydrogramme unitaire, sa forme dépendant implicitement des autres paramètres (production : θ_1 ; routage non-linéaire : θ_2 ; échanges : θ_4, θ_5). Le paramètre θ_3 (temps de base de l'HU) est supprimé.

D.3.1 Description de la procédure de calage dans l'approche testée

À chaque jeu de paramètres $\{\theta_i\}_{i=1\dots n}$ il faut calculer les ordonnées de l'hydrogramme unitaire par déconvolution de la série d'entrée Pr et de la série de sortie I (série d'entrée « idéale » pour le module de routage paramétré par le jeu $\{\theta_i\}_{i=1\dots n}$). Chaque étape (chaque jeu de paramètres testé) de la méthode « pas-à-pas » donne donc lieu à quatre opérations :

1. Calcul de la série Pr
2. Calcul de la série I

3. Calcul des ordonnées de l'hydrogramme unitaire
4. Calcul de la fonction objectif sur la période, en marche normale

Calcul de la série Pr

Cette série ne pose pas de problème à calculer, il suffit d'alimenter le module de production (défini par le paramètre θ_1 du jeu de paramètres testé) avec les pluies et ETP observées de la période de calage.

Calcul de la série I

Cette série est plus compliquée à obtenir car elle nécessite l'inversion de la fonction de routage/échanges non linéaire. La procédure que nous avons adoptée est largement inspirée de celle utilisée pour la « mise à jour » des états du modèle de prévision GR3P (Tangara, 2005).

Si l'on note θ_2 la capacité du réservoir de routage, α l'exposant de sa loi de vidange, et R_0 son niveau en début de pas de temps, on obtient comme expression du débit écoulé par le réservoir sur ce pas de temps :

$$Q_r(I, R_0) = Q_r(R'_0) = R'_0 \left(1 - \left[1 + \left(\frac{R'_0}{\theta_2} \right)^{\alpha-1} \right]^{-\frac{1}{\alpha-1}} \right) \quad (D.6)$$

$$\text{avec } R'_0 = \max \left\{ 0; R_0 + F(R_0) + \frac{9}{10} I \right\} \quad (D.7)$$

$$\text{et } F(R_0) = \theta_4 \left(\frac{R_0}{\theta_2} - \theta_5 \right) \text{ l'échange} \quad (D.8)$$

Le débit écoulé dans la branche directe vaut d'autre part :

$$Q_d(I, R_0) = \max \left\{ 0; F(R_0) + \frac{1}{10} I \right\} \quad (D.9)$$

Le débit total écoulé au cours du pas de temps est donc $Q = Q_r(I, R_0) + Q_d(I, R_0)$. On peut vérifier que c'est une fonction croissante de I et que les dérivées secondes $\frac{\partial^2 Q_r}{\partial I^2}$ et $\frac{\partial^2 Q_d}{\partial I^2}$ sont toujours positives ou nulles. Le courbe $Q(I | R_0 \text{ connu})$ a donc une convexité tournée vers le haut ce qui permet une inversion numérique assez rapide (cf. Figure D.2).

L'expression simple du débit Q_d sur la branche directe permet de trouver un majorant de I , le point $(I_{\text{majo}}, Q(I_{\text{majo}} | R_0))$ servant alors de pivot pour les cordes successives $\mathcal{C}_1, \dots, \mathcal{C}_n$

qui fournissent une série croissante I_1, \dots, I_n de minorants de la valeur I^* cherchée. À l'issue de ces itérations (lorsque la différence entre le débit cible Q^* et le débit $Q_n = Q(I_n | R_0)$ est inférieur à la précision sur le débit), on obtient l'estimation de I . Le niveau du réservoir à la fin du pas de temps vaut par ailleurs $R_1 = R'_0 - Q_r(R'_0)$, qui est la valeur de R_0 au pas de temps suivant, ce qui permet de continuer la procédure. Le seul problème dans l'inversion est qu'au

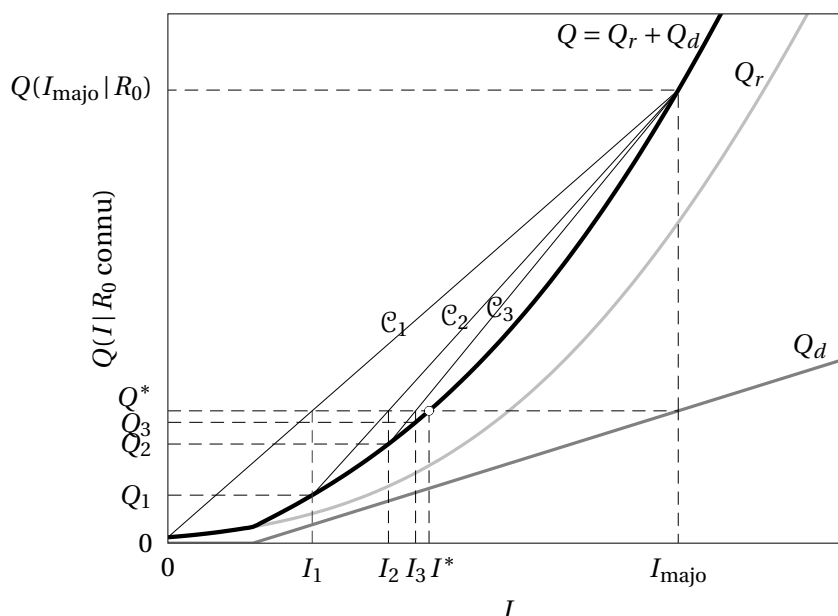


FIG. D.2: Procédure d'inversion de la fonction de routage à chaque pas de temps, permettant de trouver la valeur I^* telle que $Q(I^* | R_0) = Q^*$ (le niveau R_0 du réservoir étant connu).

premier pas de temps (ou à chaque fois que la chronique de débit reprend après une période de lacunes), on ne connaît pas le niveau R_0 du réservoir : il faut faire une hypothèse pour amorcer la procédure (problème de l'oeuf et de la poule!). Pour cela on suppose qu'à ce pas de temps tout le débit provient d'une récession (pas de débit sur la branche directe) et l'on inverse numériquement l'équation de vidange adimensionnelle $q = f(y)$ que l'on aura précédemment tabulée :

$$q = \frac{Q}{\theta_2} = y \left(-1 + [1 - y^{\alpha-1}]^{-\frac{1}{\alpha-1}} \right) \quad (\text{D.10})$$

où $y = \frac{R_1}{\theta_2}$ taux de remplissage du réservoir en fin de pas de temps ($0 \leq y < 1$).

Le niveau R_1 trouvé est utilisé comme valeur de R_0 au pas de temps suivant. Passés les premiers pas de temps, les couples $(R_0(t), I(t))$ convergent vite vers des valeurs donnant exactement le débit observé $Q(t)$.

Calcul des ordonnées de l'hydrogramme unitaire

Une fois obtenues les séries Pr et I (en pratique, on n'a besoin de recalculer qu'une seule des deux à chaque pas d'optimisation, le paramètre modifié affectant soit la série Pr soit la série I mais jamais les deux), on calcule la série d'autocorrélation $Pr \star Pr$ et la série de corrélation croisée $Pr \star I$ (Figure D.3). On calcule ensuite les transformées de Fourier de ces deux séries, puis la transformée inverse de leur quotient selon la méthode décrite dans la section D.2. L'hydrogramme unitaire est obtenu après lissage par une médiane mobile sur $2k + 1$ pas de temps, puis normalisation (Figure D.4).

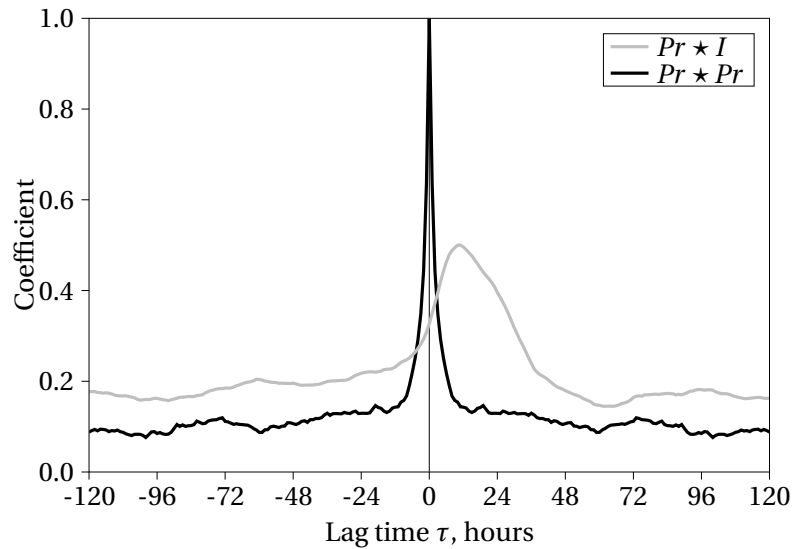


FIG. D.3: Fonction d'auto-corrélation de la production (noir) et fonction de corrélation croisée entre la production et l'entrée idéale du module de routage (gris).

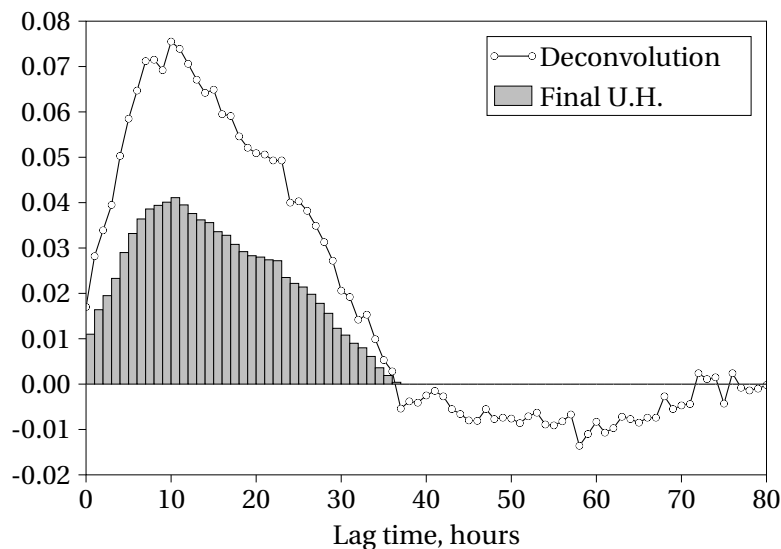


FIG. D.4: Exemple d'hydrogramme unitaire obtenu par déconvolution. L'évaluation discrète de la fonction de transfert (ligne brisée) est lissée par une médiane mobile puis normalisée après élimination des ordonnées négatives, pour obtenir l'H.U. discret final.

Calcul de la valeur de la fonction objectif

La dernière phase est le calcul de la valeur de la fonction objectif associée au jeu de paramètres courant de l'optimisation. Pour cela on procède en marche « normale » en calculant la production, puis en la faisant passer par l'Hydrogramme Unitaire précédemment identifié, puis par le module de routage en sens « direct ».

Contrôle

La procédure de calcul de la fonction objectif en contrôle suit le même schéma que celui décrit au paragraphe précédent, les ordonnées de l'hydrogramme unitaire étant récupérées à l'issue du calage.

D.3.2 Résultats sur l'échantillon de 1040 bassins versants

Distribution des critères en calage et en contrôle

La méthode proposée réduisant le nombre de paramètres effectivement optimisés (il n'y a plus de temps de base de l'hydrogramme unitaire), les résultats obtenus sont conformes à ce que l'on pouvait espérer : une légère baisse de performance en calage mais pas de perte en contrôle (cf. Figure D.5).

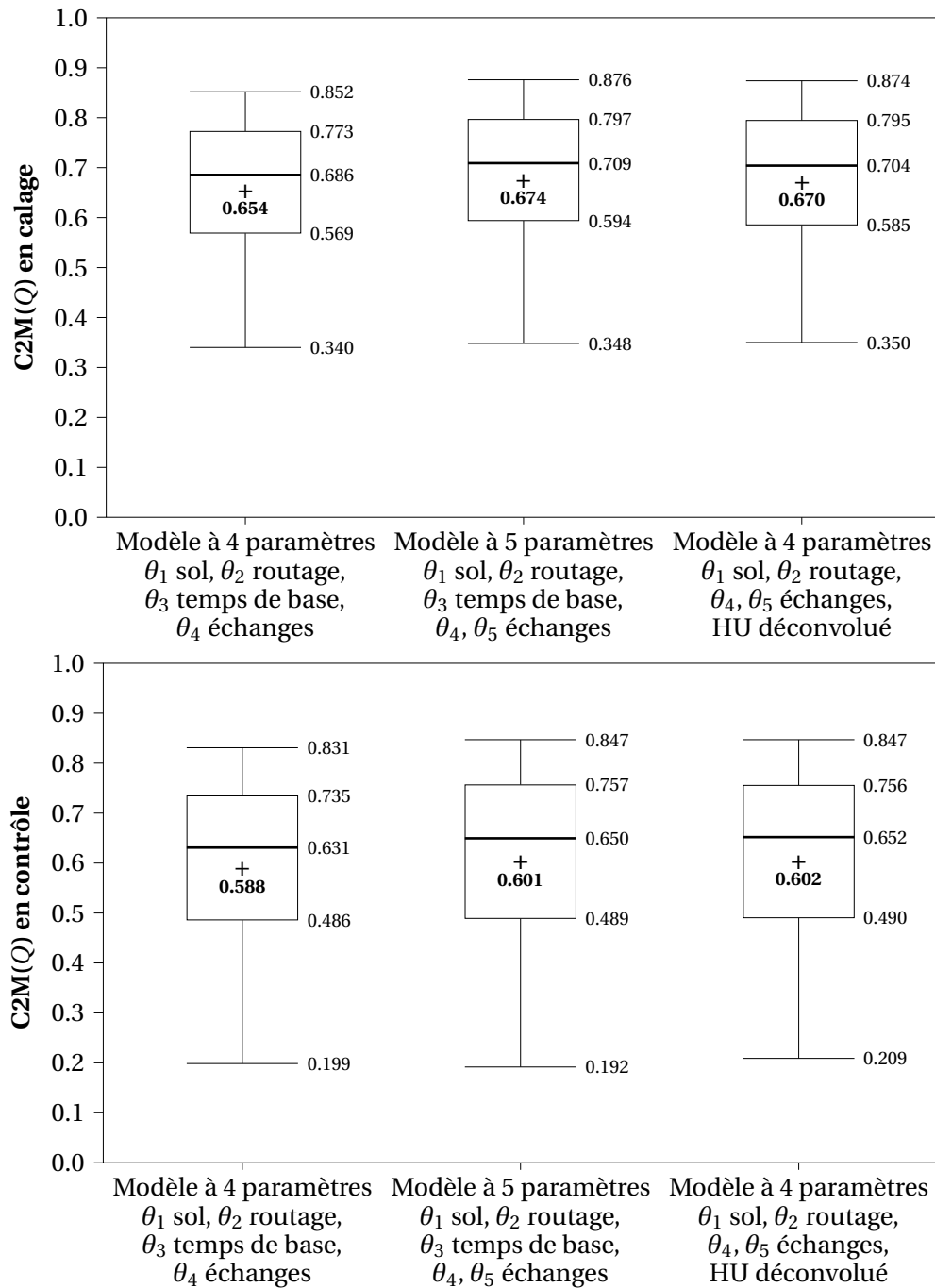


FIG. D.5: Comparaison des performances en calage et contrôle de différentes variantes du modèle horaire. Pour les modèles utilisant un temps de base θ_3 , on a représenté les distributions avec la meilleure forme d'hydrogramme trouvée (elliptique).

Distribution des paramètres

La modification de l'hydrogramme unitaire a peu d'effet sur les paramètres (cf. Figure D.6), excepté sur la capacité θ_2 du réservoir de routage qui a tendance à être nettement plus faible (en moyenne les 2/3 de la valeur d'origine, probablement parce que les récessions sont plutôt assurées par le transfert linéaire). L'amplitude θ_4 de l'échange et le seuil θ_5 ne sont pas affectés, et il n'y a finalement pas d'impact sur les valeurs interannuelles des lames

échangées (cf. Figure D.7).

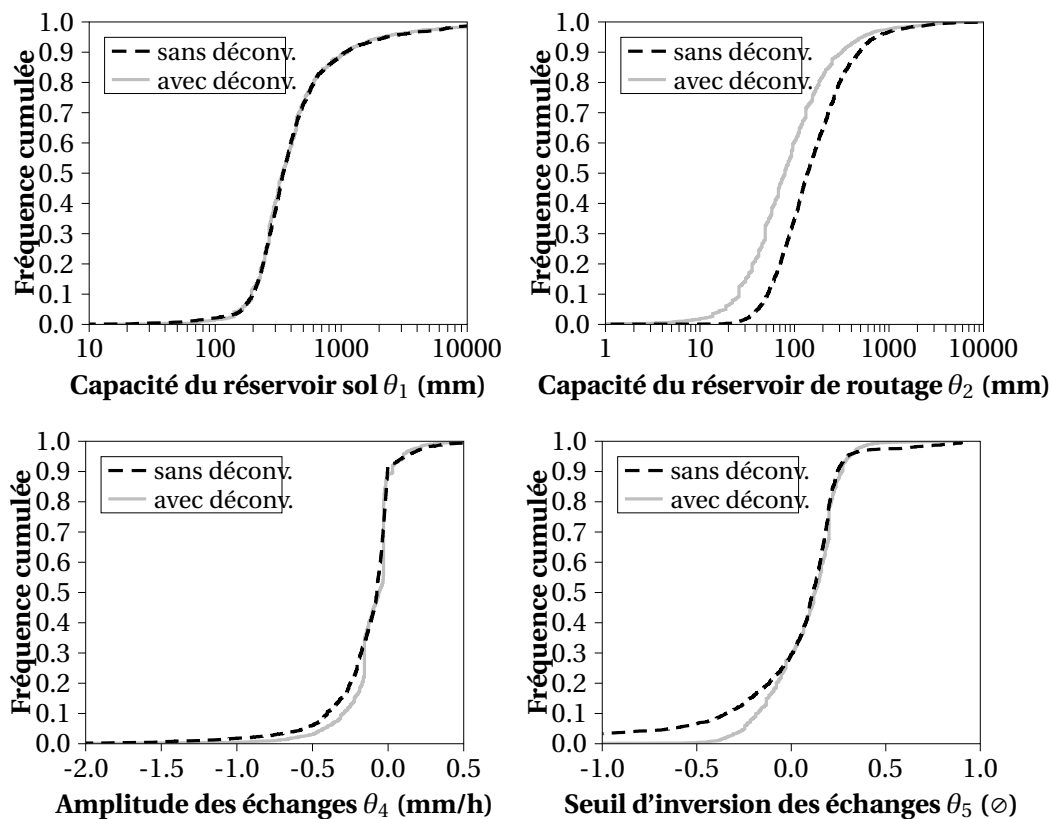


FIG. D.6: Impact de la méthode de déconvolution de l'hydrogramme unitaire sur la distribution des paramètres du modèle. L'impact le plus notable est une nette diminution de la capacité du réservoir de routage.

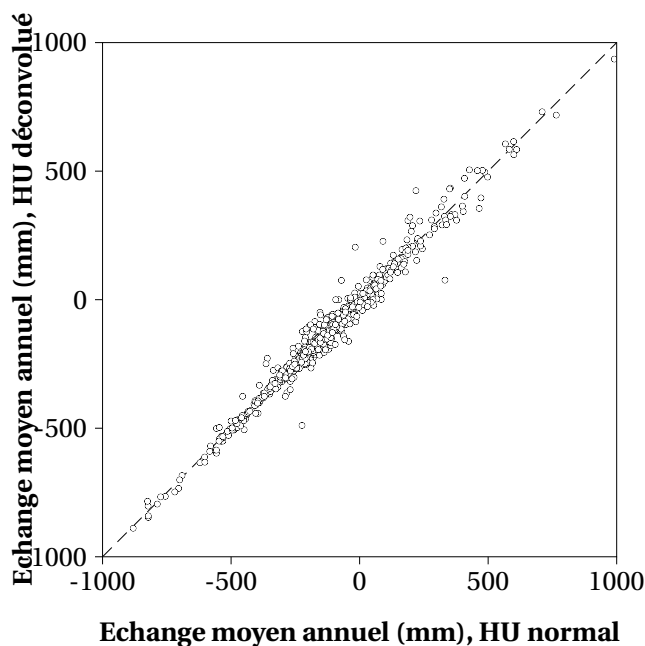


FIG. D.7: Effet de la méthode de déconvolution de l'hydrogramme unitaire sur les valeurs des échanges calculés par le modèle. On constate que la diminution de la capacité du réservoir de routage n'affecte pas le bilan interannuel calculé.

Il faut également préciser que l'exposant optimal pour la loi de vidange du réservoir de routage pourrait se trouver légèrement modifié, ce qui nécessite de poursuivre les tests.

D.4 Synthèse

Dans cette annexe, nous avons cherché à identifier la meilleure forme d'hydrogramme unitaire pour le modèle horaire. Il ressort principalement de ces tests que la paramétrisation de ce composant n'est pas très sensible.

Sur le strict plan du nombre de paramètres « optimisés » au sens classique du terme, la méthode d'identification de l'hydrogramme unitaire par déconvolution est plus parcimonieuse que l'imposition d'une forme *a priori* avec calage du temps de base. Elle évite théoriquement les problèmes d'équifinalité dans l'optimisation du routage linéaire et du routage non-linéaire, mis en évidence par la distribution du paramètre θ_2 avec et sans optimisation de l'hydrogramme unitaire. Le modèle à quatre paramètres basé sur cette méthode s'avère effectivement robuste, ses performances en contrôle étant équivalentes à celles du modèle à cinq paramètres. Malgré tout, l'apport reste assez limité et surtout ne justifie pas le coût calculatoire : calcul d'un corrélogramme (auto- ou croisé), de sa transformée de Fourier, et transformée inverse à chaque pas d'optimisation — opérations très coûteuses vue la taille des séries à la résolution horaire —, ce qui multiplie en moyenne le temps de calage par un facteur six !

Transformation en quantiles normaux



E.1 Origine et intérêt pour la transformation des débits

Au Chapitre 11, nous avons mentionné la nécessité de transformer d'une manière ou d'une autre les valeurs « brutes » des débits, pour réduire l'hétéroscédasticité des erreurs associées. Sur un plan pratique, on aime se ramener à une distribution plus ou moins symétrique et si possible pas trop éloignée d'une distribution normale. En effet, si la variable transformée ($T(Q)$) suit une loi normale, il y a de meilleures chances que les erreurs sur cette variable (i.e., $\epsilon_t = T(\hat{Q}_t) - T(Q_t)$) se rapprochent elles-aussi asymptotiquement d'une loi normale, pour laquelle la vraisemblance est simple à écrire (sous réserve de régler la question de l'autocorrélation, ce qui est une autre paire de manches).

Dans la littérature hydrologique, l'usage des transformations paramétriques de type Box-Cox (généralisation d'une transformation puissance permettant, par continuité, d'inclure la transformation logarithmique) est assez fréquent (e.g. *Thyer et al.*, 2002; *Yang et al.*, 2007). *Guillot* (1964) note d'ailleurs que pour bien des bassins, l'hypothèse log-normale est « *commode* » et suffisante pour corriger la dissymétrie des distributions de débits journaliers, et rendue selon lui « *d'autant plus tentante* » par la « loi de l'effet proportionnel » avancée par *Gibrat* (1936).

Cependant, pour certains bassins (bassins arides ou karstiques perdant une partie de leur écoulement par leur lit, cf. Figure E.1), il peut y avoir jusqu'à plusieurs mois consécutifs par an sans écoulement, et il faut trouver un moyen de prendre en compte ces débits (on peut choisir de les exclure du critère, mais dans ce cas on ne pénalise pas un modèle qui ne simule jamais de débits nuls, ou qui se trompe beaucoup sur les dates de cessation d'écoulement). Une méthode consiste à ajouter une petite valeur δ au débit, et donc calculer $\ln(Q_t + \delta)$ et $\ln(\hat{Q}_t + \delta)$. Cet expédient a l'avantage de la simplicité mais laisse ouvert le choix de la valeur

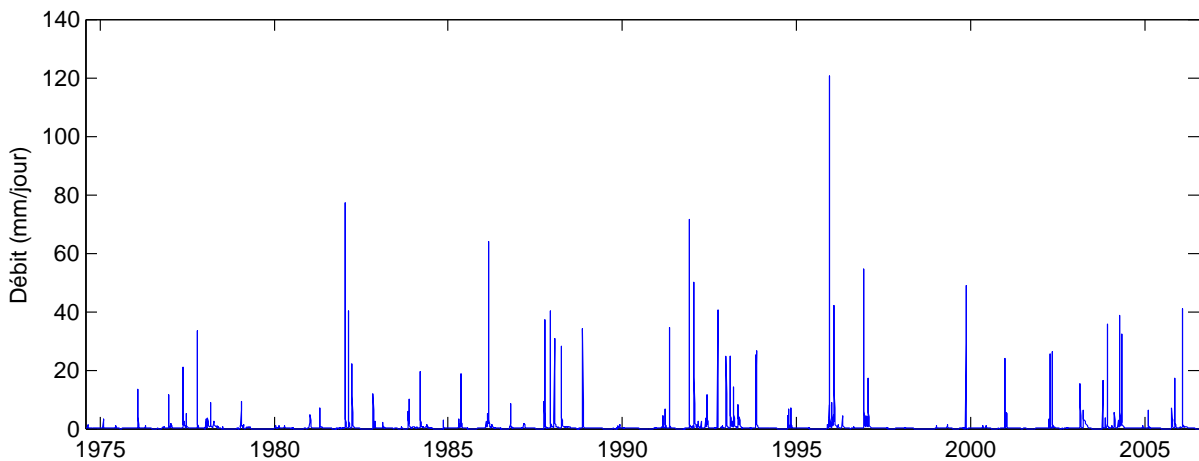


FIG. E.1: Hydrogramme des débits journaliers de la Canterrane à Terrats (Y0325010) entre 1975 et 2005. Les débits sont nuls pendant environ un quart de la chronique.

de δ . Prenons comme exemple les lames écoulées horaires qui sont écrites, dans les fichiers ASCII de notre base de données, à une précision (numérique) de quatre décimales en horaire ($p = 10^{-4}$ mm, c'est-à-dire que la valeur lue est toujours un multiple de p et qu'en particulier il n'y a pas de valeurs comprises strictement entre 0 et p) et trois décimales en journalier ($p = 10^{-3}$ mm).

- si δ est très petit devant la précision numérique, par exemple $\delta = 10^{-7}$ mm, alors on a une bonne sensibilité pour les faibles débits non-nuls mais lorsque le débit est nul, la transformation varie très fortement, les erreurs devenant potentiellement grandes. Par exemple, si à t le débit observé vaut $Q_t = p$ et le débit simulé vaut $\hat{Q}_t = \frac{p}{2}$ (erreur d'un facteur 2, le débit simulé pouvant lui varier avec des incréments bien plus petits que p), l'erreur correspondante vaut $\epsilon_t = \ln(\frac{p}{2} + \delta) - \ln(p + \delta) \simeq -\ln 2$ (soit -0.69). Si au pas de temps suivant nous avons une cessation d'écoulement à la précision numérique, i.e. $Q_t = 0$, et que le débit simulé reste environ égal à $\frac{p}{2}$, alors l'erreur vaut $\epsilon_t = \ln(\frac{p}{2} + \delta) - \ln(\delta) \simeq +6.21 \dots$ ce qui correspond à une erreur d'un facteur 500 sur une valeur quelconque de débit non-nul!
- si δ est grand devant la précision numérique, alors ce problème ne se pose plus, mais cette fois on n'a plus de sensibilité pour les faibles valeurs.

Ainsi, on pourrait penser que le choix d'une valeur de δ sensiblement égale à la précision numérique p résout le problème et permet de pénaliser de façon juste les erreurs en étiage. Cependant, c'est faire l'hypothèse implicite que *pour tout bassin versant*, le débit minimal pouvant être atteint est plus ou moins le débit nul... mais c'est là une vision d'hydrologue exclusivement de surface (voir encore *Guillot, 1964*, qui considère « *inacceptable* » le fait de translater l'origine de l'échelle logarithmique à une valeur positive)! *Lang (2007)* aborde ce problème pour le choix d'une loi statistique d'ajustement des étiages, mais il se pose de façon tout aussi aigüe pour le choix d'une fonction objectif.

Considérons l'hydrogramme de la Figure E.2. Il s'agit du bassin versant de la Moder (Schweighouse-sur-Moder, A3301010), dont les étiages sont notoirement soutenus par la nappe des Grès du Trias. Nous voyons clairement sur la distribution des débits (Figure E.3) que la valeur minimale observée (lame écoulée de 0.242 mm sur une journée) est bien plus grande que la précision numérique ($p = 10^{-3}$ mm). Plus que l'erreur sur le débit ou sur une fonction puissance du débit, une mesure de la qualité de la simulation des extrêmes serait plutôt l'erreur sur leurs probabilités de dépassement (pour les étiages) ou de non-dépassement (pour les crues) : ici, un modèle s'autorisant à simuler des débits journaliers de l'ordre de 0.1 mm devrait être sévèrement pénalisé, car cette gamme de débit est physiquement très improbable étant donnée l'importance du débit de base.

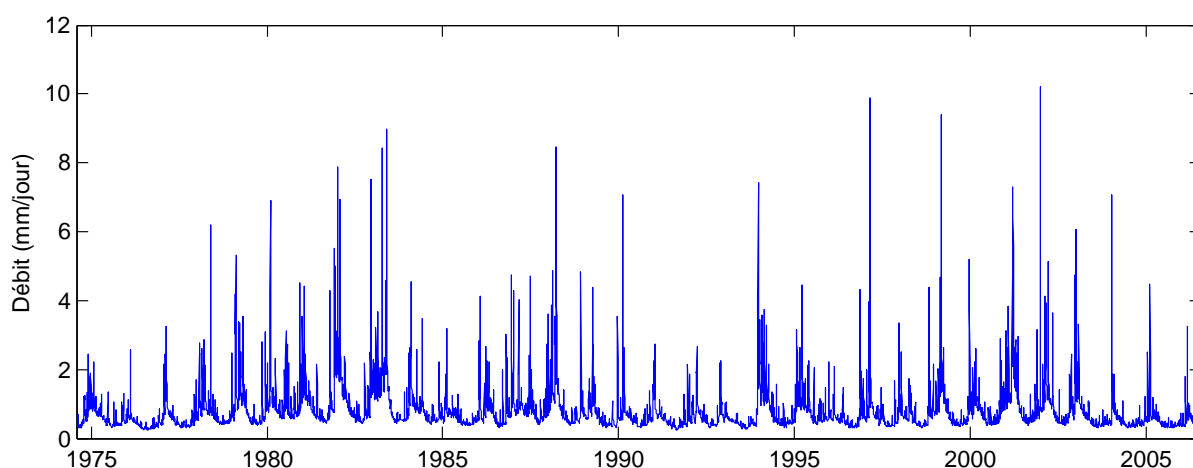


FIG. E.2: Hydrogramme des débits journaliers de la Moder à Schweighouse-sur-Moder (A3301010) : Les étiages sont fortement soutenus par la nappe des Grès Vosgiens (Trias Inférieur).

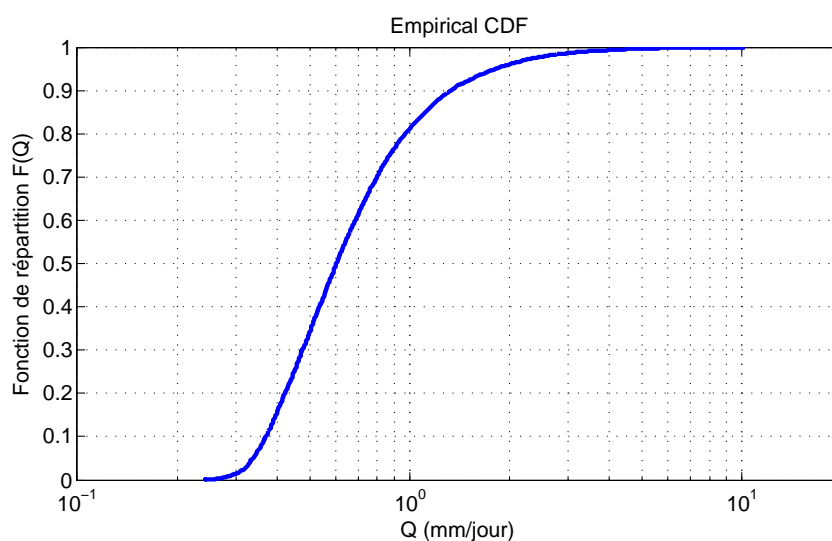


FIG. E.3: Fonction de répartition empirique des débits de la Moder à Schweighouse-sur-Moder (A3301010). L'échelle des abscisses (débits) est logarithmique.

Une dernière possibilité est donc de choisir δ en fonction des caractéristiques d'étiage, en ayant recours à une règle de pouce (par exemple une fraction du QMNA5 ou du module, comme par exemple dans *Perrin, 2000*). Mais quitte à opérer empiriquement, la transformation présentée dans cette annexe est intéressante car elle permet de s'affranchir de la forme logarithmique.

E.2 Mise en oeuvre

L'idée de la transformation par quantile normaux (*Kelly et Krzysztofowicz, 1997*) est précisément de remplacer chaque valeur par sa fréquence dans la distribution empirique, ou plus exactement par le quantile de même fréquence dans une distribution connue, en l'occurrence la loi normale $\mathcal{N}(0, 1)$, selon le principe illustré à la Figure E.4.

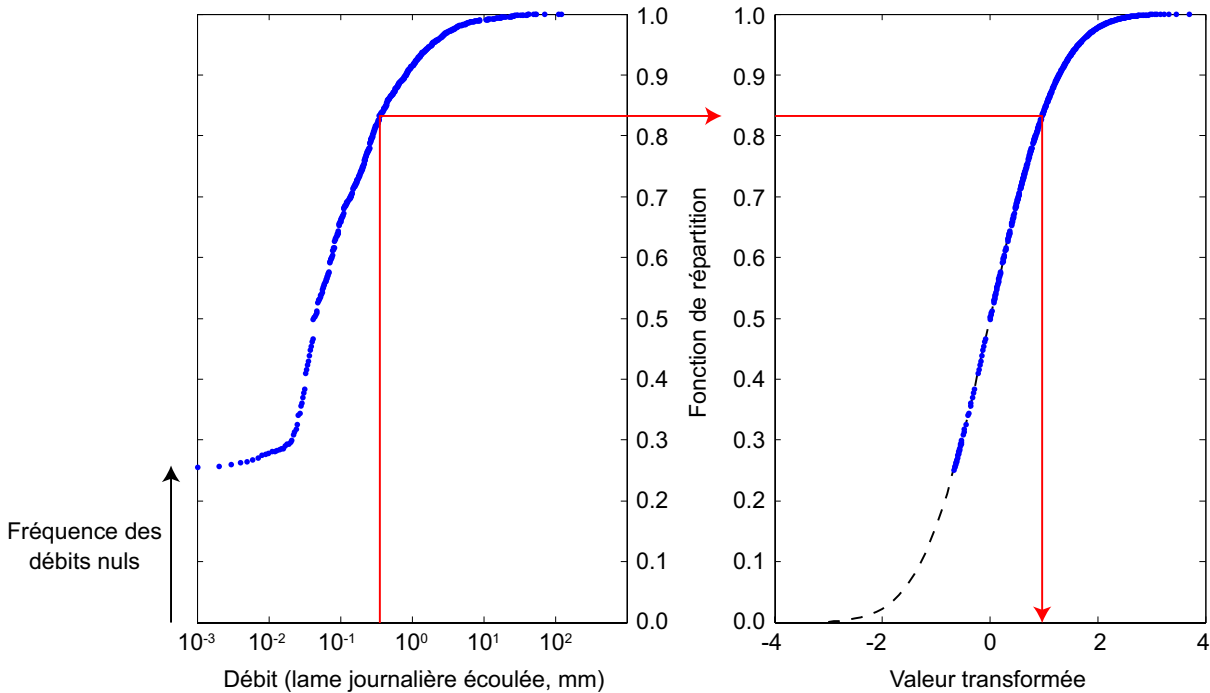


FIG. E.4: Principe de la transformation en quantiles normaux (La Canterrane à Terrats, Y0325010). Chaque valeur est associée à sa fréquence de dépassement dans la distribution empirique (graphique de gauche), puis à la déviation de même fréquence dans la loi normale $\mathcal{N}(0, 1)$ (droite).

Sur cette figure, nous voyons que nous avons déjà résolu le problème des valeurs nulles : si les débits nuls sont présent avec une fréquence cumulée $F(0)$, on associe à chaque valeur nulle la déviation normale de même fréquence cumulée. Une valeur un tout petit peu supérieure à zéro (par exemple égale à la précision numérique p) se voit affecter une valeur transformée un peu plus grande, la déviation correspondant à $F(p)$, et ainsi de suite de façon continue. Si on a N valeurs observées, on obtient la correspondance valeur–fréquence en comptant

le nombre N^* de valeurs uniques, rangées par ordre croissant $\{Q_i^*\}_{i=1\dots N^*}$, et $\{w_i\}_{i=1\dots N^*}$ leurs nombres d'occurrence (de sorte que $\sum_{i=1}^{N^*} w_i = N$, avec $N^* < N$ du fait de la précision numérique limitée). Puis on leur associe les fréquences :

$$F_i(Q_i^*) = \frac{(\sum_{j \leq i} w_j) - 0.375}{N + 0.25}$$

Il existe plusieurs autres formules possibles assez équivalentes pour de grandes valeurs de N^* , l'idée étant que les première et dernière valeurs uniques ne correspondent évidemment pas aux fréquences 0 et 1. Pour transformer une valeur de débit Q quelconque, nous cherchons l'indice i tel que $Q_i^* \leq Q < Q_{i+1}^*$, et interpolons linéairement pour trouver la valeur

$$F(Q) = F_i + \left(\frac{F_{i+1} - F_i}{Q_{i+1}^* - Q_i^*} \right) \cdot (Q - Q_i^*)$$

E.3 Modélisation des queues de distribution

Cette transformation empirique est donc extrêmement simple et, au premier abord, non-paramétrique. Cependant il faut tenir compte du fait que nous voulons l'utiliser à la fois pour transformer les débits *observés*, mais aussi les débits *simulés*, dont les valeurs peuvent se situer en dehors de la plage observée (même si nous aimerions éviter cette situation !). Il faut donc spécifier :

- dans tous les cas, un modèle pour la queue supérieure de distribution ;
- dans les cas où le plus faible débit observé n'est pas nul, un modèle pour la queue inférieure.

Constatant que la loi log-normale est une première approximation acceptable, nous modélisons la distribution des valeurs supérieures (resp. inférieures) à un seuil de crue Q_C (resp. d'étiage Q_E) par une distribution exponentielle ajustée sur $\ln Q$ (c'est-à-dire, une loi puissance, cf. Figure E.5) :

$$\begin{cases} F_C(Q | Q \geq Q_C) = 1 - e^{-\lambda_C (\ln Q - \ln Q_C)} & \text{pour les crues,} \\ F_E(Q | Q \leq Q_E) = e^{-\lambda_E (\ln Q_E - \ln Q)} & \text{pour les étiages (si besoin)} \end{cases}$$

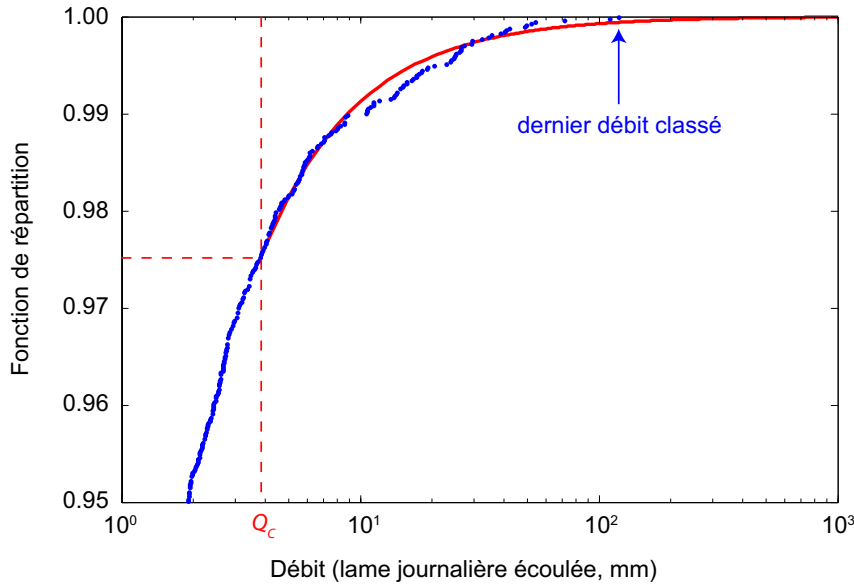


FIG. E.5: Ajustement d'une loi modélisant la queue supérieure de la distribution, de façon à pouvoir transformer des valeurs supérieures au plus fort débit observé (dernier débit classé). Dans le cas où la plus faible valeur de débit observé (premier débit classé) est non-nulle, on ajuste une loi similaire pour les étiages.

Le seuil Q_C et, le cas échéant Q_E , sont choisis de façon à pouvoir estimer robustement les paramètres λ_C et λ_E : on prendra par exemple les quantiles 0.975 et 0.025. Ces paramètres sont ajustés par la méthode des moments, ici simplement l'espérance de la loi exponentielle :

$$\left\{ \begin{array}{l} \frac{1}{\lambda_C} = \frac{1}{\sum_{Q_i^* \geq Q_C} w_i} \sum_{Q_i^* \geq Q_C} w_i (\ln Q_i^* - \ln Q_C) \\ \frac{1}{\lambda_E} = \frac{1}{\sum_{Q_i^* \leq Q_E} w_i} \sum_{Q_i^* \leq Q_E} w_i (\ln Q_E - \ln Q_i^*) \end{array} \right.$$

Drainage density and streamflow : a new look at an old debate

Nicolas Le Moine et Vazken Andréassian
Article soumis à *Water Resources Research*

Abstract

This paper sheds new light on a thirty years old debate on the hydrological relevance of the Drainage Density metric. While Drainage Density is a widely recognized catchment descriptor, Dingman (1978) had suggested that its reputation was probably overrated, because of the way it was classically computed and because of the limited hydrological information carried by this index. After reviewing Dingman's critics and the counter arguments that were opposed to him, we present an automatic approach to compute an Effective Drainage Density as advocated by Dingman. Then, we use a large dataset of 930 French catchments, which allows us to confirm empirically that Drainage Density is well correlated with the 10-year return period flood, while it has no explanatory power for low flows and mean flow. A detailed analysis on a subset of seven catchments permits to discuss the physical reasons of these results.

F.1 Introduction

The Drainage Density (DD) of a catchment can be defined as $D = L/A$ [L^{-1}] where L is the total length [L] of streams over a catchment of size A [L^2] (Figure F1). Introduced in hydrology by Horton (1932, 1945), this index has rapidly become popular in the hydrological world, being used as a building block of many empirical and semi-empirical formulas (Gardiner and Gregory, 1982 ; Wharton, 1994). But DD has also been criticized on three grounds :

- first, as a *scale-dependent* metric (Langbein, 1944 ; Gregory and Walling, 1973) : its value may indeed vary depending on the scale of observation ;
- second, as a *time-varying* characteristic (Gregory and Walling, 1968 ; Gregory and Gardiner, 1979) : in some conditions, its value may vary according to the wetness of the catchment ;
- third, as a metric of *limited hydrological relevance* (Dingman, 1978) : although DD has generally been considered as an informative hydrogeomorphological descriptor, its use has lead to numerous disappointments. Gregory and Gardiner (1975) consider it to be '*one of the most promising*' but at the same time '*one of the most elusive variables in the morphometric analysis of the drainage basins*'.

Determining whether or not the Drainage Density metric can be useful in interpreting the hydrology of a catchment first requires examining in detail these three possible flaws. It then requires choosing an appropriate technique to estimate automatically drainage density values, in order to be able to confront them to the mean-, high- and low- flows estimates of a large catchment dataset.

F.2 The three flaws of the drainage density metric

F.2.1 Flaw 1 : scale-dependence

The issue of scale dependence of the DD metric has been treated exhaustively in the literature. All authors have underlined that only DD values obtained from maps of identical scale could be compared. When working in a two-dimensional space such as a map, descriptors having the dimension of a surface [L^2] have a bounded measure when the length of discretization dx tends to zero. In contrary, descriptor representing a length (such as the cumulative length of the hydrographic network) tend to infinity when $dx \rightarrow 0$. This relations have been widely investigated (see e.g. *Mandelbrot, 1982 ; Hjelmfelt, 1988 ; Rodriguez-Iturbe and Rinaldo, 1997*)

But actually, what is needed is not a method that measures blue lines over a catchment at a given scale, but one that evaluate the *drainage efficiency of the river network* (*Dingman, 1978*), i.e. the capacity of the hydrologically active valley network to transport water towards the catchment outlet. As outlined by Dingman indeed, '*the detailed channel of a stream, however sinuous, has a generally insignificant relationship to the length of either overland or subsurface flow routes to channels*' (p. 1183). This is why we propose in section 3 an automatic method to compute Dingman's effective Drainage Density.

F.2.2 Flaw 2 : can we trust the blue-lines found on maps ?

The second critic of the DD metric is related to the hydrological relevance of blue-lines : even at a given map scale, there is no agreement on how blue-lines should be interpreted (see e.g. the most recent review by *Wharton, 1994*), and three opinions co-exist : (i) the pessimistic view is that there are so many inconsistencies and errors on the mapping of small rivers that their uncertainty turn the DD metric worthless (see e.g. *Schneider, 1961*) ; (ii) a second view is that map-derived DD figures are of little interest because the river network is essentially dynamic (streams are seen as variable objects which can move up valleys and hollows during storm events, or down during drought events, in line with the streamflow genesis views of *Hewlett and Hibbert, 1963, 1967 ; Gregory and Walling, 1968 ; Dunne and Black, 1970*) ; (iii) third, the optimistic view is that of hydrologists who consider that the topographic maps represent regionally the best available synthesis of hydrographic knowledge, and see thus some hydrologic potential in the DD metric.

Can we take position on this debate? We consider that the view of the first group is excessively pessimistic. Of course, the very nature of field-data collection imposes limits on

the use we can make of the extent of blue-lines networks (see the review by *Chorley and Dale*, 1972). Unavoidably, some subjectivity remains in the mapped river network. But statistically, the blue-lines do represent flowing channels, and thus, they encapsulate (summarize) a wealth of field information, obtained by surveyors and also by local inhabitants. Then, we do not see the 'dynamic' point of view as necessarily incompatible with the 'optimistic' point of view : the dynamic nature of river networks does of course induce some uncertainty in the computation of DD (the question being that of the wetness conditions under which a given channel was surveyed). But in many countries, where maps have been reedited several times, the blue-lines integrate surveys, field-visits and interviews which span over decades, so that it is rather reasonable to say that they represent flow in 'average' wetness conditions.

F.2.3 Flaw 3 : sheer hydrological informativity

The last and the most serious of the critics made towards the DD metric is that of its limited hydrological relevance. *Dingman* (1978) reaches this conclusion after reviewing the physical reasoning underlying the assumption that drainage density was an integrated descriptor of streamflow. He concludes that the general utility of DD as a predictor of streamflow is questionable and that there are both theoretical and empirical evidence that it has no significant explanatory power for low-flows and average flows. From his point of view, DD is '*of much more limited value as an independent variable for predicting hydrologic characteristics than has been widely assumed in the literature*' (p. 1183), and its use should be limited to the interpretation of high flows statistics. He also brings strong arguments why some care should be taken to compute truly 'effective' drainage density as mentioned previously.

F.2.4 Scope of the paper

In this paper, our aim is to address the critics to DD as a hydrologically relevant metric raised by *Dingman* (1978). The reviews of literature (*Gardiner and Gregory*, 1982 ; *Wharton*, 1994) show that most discussions have been based on an extremely limited corpus of data : rarely more than ten or twenty catchments. This is all the more surprising as *Langbein* (1944) had derived topographic characteristics for 340 US catchments, and that *Benson* in 1960 had looked for the physical determinants of floods over 164 catchments. Unfortunately, these studies remained exceptions, and we believe that there is time to confront the theory with field evidence at a larger scale.

In section 3, we describe the approach adopted to compute automatically a robust and effective drainage density, in line with *Dingman's* recommendations. We apply then this

methodology to the whole of France. The next step, in section 4, consists in confronting our effective DD estimates to the streamflow statistics of a large set of 930 French catchments. Last, in section 5 we zoom to a case study concerning a small group of seven neighbor catchments, similar in all points but differing by their drainage density, in order to discuss the physical reasons underlying the hydrological relevance of drainage density.

F.3 Automatic computation of Effective Drainage Density

F.3.1 How to measure effective drainage density ?

Measuring DD is apparently extremely easy : it just requires measuring the total length of continuous blue lines on a topographical map. But for a large number of catchments, by hand, it is an impossible task. Moreover, the question of how to extract a measure of drainage efficiency rather than a sheer density of a scale-dependent blue-line raised by *Dingman* (1978) remains.

We sought an automatic approach to derive an index of effective DD over the whole of France (550,000 km²). The best alternative we found consisted in using a 75-m Digital Elevation Model (DEM) to extract automatically the drainage network, which is then compared with a digitized version of the blue-line network. As shown on Figure F2, the extent of the DEM-derived network depends on the source area size (see e.g., *Montgomery and Foufoula-Georgiou*, 1993 ; *Moussa and Bocquillon*, 1996) : estimating the local Drainage Density (over a small homogeneous area) requires identifying the source area threshold that will yield the best match between the blue-lines network and the DEM-derived network. This could be done visually, but the result would then be somewhat subjective. For an automatic procedure, we needed an adequate metric to compare the shape and extent of both networks : over each small homogeneous area, the best guess of the local source area threshold will correspond to the threshold that yields the highest level of similarity between the two networks.

We developed an original approach to test the similarity of the map-derived (digitized blue-lines) and DEM-derived networks. It is based on the Hausdorff metric (*Hausdorff*, 1919), which is a very popular technique in image processing (see e.g. *Rucklidge*, 1996) and presents a number of practical advantages over alternative approaches :

- (i) by definition, it does not require a one-to-one correspondence between the DEM-derived network and the map-derived network ; i.e., we do not have to worry about minor

- local discrepancies between the two networks. The algorithm rather provides a robust identification of the average pattern and will not be fooled by singularities ;
- (ii) for this reason, it is not very sensitive to both DEM resolution and discretization level of the ‘blue lines’ (i.e., edgy lines will still provide good results, as long as the position of their extremities is reasonably estimated).
- (iii) subsequently, it avoids all the problems linked with multiple channels (braided rivers for example) and meandering. Both meanders and multiple channels will be located in the very neighborhood of a correspondent, unique channel of the DEM-based network (indeed, flow convergence has been ensured by the correction of the DEM using the *Garbrecht and Martz* (1997) and *Martz and Garbrecht* (1998) algorithms), locally yielding a very small Hausdorff’s distance.

F.3.2 Extraction of drainage networks from digital elevation models

The drainage network was established using a 75-m DEM and a D8 approach. The DEM was corrected using implementations of *Martz and Garbrecht* (1998) ‘breaching’ algorithm in order to remove flow sinks, and *Garbrecht and Martz* (1997) ‘double gradient’ algorithm in order to treat flat areas. There are numerous alternative methods that could have been used instead, but this one had the advantage of producing flow convergence and avoiding parallel flow lines that could have been a drawback for the subsequent operations (see 3.3).

F.3.3 How to measure the likeness of the extracted and observed networks ?

We need to define what we call ‘likeness’ in order to select the best extraction threshold. For that, we use the Hausdorff distance, which is a generalization of the Euclidean distance for two sets of point.

Let X and Y be two sets of points. The Hausdorff distance between those two sets is derived from the classical Euclidean distance $d(.,.)$ between two points, as follows :

$$H(X, Y) = \max \left\{ \underbrace{\sup_{x \in X} \inf_{y \in Y} d(x, y)}_{h_Y(X)}, \underbrace{\sup_{y \in Y} \inf_{x \in X} d(x, y)}_{h_X(Y)} \right\} \quad (\text{E.1})$$

This distance is quite easy to understand intuitively. It is the maximum of two terms, also called semi-distances :

- the semi-distance $h_Y(X)$ which is the greatest distance of an element x in X to its nearest neighbor in Y ,

- the semi-distance $h_X(Y)$ which is, conversely, the greatest distance of an element y in Y to its nearest neighbor in X .

Practically, the semi distance $h_Y(X)$ corresponds to the size of the smallest buffer zone around Y that is required to include all points of X . It is obtained by first computing an Euclidean Distance Map (EDM) to the set Y , which is then sampled at each point of X . Then we exchange the roles of X and Y and we apply the same procedure in order to compute the other semi-distance $h_X(Y)$.

We can see from Figure F3 that usually $h_Y(X) \neq h_X(Y)$: the semi-distances are not symmetric, whereas the Hausdorff distance is ($H(X, Y) = H(Y, X)$).

F.3.4 Application to the observed and extracted networks in a subcatchment

Now we can explain how to find the best extraction threshold in a given subcatchment, i.e. the average source area in this subcatchment. The two sets whose distance we measure are respectively the observed network (O) and the extracted network at threshold S (noted $E(S)$). We want to find the surface extraction threshold S_{opt} that minimizes the Hausdorff distance $H(O, E(S))$, as shown in Figure F4 :

- when the extraction threshold S is much lower than S_{opt} (top), the extracted network $E(S)$ is much more developed than the observed network (at the limit, the extracted network would expand until it covers the entire subcatchment). Apart from local discrepancies due to the precision of the DEM, the observed network is a subset of the extracted network : the semi-distance $h_{E(S)}(O)$ is very small and the Hausdorff distance $H(O, E(S))$ is equal to the semi-distance $h_O(E(S))$,
- conversely, when the extraction threshold S is much higher than S_{opt} (bottom), the extracted network $E(S)$ is much less developed than the observed network (at the limit, the extracted network would shrink to the outlet). The extracted network being now a subset of the observed network, the semi-distance $h_O(E(S))$ is very small and the Hausdorff distance $H(O, E(S))$ is then equal to the semi-distance $h_{E(S)}(O)$.

Obviously, there is a situation $S = S_{\text{opt}}$ between those two extremes where the observed and extracted networks have a similar development. We can show that this situation corresponds to $H(O, E(S)) = h_{E(S)}(O) = h_O(E(S))$. Indeed, the semi-distance $h_{E(S)}(O)$ is a monotonically increasing function of S : for each point of the observed network O , it becomes more difficult to find a neighbor in $E(S)$ as this extracted network shrinks ; conversely, this 'shrinking' effect makes the semi-distance $h_O(E(S))$ a monotonically decreasing function. As a consequence, the optimal threshold S_{opt} is located at the intersection of the two curves (Figure F5). Of course, even at extraction threshold S_{opt} there are still differences for first order streams

(some of them exist in the observed network and not in the extracted network, or conversely), but the overall drainage pattern is well matched : the threshold S_{opt} gives the order of magnitude of source areas in the subcatchment, which should depend on geomorphological thresholds (see e.g., *Montgomery and Dietrich*, 1989).

S_{opt} [L^2] can be defined as the typical size of a catchment at the beginning of a first-order stream in a given area (though it is obtained in a much convenient way than mapping each observed source on the DEM). The lower this value, the higher the drainage density. Thus, if we look at $\frac{1}{S_{\text{opt}}}$ [L^{-2}], we have a measure of DD which is dimensionally much more satisfying than the classical definition [L^{-1}]. This is due to the fact that our definition does not rely on any measure of *length*, but only on *Euclidean distance between pairs of points sampled from two sets*. Figure F.6 illustrates this fact : when we degrade the precision of the polylines describing the ‘actual’ stream network, the Hausdorff distance between the extracted network and the polyline set is almost not affected : the algorithm can operate even on very edgy lines as long as their extremities are approximately correct. In contrary, the classical DD would be strongly affected by such a degradation.

F.3.5 Mapping the perennial source area at the country scale

In order to map what we have defined as the perennial source area or its inverse (DD), we need to apply the algorithm on a set of subcatchments (topographical mesh). We define this topographical mesh according to the number of sources of the observed network. We use an algorithm that goes through the DEM-based (D8) drainage scheme from upstream to downstream, counting the number of sources upstream a given point : each time we found more than n sources in a subcatchment, we split it. The value of n needs not to be too small, so that the channel pattern identification remains robust, nor too high so that the subcatchments in which we identify the optimal threshold remain quite homogeneous. We tried with several values of n ranging between 10 and 20 ; a satisfactory value was $n = 12$: the resulting map is given in Figure E.7, together with a geological map of France in order to show the clear link with lithology. The full hydrographic network is also given in Figure E.8.

In Figure E.9 we show that, though the classical definition of DD depends on the resolution of the blue lines, there is a very good correlation between the two definitions at a given resolution.

Now that we have been able to compute a more robust estimate of drainage density, we wished to show how it could be used to narrow the predicted range of several streamflow

characteristics.

F.4 Confronting streamflow attributes to drainage density on a large catchment set

In this section, we use a dataset of 930 French catchments (see Figure F.10), over which we analyze the correlations existing between various streamflow characteristics and catchment descriptors, among which drainage density. We use classical log-regressions between the target streamflow Q and the potential *explicans* X_1, \dots, X_k :

$$\log(Q) = \beta + \sum_k \alpha_k \log(X_k) \quad (\text{E2})$$

It is worth noting that the choice of Q (raw discharge) or Q/A (specific discharge) has no importance when working with log-regressions (since $\log(Q/A) = \log(Q) - \log(A)$) as long as we keep area A among the explanatory variables (it will just change the corresponding α_A by ± 1). We wish to do so because area A can also be an *explicans* for timing effects, so our explanatory variables will be :

- catchment area A (km²),
- mean annual rainfall P_m (mm),
- mean annual potential evapotranspiration PE_m (mm),
- spatially averaged 10-year return period daily rainfall P_{10} (mm),
- drainage density DD (km⁻²)

We also derived two secondary variables :

- aridity index $AI = PE_m/P_m$,
- rainfall irregularity index $RII = P_{10}/P_m$.

We tested all regressions involving up to five explanatory variables in order to predict three characteristic streamflows :

- mean annual streamflow,
- 5-year return period minimal monthly flow (low flow),
- 10-year return period peak flow (flood flow).

F.4.1 Mean annual streamflow

Table F.1 summarizes regression results for the mean annual flow. As we can see, few importance is given to DD variable in the best regressions. If we were to pick a simple and efficient regression, we would probably pick the one involving catchment area A , mean annual P_m and mean annual potential evapotranspiration PE_m ($R^2 = 0.945$). We know that

most of long term water balance models (e.g., *Schreiber*, 1904 ; *Ol'dekop*, 1911 ; *Turc*, 1954 ; *Pike*, 1964) use these variables only. Though our formulation seems quite 'brutal', we can show that it is close to a classical water balance equation when $\frac{P_m}{PE_m}$ is not too far from unity (as it is usually the case) :

$$\begin{aligned}
 Q &= k \cdot A \cdot P_m^2 \cdot PE_m^{-1} \\
 \frac{Q}{A} &= k \cdot PE_m \left(\frac{P_m}{PE_m} \right)^2 \\
 \frac{Q}{A} &= k \cdot PE_m \left(1 + \frac{P_m - PE_m}{PE_m} \right)^2 \\
 \frac{Q}{A} &\simeq k \cdot PE_m \left(1 + 2 \frac{P_m - PE_m}{PE_m} \right) \\
 \frac{Q}{A} &\simeq k (PE_m + 2 (P_m - PE_m)) \\
 \frac{Q}{A} &\simeq k (P_m - PE_m)
 \end{aligned} \tag{F.3}$$

F.4.2 Low flows

Table E2 summarizes regression results for the 5-year return period minimal monthly flow. The results are much less satisfactory, since the best correlation coefficient is $R^2 = 0.611$. However, we can see that DD does not bring any substantial information into the regression : as for mean annual streamflow, the most parcimonious solution would be to use catchment area A , mean annual rainfall P_m and mean annual potential evapotranspiration PE_m ($R^2 = 0.598$).

F.4.3 Flood flow

Table E3 summarizes regression results for 10-year return period peak flow. This time, DD seems to be more informative since it is present in the eight best regressions with a quite important weight. The higher the value of DD, the higher the decadal flood peak, *all other things being equal*. Were we to use a suitable regression, we could for example pick the following one ($R^2 = 0.827$, see Figure E11) :

$$Q_{i \times 10} \propto A^{0.89} P_{10}^{1.14} DD^{0.33} AI^{-0.65} \tag{F.4}$$

This interesting result raises at least two questions : why is drainage density relevant for explaining flood discharges, and flood discharges only ? We propose to give some insight to this result by looking at a case study, which will make things more concrete.

F.5 Detailed analysis of a case study in Normandy

F.5.1 Seven catchments where lithology induces strong differences of drainage density

In this section we present a subset of seven catchments located in a region which is climatologically homogeneous but which exhibits, for geological reasons, high variations in DD (see Figure F.12). We can thus exemplify the sensitivity of flood flows to the geologically-controlled drainage density, and also try to understand why low flows and mean flows do not correlate with DD. Basically, in this area, DD can be interpreted as a measure of the connection between surface- and groundwater : the more permeable the lithology, the more important the groundwater drainage and the lower the drainage density. Out of these seven catchments, one is located on impervious lithology, and this is reflected in high DD (0.27 km^{-2}) : this is the Epte river, which will serve as a reference. The six others are located on extremely permeable chalk, and they consequently exhibit low DD values ($0.03\text{--}0.06 \text{ km}^{-2}$).

F.5.2 Effects on flow quantiles

Mean annual flow

Basic hydrographs of specific discharge Q/A , such as those shown in Figure F.13, clearly shows why DD cannot be an *explicans* for mean annual flow : low DD values do indicate an important groundwater-to-surface water connection but it does not allow predicting whether groundwater transfers mainly occur *from* the groundwater body *to* the stream, or the reverse. Extreme cases are provided by the Cailly and Commerce rivers, the first one with a net gain ($\frac{Q_m}{P_m} = 0.46$) and the second one with a net loss ($\frac{Q_m}{P_m} = 0.06$). As a reference for this area, we can take the 'conservative' Epte catchment ($\frac{Q_m}{P_m} = 0.29$)

Low flows

For the very same reason, DD cannot be an *explicans* for low flows : groundwater may still contribute to discharge during low flow periods, but in a proportion depending on the size of the groundwater catchment, about which DD tells nothing.

High flows

In contrary we will show that, on these seven catchments, the correlation between drainage density and flood discharges is observed. Flood discharges of the Cailly river ($Q_{ix10} = 0.17 \text{ mm/h}$) are far less than those of the reference Epte catchment ($Q_{ix10} = 0.48 \text{ mm/h}$), though its mean annual (specific) flow is 1.8 times higher due to groundwater contribution. What

is more, flood discharges of the Cailly do not exceed those of the Commerce river ($Q_{ix10} = 0.20$ mm/h), although their mean annual flows are completely different.

A convenient way to look at these behaviors is to use a Quantile-Quantile (QQ) plot of specific hourly discharges (Figure F.14). We take the Epte (impervious) catchment as the reference distribution in x-axis, and then we plot the corresponding discharge quantiles of the six chalky catchments in the y-axis. Figure F.14 shows that :

- in the left part (low flows), all curves plot higher than the 1/1 line but without converging. This indicates that low flows of all chalky catchment are sustained by groundwater (compared to the Epte river), though in very different quantitative ways. DD brings no information on these quantities.
- in the right part (high flows), all curves converge to a point located below the 1/1 line (between 1/2 and 1/4). This is a clear indication of a dampering effect on flood peaks for all chalky catchments.

This dampering effect, and the fact that DD is a statistical *explicans* for it, is quite easy to understand . During storm events, an important part of net rainfall will infiltrate on those catchments where the lithology is very permeable (and thus, the DD very low). Large amounts of water will bypass the channel network through slower subsurface or groundwater flow paths. Whether this water will ultimately reach the surface stream (e.g., in the form of baseflow during summer) or not does not affect the short-term reaction of the catchment : we will always observe a dampering effect on floods. Low DD itself is not the main cause of this attenuation (which is mainly the consequence of a decrease in surface runoff volume because of infiltration), though it can reinforce the dampering effect through delayed flood routing in unchannelized overland flow.

F.6 Limits and perspectives of the use of drainage density in hydrology

F.6.1 Synthesis

In this paper, our aim was to investigate a thirty year old unsettled debate on the hydrological relevance of the Drainage Density metric. We considered that only data could settle this controversy, and we confronted DD and streamflow characteristics over a large set of 930 catchments. Our results confirm the propositions of Dingman, in that they showed DD to be a significant explicans of the 10-year return period flood estimates, while this was not true when we examined low- and mean-flow statistics.

On a subset of seven catchments located in a climatologically homogeneous but geologically heterogeneous area, we could analyze more in detail the variability of catchment responses for different flow quantiles. We verified that convergence of behaviors for the low-DD catchments only occurs for the higher flow quantiles, while at low an mean flows, a considerable variability of behavior remains.

F.6.2 Did we settle the controversy ?

Following *Popper* (1959), we are reluctant to pretend having positively ‘demonstrated’ anything about the value of DD as a relevant hydrological metric (though we think that its relevance as regards to mean- and low-flow prediction is clearly falsified). At least, we believe that the size of the dataset, its geological and climatological variability, and the care that was taken in computing DD values allows us to write that field evidence clearly favors *Dingman’s* (1978) views. What is more, we would like to emphasize this fact : though many problems exist as regards to stream cartography (the ‘blue-line’ issue), we believe that this information is *statistically* highly relevant. Map making is an iterative process but still incorporates a priceless amount of field evidence.

Last, we must acknowledge that the fact that a good correlation exists with the 10-year return period flood flow does not imply that drainage density is the ‘physical’ cause of rainfall-runoff transformation. As mentioned by *Dingman* (1978), it may still be a substitute for other ‘more fundamental’ (lithological or morphological) attributes. It is most likely that drainage density sums up an equilibrium state between other quantities controlling energy dissipation through erosion processes (*McNamara*, 2006), such as mean annual water yield, slope, permeability, etc. Such relations are to be investigated in a similar statistical approach.

Acknowledgments

The authors wish to thank Dr Roger Moussa of INRA (especially for his introduction to the concept of variable extraction threshold), Prof. S. Lawrence *Dingman*, John B. Conroy and Robert A. Kimmel of USGS, and Prof. Desmond E. Walling of the University of Exeter for scientific exchanges which fueled the preparation of this paper.

References

Benson, M. A. (1960), Areal flood-frequency analysis in a humid region, *Bulletin of the International Association for Scientific Hydrology*, 5, 5–15.

Chorley, R. J., and P. F. Dale (1972), Cartographic problems in stream channel delineation, *Cartography*, 7, 150–162.

Dingman, S. L. (1978), Drainage density and streamflow : a closer look, *Water Resour. Res.*, 14(6), 1183–1187.

Dunne, T., and R. D. Black (1970), Partial area contributions to storm runoff in a small New England watershed, *Water Resour. Res.*, 6, 1296–1311.

Garbrecht, J., and L. W. Martz (1997), The assignment of drainage direction over flat surfaces in raster digital elevation models, *J. Hydrol.*, 193(1–4), 204–213.

Gardiner, V., and K. J. Gregory (1982), Drainage density in rainfall-runoff modelling, in *Rainfall-runoff relationships*, pp. 449–476, edited by V.P. Singh, Water Resources Publications, Littleton.

Gregory, K. J., and V. Gardiner (1975), Drainage density and climate, *Zeitschrift für Geomorphologie*, 19, 287–298.

Gregory, K. J., and V. Gardiner (1979), Comment on ‘Drainage density and streamflow : a closer look’ by S. Dingman, *Water Resour. Res.*, 15(6), 1662–1664.

Gregory, K. J., and D. E. Walling (1968), The variation of drainage density within a catchment, *Bull. Int. Assoc. Sci. Hydrol.*, 13(2), 61–68.

Gregory, K. J., and D. E. Walling (1973), *Drainage Basin Form and Process. A Geomorphological Approach*, Edward Arnold, London.

Hausdorff, F. (1919), Dimension und äusseres Mass, *Math. Ann.*, 79, 157–179.

Hewlett, J. D., and A. R. Hibbert (1963), Moisture and energy conditions within a sloping soil mass during drainage, *J. Geophys. Res.*, 68(4), 1081–1087.

Hewlett, J. D., and A. R. Hibbert (1967), Factors affecting the response of small watersheds to precipitation in humid areas, in *International symposium on forest hydrology*, pp. 275–290, edited by W.E. Sopper and H.W. Lull, Pergamon, Oxford, UK.

Hjelmfelt, A. T. (1988), Fractals and the river length-catchment area ratio, *Water Resources*

Bulletin 24 (2), 455–459.

Horton, R. E. (1932), Drainage basin characteristics, *Am. Geophys. Union Trans.*, 13, 350–361.

Horton, R. E. (1945), Erosional development of streams and their drainage basins : hydro-physical approach to quantitative morphology, *Bulletin of the Geological Society of America*, 56, 275–370.

Langbein, W. B. (1944), Topographic characteristics of drainage basins, *USGS Water Supply paper 968C*, pp. 125–157.

Martz, L. W., and J. Garbrecht (1998), The treatment of flat areas and depressions in automated drainage analysis of raster digital elevation models, *Hydrol. Processes*, 12, 843–855.

Mandelbrot, B. (1982), *The Fractal Geometry of Nature*, W. H. Freeman, San Francisco.

McNamara, J. P., A. D. Ziegler, S. H. Wood, and J. B. Vogler (2006), Channel head locations with respect to geomorphologic thresholds derived from a digital elevation model : A case study in northern Thailand, *Forest Ecology and Management*, 224(1–2), 147–156.

Montgomery, D. R., and W. E. Dietrich (1989), Source areas, drainage density, and channel initiation, *Water Resour. Res.*, 25(8), 1907–1918.

Montgomery, D. R., and E. Foufoula-Georgiou (1993), Channel Network Source Representation Using Digital Elevation Models, *Water Resour. Res.*, 29(12), 3925–3934.

Moussa, R. and C. Bocquillon (1996), Fractal analyses of tree-like channel networks from digital elevation model data, *J. Hydrol.* 187, 157–172.

Ol'dekop, E. M. (1911), On evaporation from the surface of river basins, *Trans. Meteorol. Observ. University of Tartu*, 4, pp. 200 and 20 figures.

Pike, J. G. (1964), The estimation of annual runoff from metrological data in tropical climate, *J. Hydrol.*, 2, 116–123.

Popper, K. R. (1959), *The Logic of Scientific Discovery*, Hutchinson, London, English translation, with new preface, footnotes and appendices of *Logik der Forschung*, 1934.

Rodriguez-Iturbe, I., and A. Rinaldo (1997), *Fractal River Networks : Chance and Self-Organization*, Cambridge University Press, New York, 1997.

Rucklidge, W. (1996), *Efficient Visual Recognition Using the Hausdorff Distance*, Springer.

Schneider, W. J. (1961), A note on the accuracy of drainage densities computed from topographic maps, *Journal of Geophysical Research*, 66(10), 3617–3618.

Schreiber, P. (1904), Über die Beziehungen zwischen dem Niederschlag und der Wasserführung der Flüsse in Mitteleuropa, *Z. Meteorol.*, 21, 441–452.

Turc, L. (1954), Le bilan d'eau des sols : relation entre les précipitations, l'évapotranspiration et l'écoulement, *Annales agronomiques, Série A*, 491–595.

Wharton, G. (1994), Progress in the use of drainage network indices for rainfall-runoff modelling and runoff prediction, *Progress in Physical Geography*, 18(4), 539–557.

A	P_m	PE_m	P_{10}	DD	AI	RII	R^2
0.907							0.772
0.903						-0.172	0.773
0.947				0.168			0.791
0.942				0.199		-0.413	0.798
0.925		-3.040					0.820
0.925		-3.046				0.014	0.820
0.971			0.973				0.842
0.978			0.916	0.048			0.843
0.969		-3.171		0.186			0.844
0.966		-3.071		0.202		-0.230	0.846
0.979		-2.544	0.870				0.875
0.992		-2.661	0.767	0.082			0.879
0.988	2.141						0.941
0.984					-1.784		0.941
0.986	2.180		-0.042				0.941
0.986					-1.788	0.046	0.941
0.988			0.109		-1.707		0.942
0.997	2.089			0.050			0.943
0.996	2.189		-0.119	0.060			0.943
0.988	2.003	-0.919					0.945
0.988	1.997	-0.921	0.006				0.945
1.001				0.080	-1.732		0.945
1.001			0.001	0.079	-1.731		0.945
1.000				0.085	-1.722	-0.065	0.946
1.001	1.914	-1.059		0.066			0.948
0.999	1.986	-1.033	-0.080	0.072			0.948

TAB. F.1: Regression results for mean annual streamflow values. The figures in each column are the exponents of the corresponding explanatory variable, and the value of R^2 is the last figure of each row. AI stands for aridity index ($AI = PE_m/P_m$) and RII stands for AI for rainfall irregularity index ($RII = P_{10}/P_m$).

A	P_m	PE_m	P_{10}	DD	AI	RII	R^2
1.009							0.454
0.993				-0.069			0.455
1.052			0.659				0.469
0.981						-1.007	0.474
0.983				0.008		-1.017	0.474
1.024			0.882	-0.185			0.478
1.045		-6.211					0.550
1.037		-6.186		-0.035			0.550
1.071		-5.973	0.418				0.556
1.025		-5.894				-0.648	0.558
1.028		-5.895		0.014		-0.664	0.558
1.053		-5.815	0.557	-0.110			0.559
1.102	2.463						0.560
1.058	2.693			-0.221			0.575
1.075	3.265		-0.861				0.575
1.049	3.240		-0.648	-0.167			0.582
1.110					-2.319		0.590
1.102	1.820	-4.284					0.598
1.086					-2.248	-0.734	0.600
1.070				-0.194	-2.445		0.601
1.082			-0.798		-2.878		0.604
1.070	2.042	-3.933		-0.163			0.606
1.061				-0.149	-2.364	-0.539	0.606
1.082	2.478	-3.960	-0.655				0.607
1.061			-0.617	-0.133	-2.838		0.609
1.063	2.498	-3.768	-0.508	-0.123			0.611

TAB. F.2: Regression results for 5-year return period minimal monthly streamflow (Q_{mna5}). The figures in each column are the exponents of the corresponding explanatory variable, and the value of R^2 is the last figure of each row. AI stands for aridity index ($AI = PE_m/P_m$) and RII stands for AI for rainfall irregularity index ($RII = P_{10}/P_m$).

<i>A</i>	<i>P_m</i>	<i>PE_m</i>	<i>P₁₀</i>	<i>DD</i>	<i>AI</i>	<i>RII</i>	<i>R²</i>
0.712							0.491
0.717		-0.733					0.494
0.748						1.259	0.559
0.758		-1.390				1.343	0.570
0.784					-1.657		0.642
0.832				0.510			0.672
0.839		-1.098		0.516			0.679
0.798	2.285						0.690
0.839				0.457	0.706		0.692
0.850		-1.445		0.458	0.793		0.703
0.798	2.587	2.007					0.708
0.832					-1.799	1.477	0.735
0.834			1.859				0.754
0.833		0.335	1.873				0.754
0.840			1.583		-0.547		0.764
0.875				0.440	-1.371		0.772
0.841	0.979		1.403				0.774
0.839	1.247	1.347	1.333				0.783
0.877	1.865			0.404			0.797
0.873	2.060	1.175		0.386			0.803
0.890				0.356	-1.522	1.014	0.811
0.883			1.479	0.315			0.812
0.884		-0.118	1.472	0.316			0.812
0.891			1.138	0.327	-0.646		0.827
0.891	1.027		0.994	0.320			0.835
0.888	1.197	0.863	0.962	0.310			0.838

TAB. F.3: Regression results for 10-year return period peak flow (Q_{ix10}). The figures in each column are the exponents of the corresponding explanatory variable, and the value of R^2 is the last figure of each row. *AI* stands for aridity index ($AI = PE_m/P_m$) and *RII* stands for *AI* for rainfall irregularity index ($RII = P_{10}/P_m$).

Catchment	Area km ²	<i>P_m</i> mm/yr	<i>PE_m</i> mm/yr	<i>P₁₀</i> mm/day	<i>Q_m</i> mm/yr	<i>y</i> ∅	<i>Q_{ix10}</i> mm/hr	<i>Q_{mna5}</i> mm/month
Andelle	371	857	630	54.9	339	0.40	0.16	16.3
Cailly	196	937	625	60.0	428	0.46	0.17	20.1
Commerce	116	1048	618	61.5	65	0.06	0.20	2.7
Durdent	349	985	606	61.4	346	0.35	0.13	20.3
Ganzéville	89	1031	609	60.8	280	0.27	0.16	11.5
Yeres	265	926	595	57.4	323	0.35	0.13	12.9
Epte	245	831	629	53.7	241	0.29	0.48	2.0

TAB. F.4: Characteristics of the seven catchments used in Section 5. *P_m* : mean annual rainfall; *PE_m* : mean annual potential evapotranspiration; *P₁₀* : 10-year daily rainfall; *Q_m* : mean annual discharge; *y* = Q_m/Q_m : mean annual yield; *Q_{ix10}* : 10-year flood peak; *Q_{mna5}* : 5-year minimal monthly flow.

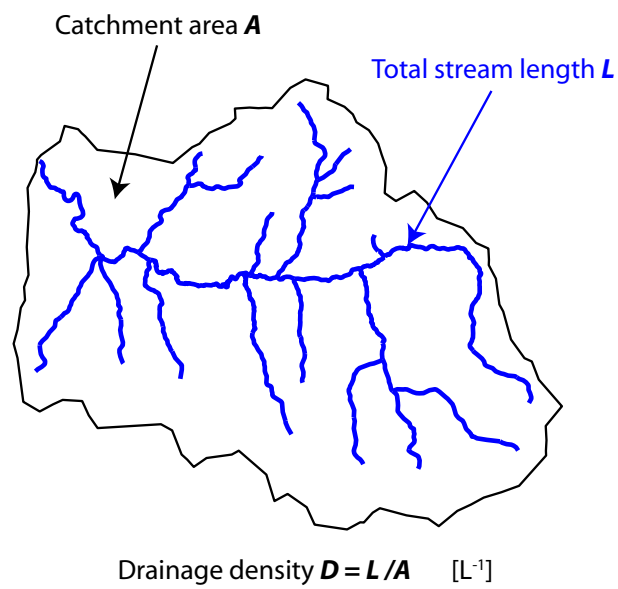


FIG. F.1: Classical definition of drainage density.

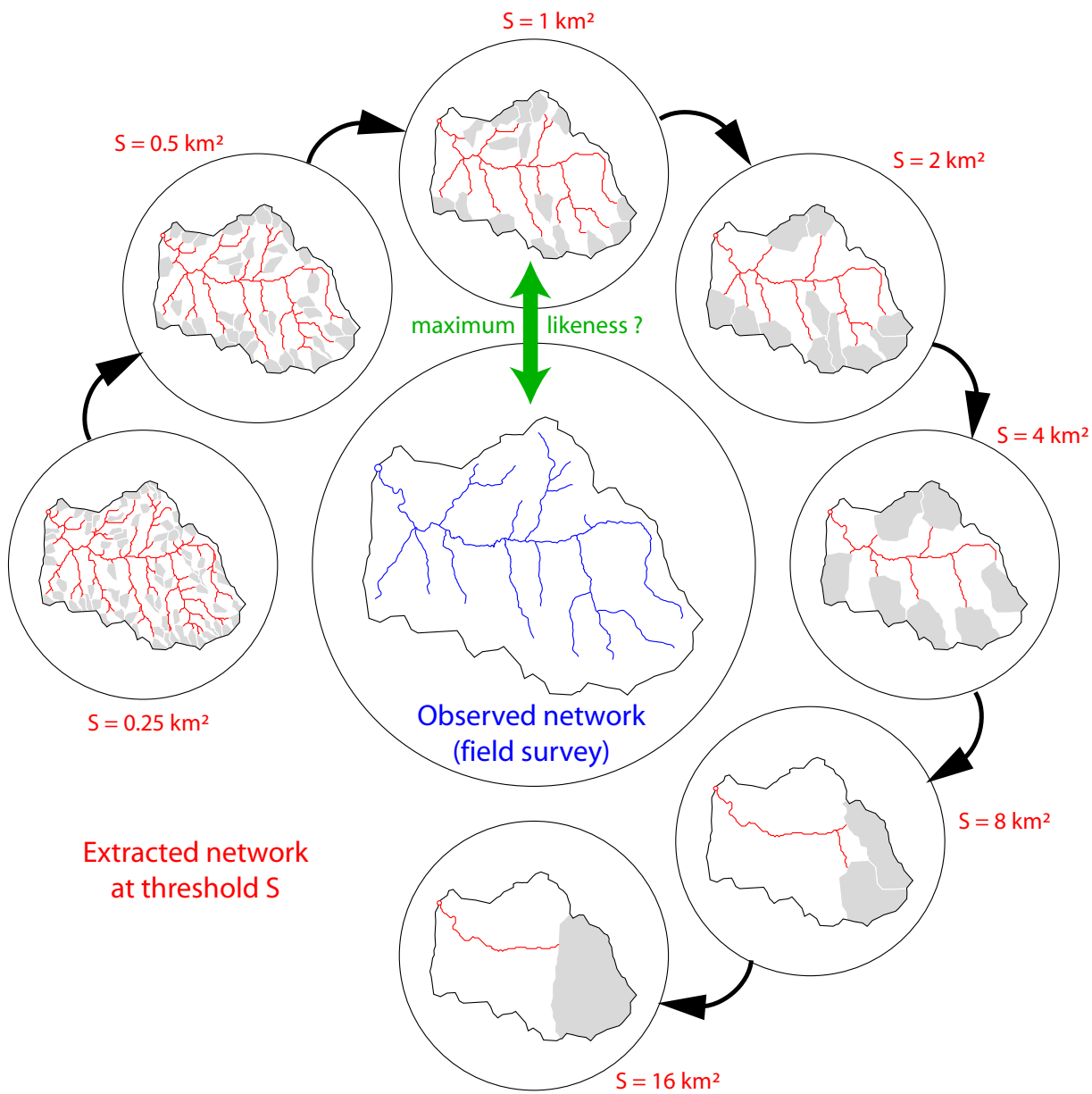


FIG. F.2: Illustration of the variable extraction threshold procedure in a homogeneous subcatchment. We vary the threshold of upstream area above which a point is considered to belong to the channel network, until reaching the maximum likeness with the actual (according to some likeness criterion that we will define later on).

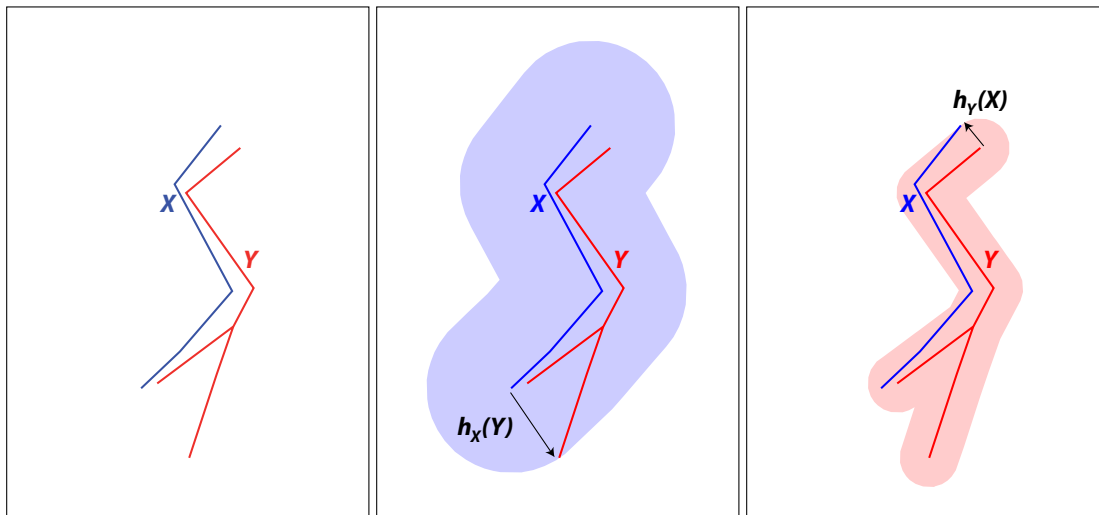


FIG. E3: Illustration of the computation of the Hausdorff distance between two sets of point X and Y . First, a distance map to the X set is computed and sampled at each point of Y : the maximum gives the semi-distance $h_X(Y)$ (width of the smallest blue buffer around X that entirely contains Y). Then the roles of X and Y are switched in order to compute the second semi-distance $h_Y(X)$ (width of magenta buffer). The Hausdorff distance $H(X, Y)$ is the maximum of $h_X(Y)$ and $h_Y(X)$ (in the present case, $H(X, Y) = h_X(Y)$)

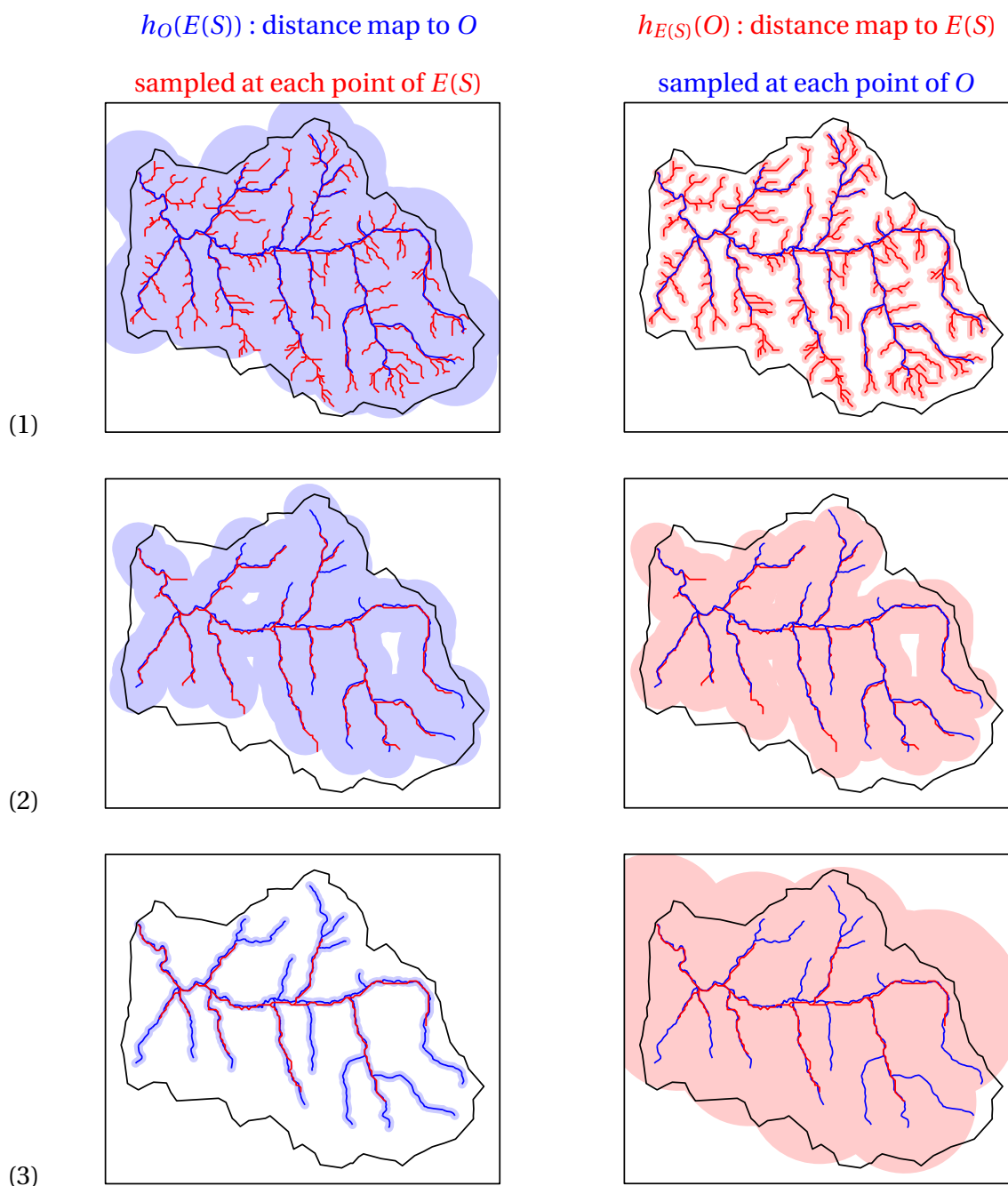


FIG. F.4: Evolution of Hausdorff semi-distances $h_O(E(S))$ (blue buffer) and $h_{E(S)}(O)$ (magenta buffer) when increasing the extraction threshold S . The observed network O is stroked in blue and the extracted network (variable) $E(S)$ in red. (1) Top : extraction with $S = 0.1 \text{ km}^2$. The extracted network is much denser than the observed network. (3) Bottom : extraction with $S = 3 \text{ km}^2$. The extracted network shrinks to a subset of the observed network. (2) Middle : the two networks have a similar development, the Hausdorff distance is minimized ($S \approx 0.8 \text{ km}^2$).

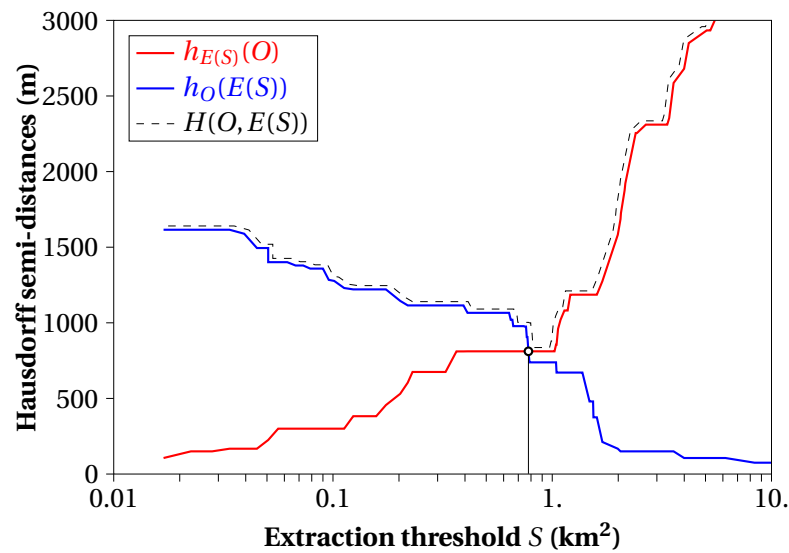
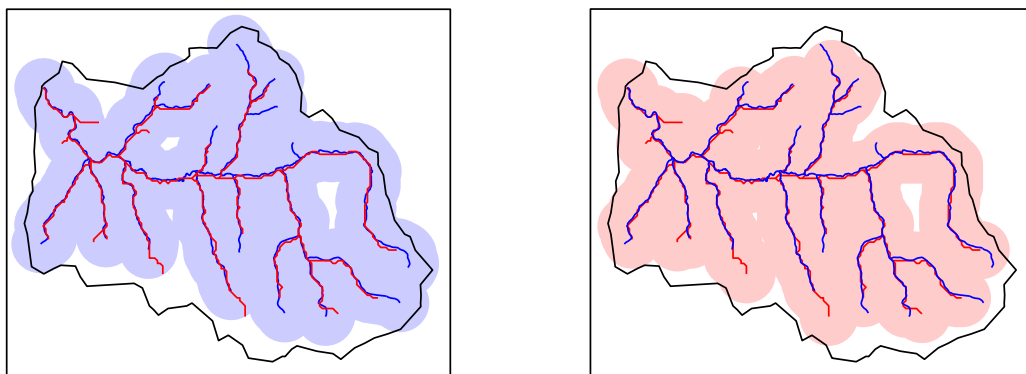


FIG. F.5: Semi-distances $h_O(E(S))$, $h_{E(S)}(O)$ and Hausdorff distance $H(O, E(S))$ as functions of extraction threshold S in the case of Figure 1. The Hausdorff distance is minimized for $S \approx 0.8$ km^2 which corresponds to $H(O, E(S)) = h_{E(S)}(O) = h_O(E(S))$.

Initial polylines for observed network (blue):



Degraded polylines for observed network (blue):

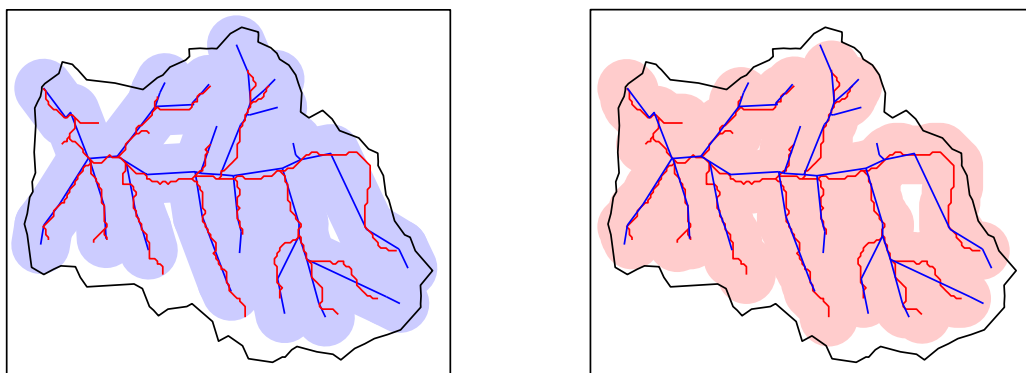


FIG. F.6: Impact of degraded precision of the polylines describing the ‘actual’ network on the calculated drainage density. We see that such a degradation does not affect the algorithm based on Hausdorff distance, which is thus much less scale-dependent than the usual approach based on the total stream length. The Hausdorff distance is sensitive to the position of the extremities of the polylines (i.e., the sources of the streams, which is what we are interested in) but not to the sinuosity of these lines (which does unintently affect the value given by the classical definition of DD).

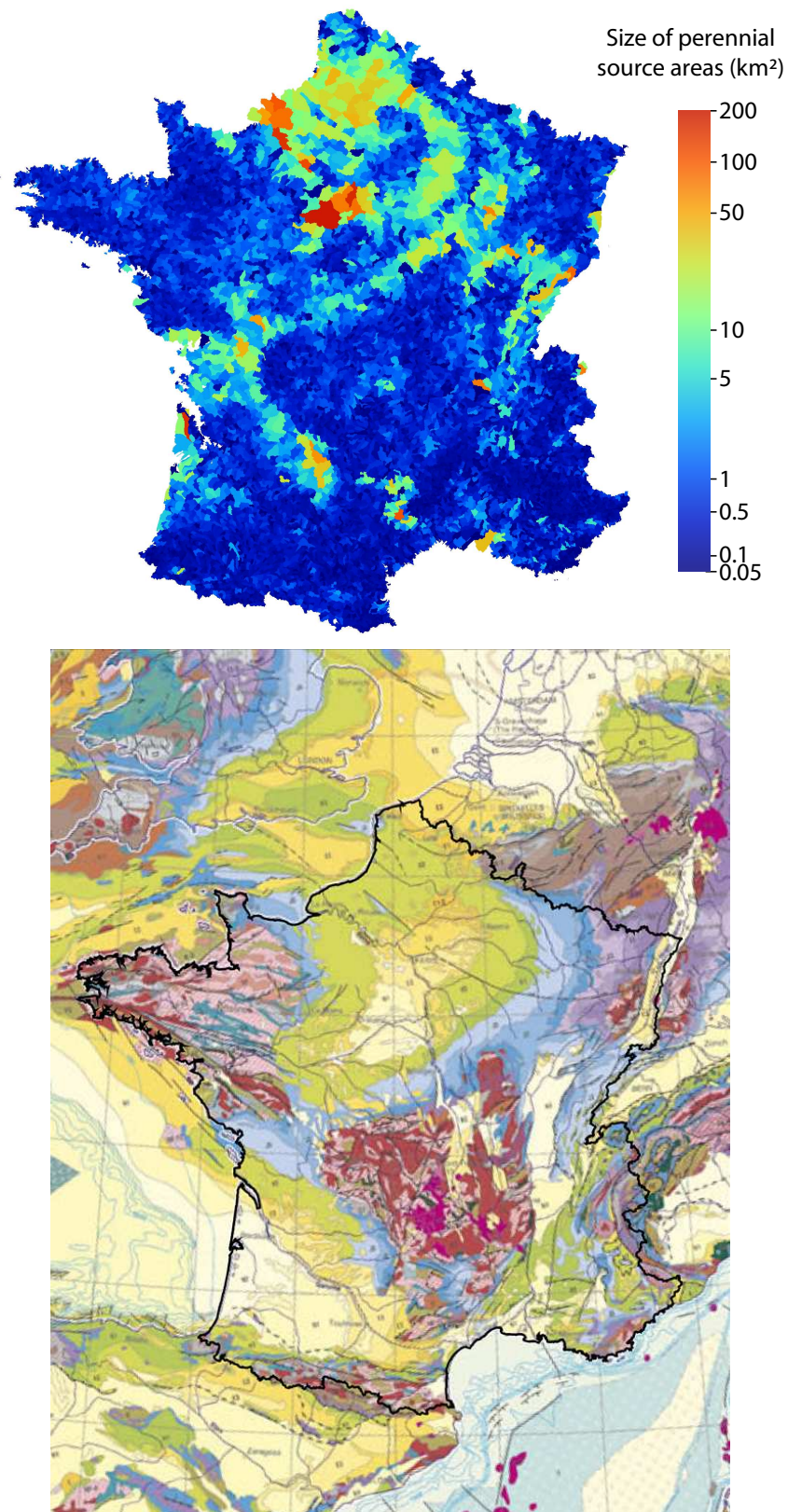


FIG. E.7: Map of perennial source area values for a topographical mesh (top) and geological map of France (bottom). We see a clear link between zones of low drainage density (i.e. high values of source areas) and permeable lithologies such as chalks (light green), middle Jurassic limestones (Dogger, deep blue) or sands (beige). Conversely, drainage density is high in crystalline bedrock regions (redish colors in Western and Central regions).

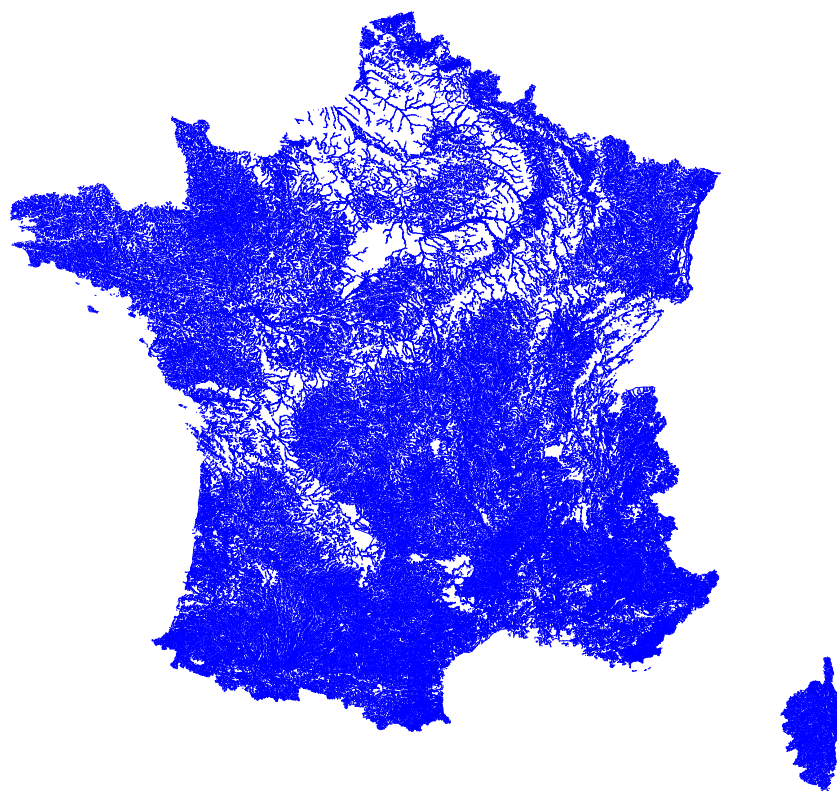


FIG. F.8: Map of the French hydrographic network at the highest level of detail given by the national database RNDE. The extension of “empty” zones matches very well the extension of specific lithological units (see Figure F.7).

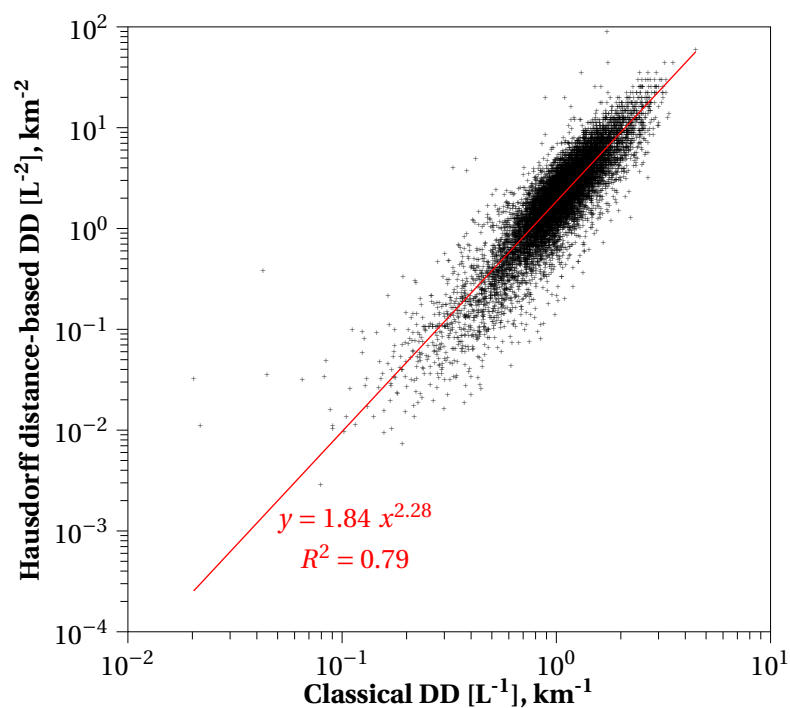


FIG. F.9: Correlation between classical drainage density $[L^{-1}]$ (x-axis) and Hausdorff distance based DD $[L^{-2}]$ (y-axis) for the 10700 elements of Figure 7.

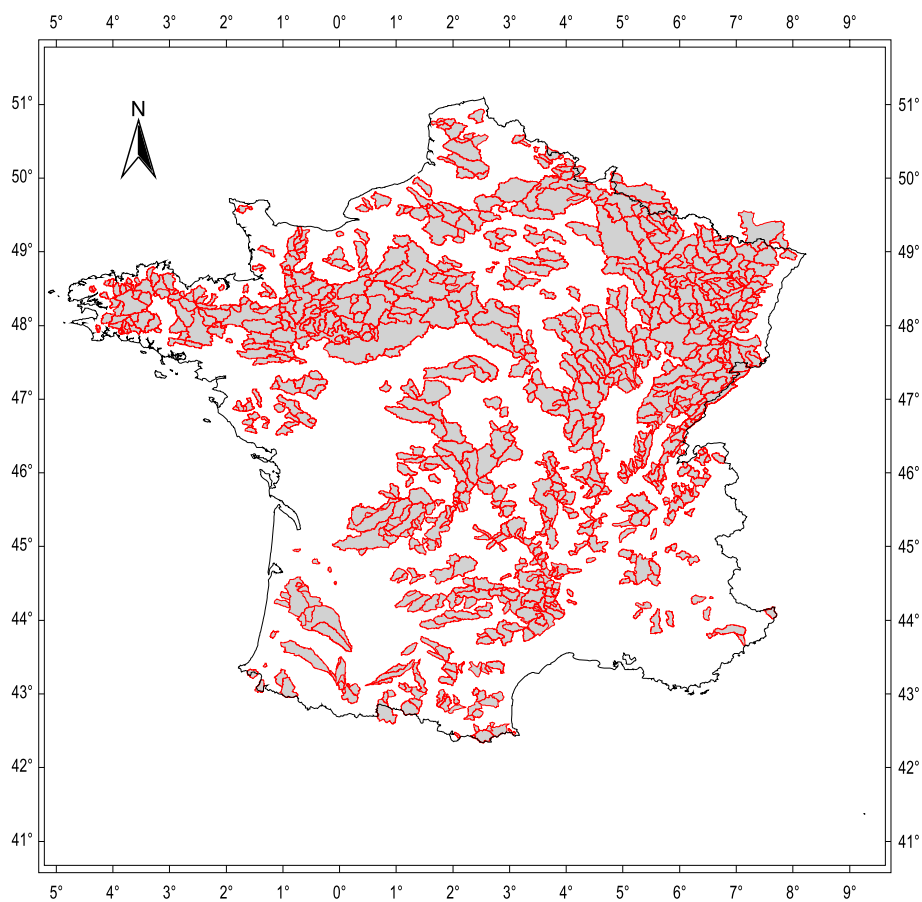


FIG. F.10: Map of the 930 French catchments used for characteristic streamflow predictions. For each of these catchments we have the following descriptors and data : A : catchment area ; P_m : mean annual rainfall ; PE_m : mean annual potential evapotranspiration ; P_{10} : 10-year daily rainfall ; DD : Drainage density (spatial geometrical mean) ; Q_m : mean annual discharge ; Q_{ix10} : 10-year flood peak ; Q_{mna5} : 5-year minimal monthly flow.

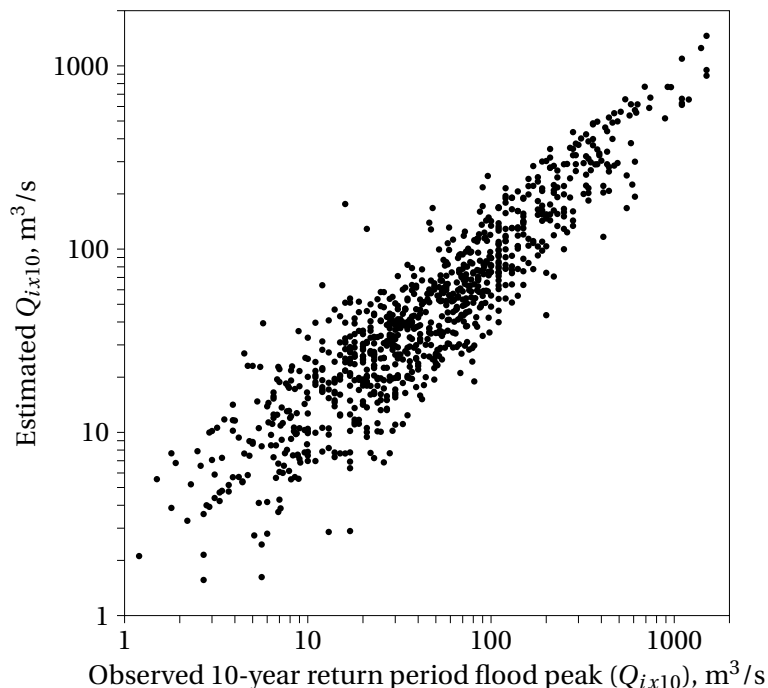


FIG. F.11: Scatter plot of observed and estimated 10-year return period flood peak using the regression of section 4.3.

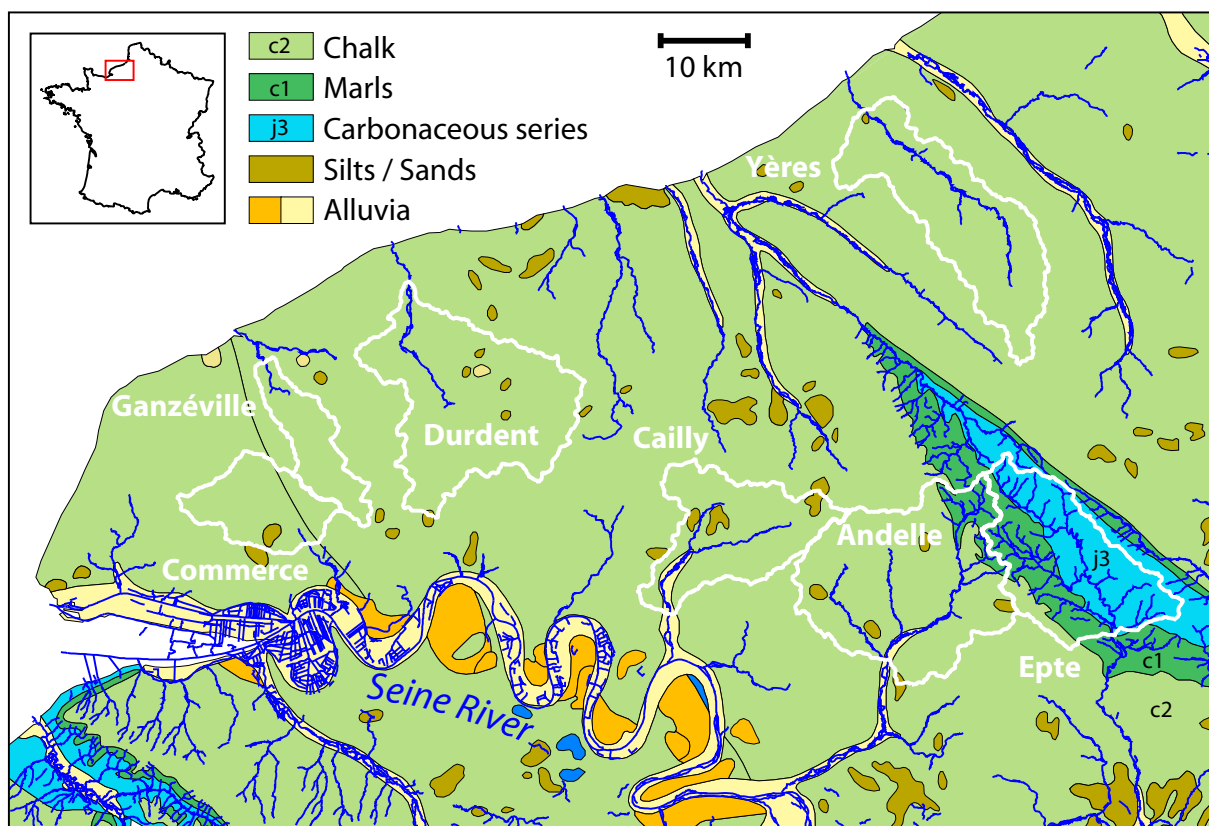


FIG. F.12: Geological setting and set of seven catchments used in Section 5. Six of those catchments are located on the Chalk (c2, green) while the remaining one —Epte river, which will be used as reference— is located on impervious substratum (marls and carbonaceous series). The structure on which the latter (eastmost) catchment is located is a small eroded anticline (*Bray anticline*), at the center of which older layers have been exposed.

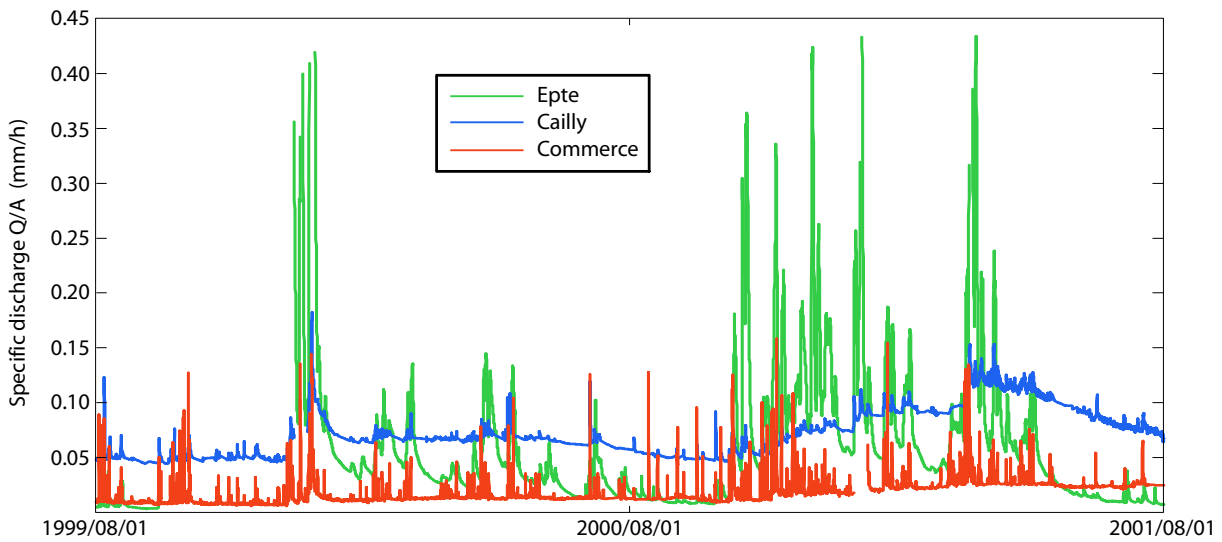


FIG. F.13: 2-year hourly hydrograph for Epte, Cailly and Commerce rivers. Y-axis is the specific discharge Q/A (mm/hr).

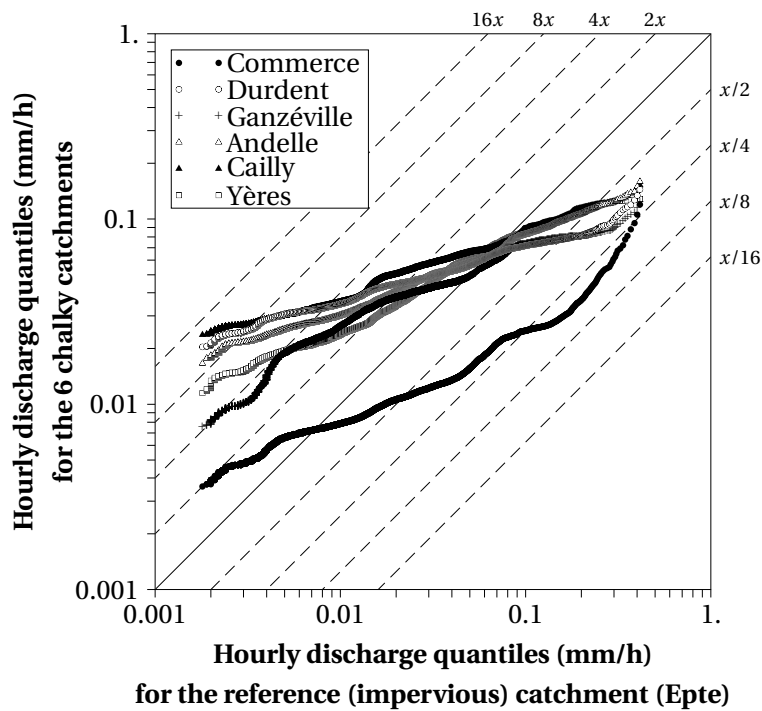


FIG. F.14: Quantile-quantile plot of hourly specific discharges, with the Epte river discharge taken as the reference in the x-axis.

Résumé

Les hydrologues dits « de surface » ont longtemps développé des représentations du fonctionnement des bassins versants basées sur une équation de bilan à trois termes : la pluie, l'évapo-transpiration, et le débit. L'équation de continuité ainsi définie stipule donc que sur le long terme, le débit écoulé à l'exutoire du bassin est égal à la différence entre précipitations et évapo-transpiration (un pas de temps suffisamment long, typiquement quelques années, permettant de négliger les variations de stock). L'une des hypothèses de ces modèles (restée souvent implicite) est donc que les infiltrations profondes, ne réapparaissant pas à l'exutoire des bassins topographiques, sont une quantité négligeable.

Ce travail a pour but de questionner la pertinence de cette hypothèse non seulement du point de vue du réalisme des modèles hydrologiques, mais plus encore du point de vue de leur efficacité. De fait, les modèles disposent souvent d'artifices numériques leur permettant d'adapter le bilan en eau, et les conséquences néfastes de cette hypothèse n'apparaissent que rarement au grand jour : en effet, tout déficit d'écoulement peut être imputé à l'autre sortie du système, l'évapotranspiration, tout aussi difficile à estimer.

Nous avons donc axé nos recherches dans deux directions complémentaires :

- d'un côté, nous avons cherché à établir des méthodes et des outils pour évaluer le réalisme des flux souterrains — ou supposés tels — simulés par les modèles,
- de l'autre, nous avons testé une large gamme de formulations mathématiques empiriques pour prendre en compte les échanges surface-souterrain dans des modèles pluie-débit globaux journaliers et horaires, dans le but principal d'améliorer la capacité de ces modèles à reproduire le seul flux mesurable : le débit de surface.

Ces travaux, basés sur un large échantillon de plus d'un millier de bassins versants français et australiens et menés au travers d'une approche comparative, nous ont permis d'aboutir plus particulièrement à une amélioration de la simulation des étiages. Ce processus a également été l'occasion de soulever des questions relatives aux critères d'évaluation des modèles et au choix d'un niveau de complexité approprié à chaque résolution temporelle.

Mots-clés

Modélisation pluie-débit; Bassin versant; Interactions eaux de surface – eaux souterraines; Hydrologie karstique; Géomorphologie; Critères d'évaluation; Sélection de modèles.