



HAL
open science

Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative

Charles Perrin

► **To cite this version:**

Charles Perrin. Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative. Hydrologie. Institut National Polytechnique de Grenoble - INPG, 2000. Français. NNT : 2000INPG0105 . tel-00006216

HAL Id: tel-00006216

<https://theses.hal.science/tel-00006216v1>

Submitted on 10 Jun 2004

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

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



INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque

□□□□□□□□□□

THESE

pour obtenir le grade de

DOCTEUR de l'INPG

Spécialité: Mécanique des Milieux Géophysiques et Environnement

préparée dans l'Unité de Recherche Qualité et Fonctionnement Hydrologique
des Systèmes Aquatiques, *Cemagref*, Antony

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

présentée et soutenue publiquement par

Charles Perrin

le 20 octobre 2000

Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative

Directeur de thèse :
M. Jean-Michel Grésillon

JURY

M. Bruno Ambroise
M. Claude Thirriot
M. Eric Servat
M. Jean-Michel Grésillon
M. Claude Michel
M. Rémy Garçon
M. Ian Littlewood
M. Thierry Leviandier

Président
Rapporteur
Rapporteur
Directeur de thèse
Co-encadrant
Examineur
Membre invité
Membre invité

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque

□□□□□□□□□□

THESE

pour obtenir le grade de

DOCTEUR de l'INPG

Spécialité: Mécanique des Milieux Géophysiques et Environnement

préparée dans l'Unité de Recherche Qualité et Fonctionnement Hydrologique
des Systèmes Aquatiques, *Cemagref*, Antony

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

présentée et soutenue publiquement par

Charles Perrin

le 20 octobre 2000

**Vers une amélioration d'un modèle global pluie-débit
au travers d'une approche comparative**

Directeur de thèse :
M. Jean-Michel Grésillon

JURY

M. Bruno Ambroise	Président
M. Claude Thirriot	Rapporteur
M. Eric Servat	Rapporteur
M. Jean-Michel Grésillon	Directeur de thèse
M. Claude Michel	Co-encadrant
M. Rémy Garçon	Examineur
M. Ian Littlewood	Membre invité
M. Thierry Leviandier	Membre invité

Remerciements

Je remercie chaleureusement Jean-Michel Grésillon, Directeur de l'Ecole Nationale Supérieure d'Hydraulique et de Mécanique de Grenoble (ENSHMG), pour avoir accepté la direction de cette thèse, pour le suivi de mon travail, ses conseils au cours de ces trois années et ses suggestions pour améliorer la qualité de mes recherches.

Toute ma reconnaissance va à Claude Michel, hydrologue au Cemagref d'Antony, qui a encadré ce travail de recherche. Il a su par ses idées, ses conseils, ses explications et ses critiques, insuffler constamment de l'élan à cette recherche. J'ai bénéficié de son expérience en hydrologie et de la clairvoyance de ses opinions. La confiance qu'il m'a accordée et son soutien constant ont été précieux dans la conduite de cette thèse.

Je tiens à remercier Thierry Leviandier, Directeur de la Recherche de l'Ecole Nationale du Génie de l'Eau et de l'Environnement de Strasbourg (ENGEES), pour l'intérêt qu'il a montré pour cette thèse, pour ses remarques sur mon travail, et pour sa participation au comité de suivi de la thèse et à ce jury.

Ian Littlewood, ingénieur-chercheur du Centre for Ecology and Hydrology (CEH) de Wallingford en Grande-Bretagne, a accepté de suivre mon travail et d'être membre du jury. C'est à ses côtés, au cours de mon stage de DEA, que j'ai effectué mes premiers pas en modélisation. J'ai pu bénéficier de ses idées et de son expérience. Je lui adresse ma sincère gratitude.

Vazken Andréassian et Etienne Leblois, ingénieurs-chercheurs au Cemagref d'Antony et de Lyon, ont également, par leurs commentaires et leurs encouragements, contribué à améliorer la qualité de ces travaux. Le temps qu'ils ont accepté de me consacrer, malgré leurs agendas chargés, notamment dans leur participation au comité de suivi, a été précieux pour moi. Je les en remercie chaleureusement.

Bruno Ambroise, Directeur de Recherche du Centre d'Etudes et de Recherches Eco-Géographiques (CEREG) à Strasbourg m'a fait l'honneur d'accepter de présider ce jury. C'est en écoutant ses cours sur les bancs de l'école que j'ai commencé à m'intéresser à l'hydrologie et à la modélisation. Je l'en remercie vivement.

J'adresse toute ma gratitude à Claude Thirriot, Professeur de l'Ecole Nationale Supérieure d'Electrotechnique, d'Electronique, d'Informatique, d'Hydraulique de Toulouse (ENSEEIH) pour avoir accepté d'être rapporteur de ces travaux.

J'adresse également ma profonde reconnaissance à Eric Servat, Directeur de Recherche de l'Institut de Recherche pour le Développement (IRD) à Montpellier, pour sa participation au jury en temps que rapporteur et pour le temps qu'il a consacré à la critique de ce manuscrit. Je lui adresse aussi mes remerciements, ainsi qu'à Brou Kouamé de l'antenne d'hydrologie de

l'IRD à Abidjan, pour avoir mis à ma disposition les données hydro-climatiques sur des bassins de Côte-d'Ivoire.

Rémy Garçon, responsable de l'équipe Hydrologie et Gestion de l'Eau à la Direction Technique Générale d'Electricité de France à Grenoble, a accepté d'être examinateur de ces travaux. Je lui exprime ma sincère reconnaissance.

Je remercie Daniel Loudière, Directeur de l'ENGEES, pour m'avoir permis de réaliser ma thèse en formation continue par la recherche au Cemagref. Je tiens à remercier également Gérard Sachon (actuellement chef du Département Gestion des Milieux Aquatiques du Cemagref) et Jean-Luc Pujol, successivement chefs de l'Unité de Recherche Qualité et Fonctionnement Hydrologique des Systèmes Aquatiques (QHAN) du Cemagref d'Antony, ainsi que Jacques Joly, chef du groupement du Cemagref d'Antony, pour leur accueil, leur soutien et les bonnes conditions dans lesquelles j'ai pu réaliser mes travaux.

Les données des bassins versants australiens ont été fournies par Francis Chiew du Département de Génie Civil et Environnemental de l'Université de Melbourne en Australie. Les données sur les bassins versants américains provenant de la base de données du Water Data Center de Beltsville, service du Ministère de l'Agriculture des Etats-Unis, ont été fournies par Jane Thurman. Enfin, les données des bassins versants brésiliens ont été fournies par Nilo Oliveira Nascimento de l'Université de Minas Gerais à Belo Horizonte au Brésil. Je voudrais à tous les trois adresser ma reconnaissance pour la gracieuse mise à disposition de ces données.

Je voudrais remercier toutes les personnes de l'Unité de Recherche QHAN qui ont contribué à ce travail, par leur aide et leurs encouragements. Je me permets de distinguer Michel Poirson, hydraulicien, pour ses précieux conseils en programmation et son indulgence vis-à-vis de l'absence répétée de commentaires dans mes programmes; Jean-Louis Rosique pour sa contribution à la mise en place de la base de données et son aide efficace pour la gestion au quotidien des problèmes informatiques; Mamoutou Tangara pour ses conseils en statistiques; Michel Ferry pour avoir mis à ma disposition des programmes de conversion de format de données; Sylvie Tonachella, Sophie Morin et Ana Badji (actuellement à l'INRA) du secrétariat pour leur aide dans les démarches administratives. Je n'oublierai pas non plus Michel Normand (qui profite maintenant de sa retraite), Cécile Jouanicot-Loumagne et Anne Weisse qui m'ont encouragé durant ces années de recherche. Merci également à tous les thésards et stagiaires de l'unité et à ceux de l'unité Drainage et Etanchéité, que j'ai côtoyés au cours de ces trois années, en particulier ceux de l'équipe hydrologie Manuelle, Marine, Valérie, Julien, Frédéric, Ahmed, Nicolas, Rhéda, Xavier et Laurent, pour leur bonne humeur et pour m'avoir tiré de temps en temps de mon écran d'ordinateur...

J'adresse également toute ma reconnaissance à Geneviève Michel, Agnès Dao, Danièle Cussatlegras et Hermine Bartolome du service de documentation du Cemagref d'Antony, pour leur disponibilité, leur aide et leur efficacité dans la recherche de documents bibliographiques.

Enfin, j'adresse mes chaleureuses pensées à ma famille, Sonia et mes amis, pour leur soutien et leurs encouragements tout au long de ces trois années, ainsi que mes remerciements aux personnes, notamment Muriel et mes parents, qui m'ont permis d'améliorer la qualité de ce mémoire par leur relecture attentive et leurs observations.

Résumé
Abstract

Résumé

La simulation de la transformation de la pluie en débit à l'échelle du bassin versant par des modèles mathématiques a connu un fort essor depuis le début des années 60 grâce notamment à l'accroissement des capacités de calcul. Il existe aujourd'hui un grand nombre de modèles, parmi lesquels on trouve les modèles conceptuels ou empiriques globaux qui représentent le lien entre la pluie et le débit par des agencements variés de réservoirs. Les études comparatives ayant impliqué de tels modèles ont été une réponse à la difficulté d'appréciation de leurs qualités et de leurs faiblesses respectives au travers de simples exercices de validation individuels. Cependant, les comparaisons réalisées jusqu'à présent, en ne considérant qu'un faible nombre de modèles et de bassins test et en n'appliquant pas toujours une procédure d'évaluation identique pour chaque modèle, ne permettent pas de dégager de conclusions claires quant aux valeurs respectives des différentes structures existantes.

Le cadre comparatif établi ici a pour but de tester un grand nombre de structures de modèles sur un large échantillon de bassins versants, afin d'explorer le rôle du nombre de paramètres optimisables et celui de la formulation du modèle sur la qualité des simulations de débit. Ainsi, 38 structures dérivées de modèles existants et comprenant au maximum neuf paramètres libres ont été évaluées au pas de temps journalier par une procédure de 'split-sample test' sur des données de 429 bassins versants situés en France, aux Etats-Unis, en Australie, en Côte-d'Ivoire et au Brésil. Toutes les structures ont reçu les mêmes données et leurs paramètres ont été calés à l'aide d'une même procédure locale d'optimisation, la méthode 'pas-à-pas', qui s'est montrée fiable pour localiser des optima satisfaisants. La qualité des simulations a été mesurée en phase de contrôle grâce à six critères quantitatifs faisant appel à diverses prises en compte de l'erreur du modèle et de la variable cible (le débit). De nouvelles formulations du critère de Nash et du critère de bilan, jugées plus satisfaisantes, ont été proposées.

Les résultats des tests indiquent tout d'abord que les modèles 'à réservoirs' sont nettement plus satisfaisants qu'un modèle de type 'boîte noire' ne comprenant pas de schéma interne de suivi d'humidité du bassin. L'ensemble des 38 structures testées ont paru assez proches dans leurs performances, suggérant en première approche une grossière équivalence de tous ces modèles. Néanmoins, des structures contenant de trois à cinq paramètres arrivent à obtenir d'aussi bons résultats que des structures ayant un plus grand nombre de degrés de libertés, cet avantage tenant d'une meilleure robustesse lors du passage des phases de calage au contrôle. Par ailleurs, la structure du modèle, c'est-à-dire sa formulation interne, est apparue déterminante pour le succès du modèle. Des complémentarités entre différentes structures ont été mises en évidence et ont suggéré des voies de modification de certaines d'entre elles. Partant de la structure simple du modèle GR3J, qui est initialement parmi les plus performantes, une nouvelle structure contenant quatre paramètres a été proposée. Elle a été jugée plus satisfaisante que des versions antérieures, notamment dans la simulation des étiages.

Abstract

Simulating the transformation of rainfall into runoff at the catchment scale using mathematical models has seen considerable developments since the early 1960s due to increasing computing capacities. There is today a large number of existing models, among which are the spatially lumped conceptual or empirical types that represent the link between rainfall and streamflow by a series of interconnected storages. Comparative studies on these types of models have evaluated their strengths and weaknesses. However comparisons carried out so far do not give clear conclusions on the worth of existing rainfall-runoff model structures, because typically they have included only a few models tested on a few catchments, and not always using homogeneous comparative frameworks.

The objective of the comparison scheme developed here is to assess a large number of model structures on a wide sample of catchments, to investigate the roles of the number of optimized parameters and model mathematical formulation on the quality of streamflow simulations. Thus 38 model structures derived from existing models and including at most nine free parameters were evaluated using a split-sample testing scheme on data from 429 catchments located in France, the United-States, Australia, the Ivory Coast and Brazil. All structures were fed with the same amount of data and their parameters were calibrated by a reliable ‘step-by-step’ local optimization algorithm. Simulation quality was assessed in validation by six quantitative criteria that appeal to various types of model error and different mathematical forms of the target variable (streamflow). Variants of the established Nash-Sutcliffe and other (water balance) criteria were proposed and applied.

Results first indicate that all storage-type models are far better than a simple ‘black box’ model that does not include any soil moisture accounting procedure. All 38 tested models obtain quite similar results, which suggest that they are all roughly equivalent. Nevertheless, some structures with only three to five parameters give results as good as those from structures with a larger number of degrees of freedom. The advantage of models having only a few parameters is shown to be their robustness when going from model calibration to model verification. Model structure, i.e. the internal mathematical formulation, is of prime importance for the success of the model. It was shown that model structures are complementary, leading to modifications of some of them. Starting from the simple GR3J model that proved initially to be among the best performing models, a modified structure with four parameters was proposed. It was found to be more satisfactory than previous versions, especially for low flow simulations.

Table des matières

Liste des tableaux

Liste des figures

Table des matières

Remerciements	3
Résumé – Abstract	7
Introduction générale	25
Chapitre 1. Modélisation hydrologique et modèles pluie-débit - Contexte d'un travail de comparaison	31
<hr/>	
1.1. Introduction.....	31
1.2. Hydrologie et modélisation pluie-débit	31
1.2.1. Objet de l'hydrologie	31
1.2.2. Qu'est qu'un modèle ?	33
1.2.3. Objet et enjeux de la modélisation pluie-débit.....	34
1.2.4. Une classification des modèles pluie-débit.....	36
1.2.5. Etudes comparatives: une évaluation objective des modèles ?.....	42
1.2.6. Quels objectifs pour une nouvelle étude comparative ?	45
1.3. Sélection de structures de modèles	46
1.3.1. Choix des modèles	46
1.3.2. Les enjeux de la parcimonie.....	47
1.3.3. Modifications des modèles et présentation des versions retenues	49
1.4. Méthodologie de comparaison.....	52
1.4.1. Choix d'une démarche d'évaluation	52
1.4.2. Stratégie de comparaison	53
1.5. Conclusion	53
Chapitre 2. Description de l'échantillon de données utilisé pour tester les modèles pluie-débit	57
<hr/>	
2.1. Introduction.....	57
2.2. Observations, instruments de connaissance du bassin et de construction de modèles ...	57
2.3. Description de l'échantillon.....	58
2.3.1. Données en France	58
2.3.2. Données aux Etats-Unis	60
2.3.3. Données en Australie	62
2.3.4. Données en Côte d'Ivoire	63
2.3.5. Données au Brésil	64
2.4. Quelques caractéristiques des bassins.....	64
2.5. Quelle critique des données dans le contexte de notre étude ?	66
2.5.1. Les facteurs limitant l'application des modèles.....	66

2.5.2.	Qualité des données	68
2.5.3.	Quelle attitude adopter dans notre étude ?	68
2.6.	Identification de périodes pour tester les modèles	69
2.6.1.	Choix des périodes	69
2.6.2.	Initialisation des modèles	71
2.7.	Classification des bassins en fonction de descripteurs hydro-climatiques	72
2.7.1.	Introduction	72
2.7.2.	Choix de variables explicatives	73
2.7.3.	Résultats de la classification	74
2.8.	Conclusion	76

Chapitre 3. Choix et évaluation d'une méthode pour le calage des paramètres des modèles pluie-débit: la méthode 'pas-à-pas'

3.1.	Introduction	81
3.2.	Optimisation en hydrologie: une problématique complexe	82
3.3.	Quelques méthodes de calage	83
3.3.1.	Les méthodes locales	84
3.3.2.	Les méthodes globales	85
3.4.	Avantages et inconvénients des méthodes de calage: des tests comparatifs	87
3.4.1.	Quelques avantages et inconvénients des méthodes locales	87
3.4.2.	Quelques avantages et inconvénients des méthodes globales	88
3.4.3.	Méthodes locales contre méthodes globales	89
3.5.	Choix d'une méthode d'optimisation	90
3.5.1.	Vers une méthode locale	90
3.5.2.	La méthode 'pas-à-pas'	90
3.6.	Une évaluation de la méthode 'pas-à-pas': influence du point de départ	93
3.6.1.	Introduction	93
3.6.2.	Méthodologie d'évaluation	93
3.6.3.	Classification et points de départ des optimisations	94
3.6.4.	Analyse des paramètres calés	98
3.6.5.	Performances en calage	100
3.6.6.	Performances en contrôle	104
3.6.7.	Conclusion	106
3.7.	Effet de la longueur de la période de calage	107
3.8.	Conclusion	110

Chapitre 4. Sélection de critères d'évaluation des performances des modèles pluie-débit

4.1.	Introduction	115
4.2.	Première sélection de critères et transformation sur les débits	117
4.2.1.	Critère d'erreur quadratique	117
4.2.2.	Critère d'erreur absolue	118
4.2.3.	Critère d'erreur cumulée	119
4.2.4.	Transformations préalables des débits	119
4.2.5.	Relations entre critères	120
4.3.	Critère de Nash: analyse et proposition d'une nouvelle formulation	121
4.3.1.	Quelques réflexions sur le critère de Nash-Sutcliffe	121
4.3.2.	Objectifs de la mise au point d'un nouveau critère	124
4.3.3.	Analyse de la problématique	125

4.3.4.	Evaluation de différents critères et des modèles de référence associés	127
4.3.5.	Quelques conclusions.....	137
4.4.	Proposition d'un nouveau critère de bilan	138
4.5.	Optimisation et évaluation multi-critère	140
4.6.	Conclusion	142
Chapitre 5. Présentation et analyse des résultats de la comparaison de structures.....		147
5.1.	Introduction.....	147
5.2.	Résultats et discussions.....	147
5.2.1.	Quel niveau de performance peut-on attendre des modèles testés ?.....	149
5.2.2.	Comment les résultats en calage et contrôle illustrent la robustesse des modèles ?	156
5.2.3.	Certaines structures peuvent-elles garantir une meilleure fiabilité ?	157
5.2.4.	Peut-on parler d'équifinalité entre structures de modèles ?.....	159
5.2.5.	Est-ce que le nombre de paramètres accroît le niveau de performance ?	159
5.2.6.	Complémentarité de structures de modèles: est-il possible de se hisser jusqu'au modèle 'idéal' ?.....	161
5.2.7.	Recherche d'une typologie bassins-modèles	164
5.3.	Conclusion	169
Chapitre 6. Exploration du modèle GR3J au travers d'une comparaison de deux méthodes d'analyse d'incertitude des paramètres.....		173
6.1.	Introduction.....	173
6.2.	Erreurs et incertitudes en modélisation pluie-débit	174
6.3.	Méthodes pour l'analyse d'incertitudes sur les paramètres	176
6.3.1.	Analyse d'incertitudes par approximation linéaire	176
6.3.2.	La procédure 'multi-calages'	177
6.4.	Le modèle GR3J	178
6.5.	Bassins, données et périodes.....	179
6.5.1.	Choix d'un échantillon de bassins	179
6.5.2.	Brève analyse sur les périodes sélectionnées.....	181
6.6.	Méthodologie	182
6.7.	Résultats et discussion	184
6.7.1.	Performances des différentes structures du modèle	184
6.7.2.	Valeurs des paramètres calés	189
6.7.3.	Corrélations entre paramètres	193
6.7.4.	Incertainitude sur les paramètres	196
6.8.	Conclusion	198
Chapitre 7. Tentative d'amélioration du modèle GR3J par une approche empirique..		203
7.1.	Introduction.....	203
7.2.	Rappels sur les résultats de la comparaison de modèles.....	204
7.2.1.	Performances et fiabilité des modèles.....	204
7.2.2.	Complémentarité entre modèles.....	205
7.2.3.	Sélection des structures et commentaires	206
7.3.	Méthodologie de recherche.....	208
7.4.	Vers une nouvelle version ?.....	208
7.4.1.	Les versions testées.....	208
7.4.2.	Choix d'un niveau de complexité	210

7.4.3.	Description de la structure modifiée retenue	212
7.5.	Comparaison aux versions antérieures de GR – Critique de la nouvelle structure.....	215
7.5.1.	Vérification des résultats en utilisant d'autres fonctions objectif.....	215
7.5.2.	Version modifiée et versions antérieures du modèle GR journalier.....	217
7.5.3.	Comparaison aux autres structures testées.....	220
7.5.4.	Stabilité des paramètres	221
7.6.	Performance de la nouvelle structure sur des classes de bassins.....	223
7.7.	Conclusion	224

Chapitre 8. Utilisation dans un modèle hydrologique de l'information contenue dans des descripteurs du bassin versant 229

8.1.	Introduction.....	229
8.2.	Approches de régionalisation.....	230
8.3.	Description d'une méthodologie améliorée.....	232
8.4.	Choix du modèle et travaux antérieurs	235
8.5.	Application.....	239
8.5.1.	Choix des bassins.....	239
8.5.1.1.	Calage du modèle et sélection des bassins	239
8.5.1.2.	Analyse de l'échantillon des paramètres	239
8.5.2.	Choix et analyse de variables explicatives.....	240
8.5.3.	Recherche de relations entre paramètres du modèle et variables explicatives ..	242
8.5.3.1.	Méthodologie de recherche	242
8.5.3.2.	Régressions sur les paramètres transformés	243
8.5.3.3.	Régressions sur les paramètres réels	247
8.5.4.	Résultats de l'application de la méthode proposée	251
8.5.4.1.	Vérification directe	252
8.5.4.2.	Recherche des coefficients constants les plus efficaces	253
8.5.4.3.	Recherche de coefficients de régression plus efficaces.....	254
8.5.4.4.	Vérification sur l'échantillon des 429 bassins.....	257
8.5.5.	Discussion	258
8.6.	Conclusion	260

Conclusion générale 265

Bibliographie..... 273

Annexe 1. Description des modèles.....	291
Fiche analytique 1 ABCD.....	293
Fiche analytique 2 Arno	297
Fiche analytique 3 Modèle B et C de Bonvoisin et Boorman.....	302
Fiche analytique 4 mSFB.....	308
Fiche analytique 5 Bucket.....	313
Fiche analytique 6 CATPRO	317
Fiche analytique 7 Cequeau	322
Fiche analytique 8 CREC.....	327
Fiche analytique 9 Gardenia	333
Fiche analytique 10 Modèle de Georgakakos and Baumer	337
Fiche analytique 11 GR3J et GR4J	341
Fiche analytique 12 GR5J.....	352
Fiche analytique 13 GRHUM.....	356
Fiche analytique 14 Modèle de Haan.....	362

Fiche analytique 15	HBV	366
Fiche analytique 16	HMS	372
Fiche analytique 17	IHACRES	376
Fiche analytique 18	Institute of Hydrology Lumped Model	383
Fiche analytique 19	Martine	387
Fiche analytique 20	MHR	391
Fiche analytique 21	MODALP	397
Fiche analytique 22	MODGLO	401
Fiche analytique 23	MODHYDROLOG	406
Fiche analytique 24	NAM	411
Fiche analytique 25	Dawdy et O'Donnell	416
Fiche analytique 26	PDM	420
Fiche analytique 27	Sacramento	425
Fiche analytique 28	SDI	430
Fiche analytique 29	SIXPAR	434
Fiche analytique 30	SMAR	438
Fiche analytique 31	TANK	443
Fiche analytique 32	TMWAM	448
Fiche analytique 33	TOPMODEL	452
Fiche analytique 34	Wageningen	459
Fiche analytique 35	Xinanjiang	463
Annexe 2. Présentation des bassins versants		471
Annexe 2.1. Présentation des bassins versants en France		471
Annexe 2.2. Présentation des bassins versants aux Etats-Unis (ARS et MOPEX)		477
1. Données ARS		477
2. Données MOPEX		480
Annexe 2.3. Présentation des bassins versants en Australie		483
Annexe 2.4. Présentation des bassins versants en Côte-d'Ivoire		485
Annexe 2.5. Présentation des bassins versants au Brésil		488
Annexe 3. Article de présentation des résultats soumis pour publication au Journal of Hydrology (Deuxième version – Juillet 2000 – En cours de révision)		493
Annexe 4. Description de la méthode d'analyse d'incertitudes par approximation linéaire		523
Annexe 5. Rôle de l'initialisation sur les performances du modèle		527

Liste des tableaux

Tableau 1.1: Liste des modèles sélectionnés, avec le code et le nombre de paramètres de la version retenue	50
Tableau 2.1: Répartition des bassins, périodes de calage et tests en contrôle en fonction des différents pays d'origine (entre parenthèses, proportion du total en %).....	70
Tableau 2.2: Composition des classes de bassins versants.....	75
Tableau 2.3: Caractéristiques des cinq classes de bassins versants (en gras, valeurs maximales pour les 5 classes; en italique, valeurs minimales)	75
Tableau 3.1: Caractéristiques des trois classifications successives effectuées sur le nuage des paramètres optimisés	95
Tableau 3.2: Modifications relatives moyennes (en %) des coordonnées des centres de gravités des classes.....	97
Tableau 3.3: Différences absolues moyennes et relatives (en %) des paramètres d'une optimisation à l'autre....	98
Tableau 3.4: Nombre de convergences vers le même optimum (sur un total de 856 optimisations)	99
Tableau 3.5: Nombre de convergence vers des optima similaires (sur un total de 856 optimisations).....	99
Tableau 3.6: Performances globales sur l'échantillon pour les trois séries de calage (maximum, moyenne, minimum des performances et pourcentage F(0) de critères négatifs).....	102
Tableau 3.7: Différences dans les valeurs de la fonction objectif entre les trois optimisations	102
Tableau 3.8: Performances des quatre structures de modèle en contrôle (performances maximales, minimales et moyennes et pourcentage F(0) des critères négatifs)	104
Tableau 3.9: Différences de performances entre les trois séries de contrôles	105
Tableau 3.10: Proportions des calages (sur un total de 856) et de contrôles (sur un total de 1780) pour lesquels il y a eu amélioration et dégradation du critère de performance en calage et en contrôle, entre la première et deuxième série de tests.....	106
Tableau 4.1: Critères d'évaluation utilisés dans des études comparatives de modèles pluie-débit	116
Tableau 5.1: Ecart maximaux entre modèles pour la moyenne et les quantiles 0,1 et 0,9 en calage et en contrôle sur les six critères de qualité	153
Tableau 5.2: Diminutions maximales, minimales et moyennes du critère CR2 pour les 38 structures de la phase de calage à la phase de contrôle.....	157
Tableau 5.3: Partie de la matrice de complémentarité montrant les meilleures associations entre structures	164
Tableau 5.4 : Liste des modèles apparaissant pour N1 critères de performances parmi les six meilleurs modèles (par classes de bassins)	167
Tableau 5.5: Performances moyennes (critère CR2) des différentes associations de modèles sur les 429 bassins (calculées à partir des performances moyennes par bassin)	167
Tableau 6.1: Liste des 17 bassins américains étudiés	180
Tableau 7.1: Structures ayant obtenu les meilleurs résultats moyens suivant les cinq critères de qualité	205
Tableau 7.2: Structures classées le plus grand nombre de fois parmi les six meilleures structures	205
Tableau 7.3: Meilleures associations de modèles incluant GR3J et nombres de bassins correspondants pour les critères (a) CR1, (b) CR2, (c) CR3, (d) CR4 et (e) CR5.....	206
Tableau 7.4: Description des versions de GR utilisées à titre comparatif.....	216
Tableau 7.5: Caractéristiques des distributions des superficies des deux échantillons de bassins	219
Tableau 7.6: Performances moyennes par bassin du modèle modifié et du modèle composite.....	224

Tableau 8.1: Travaux de régionalisation sur différents modèles pluie-débit (*: travaux en transposition).....	231
Tableau 8.2: Matrice de corrélation entre paramètres	240
Tableau 8.3: Caractéristiques de l'échantillon des 131 jeux de paramètres	240
Tableau 8.4: Caractéristiques des bassins utilisés pour établir les équations de prédétermination des paramètres.....	241
Tableau 8.5: Matrice de corrélation entre variables explicatives	241
Tableau 8.6: Relations entre paramètre transformé $\ln(A)$ et variables explicatives	243
Tableau 8.7: Relations entre paramètre transformé $\ln(B)$ et variables explicatives	244
Tableau 8.8: Relations entre paramètre transformé $\ln(C-0,5)$ et variables explicatives.....	245
Tableau 8.9: Relations entre paramètre transformé $\operatorname{argsh}(D)$ et variables explicatives	246
Tableau 8.10: Relations entre paramètre A et variables explicatives	247
Tableau 8.11: Relations entre paramètre B et variables explicatives	248
Tableau 8.12: Relations entre paramètre C et variables explicatives	249
Tableau 8.13: Relations entre paramètre D et variables explicatives	251
Tableau 8.14: Solutions de prédétermination des paramètres	252
Tableau 8.15: Valeur du critère Cr pour les quatre solutions initiales	253
Tableau 8.16: Solution après ajustement des paramètres constants	254
Tableau 8.17: Solution 2' après ajustement des coefficients	255
Tableau 8.18: Solution 4' après ajustement des coefficients	255
Tableau 8.19: Valeurs du critère Cr pour les différentes solutions sur l'échantillon des 429 bassins	258

Liste des figures

Figure 1.1: Le cycle hydrologique, avec les flux moyens annuels en pourcentage du volume annuel total des précipitations sur les terres émergées (d'après Maidment, 1992).....	32
Figure 1.2: Représentation schématique de TOPMODEL (d'après Beven et Kirkby, 1979).....	40
Figure 2.1: Localisation des 307 stations hydrométriques françaises	59
Figure 2.2: Localisation des sites des bassins versants américains expérimentaux issus de la base de données ARS	60
Figure 2.3: Localisation des stations hydrométriques des bassins versants issus de la base MOPEX	61
Figure 2.4: Localisation des 26 stations hydrométriques australiennes	62
Figure 2.5: Localisation des bassins de Côte-d'Ivoire (adapté d'après Dezetter, 1991 et Ouédraogo, 1996).....	63
Figure 2.6: Localisation des bassins versants brésiliens (adapté d'après Melo et Nascimento, 1999).....	64
Figure 2.7: Distribution des superficies des 429 bassins versants.....	65
Figure 2.8: Pluie et débit annuels moyens des 429 bassins versants.....	65
Figure 2.9: Chroniques pluie-débit pour le bassin de la Mountain Fork à Eagletown, Oklahoma, aux Etats-Unis (bassin US339000, MOPEX).....	67
Figure 2.10: Répartition des bassins en fonction de la longueur des séries disponibles	69
Figure 2.11: Répartition des différentes périodes en fonction de leur longueur	70
Figure 2.12: Exemple de division de la période d'étude totale en quatre sous-périodes de 9 ans, avec une année de mise en route	72
Figure 2.13: Répartition des bassins pour deux variables NCRU et BFI.....	76
Figure 3.1: Diagramme schématique de la méthode de calage 'pas-à-pas'.....	91
Figure 3.2: Projections du nuage de points sur des plans de l'espace des paramètres transformés du modèle GR4J, avec les deux centres de gravité des classes	95
Figure 3.3: Bassins versants appartenant à l'une des classes de la classification.....	98
Figure 3.4: Distributions des critères d'optimisation pour les structures (a) GR4J et (b) TOPM	101
Figure 3.5: Distributions des critères de performance en contrôle pour les structures (a) GR4J et (b) TOPM... ..	105
Figure 3.6: Nombre de périodes de calage par bassin	108
Figure 3.7: Répartition des périodes en fonction de leur longueur	109
Figure 3.8: Performances des modèles en contrôle en fonction du nombre de paramètres pour (a) le critère Nash(Q), et (b) le critère Nash(\sqrt{Q})	110
Figure 4.1: Critères obtenus par le modèle GR4K en contrôle sur l'échantillon de bassins après calage en utilisant le critère Nash(\sqrt{Q}): comparaison (a) des critères Nash(Q) et Nash(ln(Q)), et (b) du critère Nash(Q) et du critère d'erreur absolue défini à l'Eq. (4.5).....	121
Figure 4.2: Hydrogrammes observés et simulés sur la période 1989-1992 pour le bassin versant du Loing à Chalette-sur-Loing.....	122
Figure 4.3: Comparaison des valeurs du critère de Nash avec comme modèle de référence le débit moyen de la période considérée (cas 1) et le débit moyen de la période de calage (cas 2) sur les 3204 tests en contrôle	122
Figure 4.4: Distributions des critères de Nash (%) obtenus en contrôle par le modèle GR4K sur l'échantillon de 429 bassins, avec distinction des bassins à faible et fort rendements Cr.....	123
Figure 4.5: Distributions des valeurs du critère de Nash obtenues par le modèle GR4K sur les cinq classes de bassins en contrôle	124

Figure 4.6: Comportement schématique du critère de Nash en fonction de la qualité ‘réelle’ des simulations du modèle.....	125
Figure 4.7: Schéma des plages de variation des opérateurs de production et de routage pour le choix d’un modèle de référence	126
Figure 4.8: Comparaison des modules observés et simulés d’après le modèle de Tixeront (1964) pour les 429 bassins versants (d’après Mouelhi, 2000).....	127
Figure 4.9: Comparaison des valeurs du critère $C1$ et du critère de Nash sur l’échantillon de bassins en contrôle ($\alpha=1$ et $\beta=1/4$)	128
Figure 4.10: Distribution des valeurs du critère 1 ($\alpha=1$ et $\beta=1/4$) obtenues en contrôle, avec distinction des bassins à faible et fort rendements Cr.....	129
Figure 4.11: Comparaison des valeurs du critère $C1$ et du critère de Nash sur l’échantillon de bassins en contrôle ($\alpha=1/10$ et $\beta=2/3$)	129
Figure 4.12: Distribution des valeurs du critère 1 ($\alpha=1/10$ et $\beta=2/3$) obtenues en contrôle, avec distinction des bassins à faible et fort rendements Cr.....	130
Figure 4.13: Distributions des valeurs du critère 1 obtenues sur les cinq classes de bassins en contrôle ($\alpha=1/10$ et $\beta=2/3$)	130
Figure 4.14: Distributions des valeurs du critère 5 obtenues sur les cinq classes de bassins en contrôle ($\alpha=1/3$ et $\beta=2/3$)	132
Figure 4.15: Distribution des valeurs du critère 5 ($\alpha=1/3$ et $\beta=2/3$) obtenues en contrôle, avec distinction des bassins à faible et fort rendements Cr.....	133
Figure 4.16: Comparaison des valeurs du critère 5 et du critère de Nash sur l’échantillon de bassins en contrôle ($\alpha=1/5$ et $\beta=4/5$)	133
Figure 4.17: Distribution des valeurs du critère 5 ($\alpha=1/5$ et $\beta=4/5$) obtenues en contrôle, avec distinction des bassins à faible et fort rendements Cr.....	134
Figure 4.18: Distributions des valeurs du critère 5 obtenues sur les cinq classes de bassins en contrôle ($\alpha=1/5$ et $\beta=4/5$)	134
Figure 4.19: Comparaison des valeurs du critère 6 ($\alpha=1/3$ et $\beta=4/5$) au critère 5 sur l’échantillon de bassins en contrôle ($\alpha=1/5$ et $\beta=4/5$).....	135
Figure 4.20: Comparaison des valeurs du critère $C7$ et du critère de Nash sur l’échantillon de bassins en contrôle ($\alpha=1/2$ et $\beta=1$)	136
Figure 4.21: Distributions des valeurs du critère 7 obtenues sur les cinq classes de bassins en contrôle ($\alpha=1/2$ et $\beta=1$)	137
Figure 4.22: Distribution des valeurs de B au contrôle sur l’échantillon de bassins versants pour le modèle GR4K.....	138
Figure 4.23: Courbes des critères $B1$, $B2$ et $B3$ en fonction de B	139
Figure 4.24: Distributions des critères $B1$, $B2$ et $B3$ obtenus en contrôle sur l’échantillon test	139
Figure 5.1: Exemple de distribution expérimentale des résultats d’un modèle, avec détermination des quantiles 0,1, 0,3 et 0,9	148
Figure 5.2: Performances des structures de modèles au calage pour le critère de calage $CR2$ (Nash(\sqrt{Q})), avec moyennes et quantiles 0,1, 0,3 et 0,9.....	149
Figure 5.3: Performances des structures de modèles au contrôle pour les critères (a) $CR1$ (Nash(Q)) et (b) $CR2$ (Nash(\sqrt{Q})), avec moyennes et quantiles 0,1, 0,3 et 0,9	150
Figure 5.4: Rang des structures suivant les six critères de qualité au contrôle	154
Figure 5.5: Comparaison des performances avec et sans critique préalable des données pour les critères (a) $CR1$ et (b) $CR2$	155
Figure 5.6: Lien entre superficie des 429 bassins versants de l’échantillon et performances moyennes par bassin obtenues en contrôle par le modèle GR4K suivant le critère $CR2$	156
Figure 5.7: Nombres de bassins pour lesquels le modèle est de classes 1 et 2.....	157
Figure 5.8: Performances moyennes des structures suivant le critère $CR2$ (Nash(\sqrt{Q})) sur les sous-échantillons de données (a) françaises et (b) étrangères	158
Figure 5.9: Performances moyennes des modèles en fonction du nombre de paramètres optimisés, suivant le critère $CR2$ (Nash(\sqrt{Q})) (a) au calage et (b) au contrôle; (c):diminution des performances entre calage et contrôle	160
Figure 5.10: Distribution des performances des modèles TSYK et IDEA et des 38 structures sur l’échantillon des 429 bassins en contrôle suivant le critère $CR1$	162
Figure 5.11: Composition du modèle idéal suivant le critère $CR1$: (a) nombre de bassins où les modèles ont été classés premiers ; (b) graphe de correspondance bassins-modèles	163
Figure 5.12: Performances moyennes des modèles sur les cinq classes de bassins pour le critère $CR2$	165

Figure 5.13: Evolution des classements des performances du modèle WAGE sur chacune des classes pour les six critères de performance.....	166
Figure 5.14: Distributions des performances suivant le critère <i>CR2</i> de diverses associations de modèles sur l'échantillon des 429 bassins	167
Figure 5.15: Projection des 429 bassins sur les premiers axes principaux de l'ACP bassins-modèles	168
Figure 6.1: Structure du modèle GR utilisé pour tester les méthodes d'analyse d'incertitudes	178
Figure 6.2: ETP moyenne annuelle et coefficient d'écoulement moyen en fonction de la pluie moyenne annuelle sur les 17 bassins versants étudiés.....	180
Figure 6.3: Variations des pluies et ETP moyennes pour des périodes de (a) 9 ans et (b) 1 an sur les 17 bassins versants (avec une seule valeur d'ETP interannuelle par bassin).....	181
Figure 6.4: Rapport du débit à la pluie (Q/P) en fonction du rapport de la pluie à l'ETP (P/E) pour des périodes de (a) 9 ans et (b) 1 an sur les 17 bassins versants	182
Figure 6.5: Schéma de la méthodologie de comparaison des deux méthodes d'analyse d'incertitudes (partie gauche du graphique: procédure 'multi-calage'; partie droite: procédure AIAL)	183
Figure 6.6: Distributions des résultats obtenus par le modèle GR41 (a) en calage et (b) en contrôle.....	184
Figure 6.7: Distributions des résultats en calage et en contrôle pour le modèle GR41 en utilisant des périodes de 9 ans (a) indépendantes et (b) non indépendantes	186
Figure 6.8: Distributions des résultats en calage et en contrôle pour les structures GR31, GR41, GR42, GR51 dans les cas de périodes de (a) 9 ans, (b) 4 ans et (c) 1 an.....	188
Figure 6.9: Comparaison des paramètres moyens \bar{X}_4 et des paramètres de référence X_4^0	190
Figure 6.10: Evolution des paramètres 'régionaux' X_1 (a) et X_2 (b) avec la longueur de la période.....	191
Figure 6.11: Distributions des valeurs de X_1 pour la structure GR51 pour différentes longueurs de périodes ..	191
Figure 6.12: Liens entre les paramètres X_1 et X_4 de la structure GR51 et le rendement du bassin pour des périodes (a1,a2) de 9 ans et (b1,b2) de 1 an.....	192
Figure 6.13: Corrélations entre les paramètres X_1 et X_2 de la structure GR31 fournies par approche 'multi-calage' et méthode AIAL (valeurs moyennes par bassin) pour les cinq longueurs de périodes	194
Figure 6.14: Corrélations entre les paramètres X_1 et X_2 de la structure GR31 fournies par approche 'multi-calage' et méthode AIAL (valeurs non moyennées par bassin) pour les cinq longueurs de périodes.....	194
Figure 6.15: Corrélations moyennes entre paramètres X_1 et X_2 obtenues par AIAL (s) et par 'multi-calage' (m).....	195
Figure 6.16: Comparaison des écarts-types du paramètre X_4 de la structure GR41 fournis par les méthodes AIAL et 'multi-calage'	196
Figure 6.17: Ecarts-types du paramètre X_1 fournis par la méthode AIAL (s) et par 'multi-calage' (m) en fonction de la longueur de la période.....	197
Figure 6.18: Distributions des valeurs du paramètre X_1 du modèle GR41 en prenant 4 périodes de 9 ans, 28 périodes de 9 ans ou 36 périodes de 1 an.....	197
Figure 7.1: Schéma de la structure du modèle GR3J et modifications testées en pointillés.....	209
Figure 7.2: Evolution des performances des structures modifiées en fonction du nombre de paramètres optimisés pour les critères d'évaluation <i>CR1</i> , <i>CR2</i> , <i>CR3</i> et <i>CR6</i>	212
Figure 7.3: Performances moyennes suivant quatre critères de 30 structures modifiées à quatre paramètres, ordonnées par performance décroissante suivant <i>CR2</i>	213
Figure 7.4: Schéma structurel du modèle GR4J', avec en pointillés les modifications introduites par rapport à la version GR3J (X_1 , X_2 , X_3 et X_4 sont les paramètres du modèle).....	214
Figure 7.5: Représentation de l'hydrogramme HU2	215
Figure 7.6: Performances comparées en contrôle sur les critères (a) <i>CR1</i> , (b) <i>CR2</i> , (c) <i>CR3</i> et (d) <i>CR6</i> , de plusieurs structures calées sur différentes fonctions objectif.....	216
Figure 7.7: Comparaison des valeurs du paramètre X_2 obtenues par calage avec les fonctions objectif <i>CR1</i> et <i>CR3</i>	217
Figure 7.8: Performances comparées (quantiles 0,3) la version hybride GR4J' et de plusieurs versions antérieures du modèle GR journalier à trois ou quatre paramètres sur (a) l'échantillon total de bassins, (b) sur l'échantillon de bassins français et (c) sur l'échantillon de bassins étrangers.....	218
Figure 7.9: Comparaison des distributions des superficies des 60 bassins de l'échantillon de Rakem (1999) et de l'échantillon des 429 bassins utilisés ici.	219
Figure 7.10: Comparaison des performances de GR3J et GR4J' (a) sur tous les contrôles pour le critère <i>CR1</i> , sur les performances moyennes par bassin (b) sur le critère <i>CR1</i> et (c) sur le critère <i>CR3</i>	220
Figure 7.11: Distribution des résultats moyens par bassins suivant le critère <i>CR2</i> pour les modèles GR4K, une association de modèles, le modèle IDEA et le modèle GR4J'	221

Figure 7.12: Comparaison des fonctions de répartition des quatre paramètres des versions GR4K et GR4J'	222
Figure 7.13: Comparaison des valeurs de la capacité du réservoir de production (X4) et de la capacité à un jour (X5) intervenant dans le calcul de la percolation optimisées indépendamment	222
Figure 7.14: Performances des versions modifiées sur les cinq sous-échantillons de bassins (avec signe distinctif pour GR4J')	223
Figure 8.1: Représentation schématique des procédures de régionalisation classique et améliorée	234
Figure 8.2: Lien entre les variables <i>EAM</i> et <i>CP</i>	242
Figure 8.3: Graphe de corrélation entre valeurs transformées de <i>A</i> calées et calculées par l'équation de prédétermination	244
Figure 8.4: Graphe de corrélation entre valeurs transformées de <i>B</i> calées et calculées par l'équation de prédétermination	245
Figure 8.5: Graphe de corrélation entre valeurs transformées de <i>C</i> calées et calculées par l'équation de prédétermination	246
Figure 8.6: Graphe de corrélation entre valeurs transformées de <i>D</i> calées et calculées par l'équation de prédétermination	247
Figure 8.7: Graphe de corrélation entre valeurs réelles de <i>A</i> calées et calculées par l'équation de prédétermination	248
Figure 8.8: Graphe de corrélation entre valeurs réelles de <i>B</i> calées et calculées par l'équation de prédétermination	249
Figure 8.9: Graphe de corrélation entre valeurs réelles de <i>C</i> calées et calculées par l'équation de prédétermination	250
Figure 8.10: Graphe de corrélation entre valeurs réelles de <i>D</i> calées et calculées par l'équation de prédétermination	250
Figure 8.11: Distributions des résultats obtenus avec les paramètres calés, constants et prédéterminés	253
Figure 8.12: Distributions des résultats obtenus avec des paramètres constants	254
Figure 8.13: Distributions des résultats avant et après ajustement des coefficients des formulations de prédétermination de la solution 2	255
Figure 8.14: Distributions des résultats avant et après ajustement des coefficients des formulations de prédétermination de la solution 4	256
Figure 8.15: Comparaison des distributions des résultats avec les solutions de prédétermination après ajustement des coefficients, sur l'échantillon des 131 bassins	256
Figure 8.16: Distributions des résultats obtenus sur les 429 bassins en utilisant les solutions initiales de prédétermination	257
Figure 8.17: Distributions des résultats obtenus sur les 1284 périodes de calage après ajustement des formules de prédétermination	258
Figure 8.18: Distribution des résultats obtenus au contrôle pour les modèles testés dans la comparaison et le modèle GR4J' avec paramètres constants	260

Introduction générale

Introduction générale

La Terre est la seule planète du système solaire à disposer, grâce à sa position privilégiée, d'eau liquide à sa surface. La petite molécule H₂O y est indispensable pour le développement de la vie. Habitat d'une partie de la faune et de la flore, vecteur de matières, érodant et modifiant les paysages, élément vital d'alimentation, l'eau devient de plus en plus, avec l'accroissement de la population humaine sur Terre, une denrée convoitée. L'eau douce, objet principal d'enjeux, ne constitue que 3 % de la ressource mondiale (le restant étant formé par les mers et les océans), dont les trois quarts sont stockés sous forme de glace. Dans son ouvrage *La Bataille de l'Eau*, Roger Cans (1994) dresse un tableau parfois alarmant de l'avenir de cette ressource. Sa répartition quantitative à l'échelle du globe est très hétérogène, et souvent très différente de la répartition des populations. Enjeu de vie, l'eau devient de plus en plus une question politique, source de tensions croissantes entre Etats, notamment au Proche-Orient. La variabilité de sa disponibilité est également source de problèmes: trop rare, elle entraîne des situations de pénurie, de désertification, d'exode de populations; trop abondante, elle est la cause d'inondations catastrophiques et rappelle à l'homme l'impossibilité d'en maîtriser les forces. Sa qualité est également très variable, soumise de plus en plus aux déséquilibres des écosystèmes aquatiques liés aux pollutions d'origine humaine. Quant à la qualité de l'eau de distribution, elle est l'enjeu d'une compétition cette fois économique que se livrent de grands groupes industriels, pour ce que certains nomment l'or bleu.

Devant cet accroissement de la demande en eau lié à la multiplication des usages (consommation domestique, industrie, agriculture, loisirs, etc.), le développement des problèmes environnementaux dus aux pollutions d'origine humaine, et les conséquences de choix d'aménagements hydrauliques faits par le passé apparaissant aujourd'hui peu judicieux, les gestionnaires de cette ressource doivent faire face actuellement à des situations de plus en plus complexes, où interviennent de multiples acteurs, aux intérêts et objectifs parfois opposés. Dans ce contexte, on comprend alors aisément la nécessité de mettre au point des outils d'aide à la gestion et à la décision qui permettent de mieux cerner le fonctionnement des hydro-systèmes naturels et le devenir de l'eau dans son environnement. L'utilisateur de ces instruments peut bénéficier ainsi d'une meilleure connaissance de la répartition spatiale et temporelle des flux d'eau et des matières et composés qu'elle véhicule à l'échelle du bassin.

Le cours d'eau constitue, dans cette démarche d'appréhension des hydrosystèmes, un objet d'intérêt privilégié, en raison de l'accès aisé qu'il offre à la ressource. La connaissance de son débit est aujourd'hui un instrument indispensable à la gestion des systèmes aquatiques (notamment pour la prévision de débits, la gestion d'ouvrages de retenue, le dimensionnement d'ouvrages hydrauliques, etc.). De nombreuses recherches se sont donc attachées, depuis plus d'un siècle, à essayer de comprendre les processus de génération des débits et le fonctionnement du bassin versant, entité hydrologique de production et de concentration des

écoulements. Un moyen pour arriver à comprendre la génération des débits et pour en établir des simulations est de remonter jusqu'à leur cause première, les pluies. Ainsi, on a essayé de construire des modèles permettant d'obtenir une représentation simplifiée et facilement utilisable de ce lien entre pluie et débit. Cependant, comme d'autres disciplines s'attachant à comprendre et représenter des systèmes naturels, la modélisation pluie-débit et plus généralement l'hydrologie continentale sont confrontées à la difficulté d'appréhension des systèmes étudiés du fait de leur complexité intrinsèque: leurs caractéristiques spatiales sont très hétérogènes, de nombreux facteurs influencent leur dynamique temporelle et leur observation reste encore aujourd'hui difficile et coûteuse.

Notre travail de recherche s'inscrit dans le domaine de la modélisation de la transformation de la pluie en débit et de sa représentation à l'échelle du bassin versant. Notre objectif principal est d'améliorer le fonctionnement d'un modèle de simulation des débits, celui du Génie Rural (GR), au travers d'une approche comparative. Pour cela, nous avons voulu réaliser une revue des différentes approches de modélisation et mettre en place un cadre de comparaison pour apprécier les forces et faiblesses de différentes représentations mathématiques existantes de la transformation pluie-débit.

Après avoir précisé la place qu'occupe la modélisation pluie-débit au sein des sciences hydrologiques, nous présentons dans le premier chapitre les approches de modélisation qui existent actuellement pour représenter le bassin versant en tant qu'entité hydrologique et simuler les débits. Nous discutons des problèmes rencontrés dans cette démarche, pour la représentation des processus et des facteurs influençant les écoulements. Nous détaillons davantage la description des modèles mathématiques globaux, conceptuels ou empiriques. La transformation de la pluie en débit y est simulée par des enchaînements de réservoirs et le bassin versant y est considéré comme une entité aux caractéristiques homogènes, sans prise en compte explicite de son hétérogénéité spatiale. De tels modèles donnent actuellement des résultats satisfaisants, en comparaison à des approches plus complexes, comme les modèles physiques qui découlent de l'application des lois physiques à des milieux dont les caractéristiques sont mesurées. Cependant, les paramètres des modèles globaux n'ont pas de signification physique a priori et doivent être calés par une procédure numérique.

Leur faible exigence en données et leur facilité de mise en œuvre les rendent compatibles avec les exigences des études d'ingénierie et d'hydrologie opérationnelle (dimensionnement d'ouvrage, gestion de la ressource, prévision de crues ou d'étiages...). Comme la plupart des modèles existants, ils ne sont pas encore aptes à pouvoir prévoir avec fiabilité des effets de changements d'occupation des sols ou de climat, mais ils peuvent a posteriori les mettre en évidence. Enfin, l'approche de modélisation globale renonce à vouloir décrire (faute de le pouvoir actuellement à l'échelle du bassin versant) la mécanique interne du bassin et les flux de matière et d'énergie en son sein.

Les modèles globaux, dont fait partie le modèle GR, existent aujourd'hui en très grand nombre et nous nous proposons d'évaluer les qualités de la structure de 38 d'entre eux par une approche comparative. Nous en exposons les modalités dans cette première partie. Dans tout notre travail, le terme de 'structure' sera utilisé pour désigner l'assemblage des différents composants du modèle, c'est-à-dire l'association de tous les outils mathématiques avec lesquels il est construit (réservoirs d'interception, de suivi d'humidité, de vidange, hydrogramme unitaire, etc.).

Pour réaliser ces tests, nous avons rassemblé tout d'abord un large échantillon de données. L'objectif est de soumettre les modèles à des conditions climatiques et hydrologiques très variées, et de leur demander de reproduire la transformation pluie-débit opérée par des bassins versants aux caractéristiques très diversifiées. Les 429 bassins versants étudiés,

principalement situés en France mais également aux Etats-Unis, en Australie, en Côte-d'Ivoire et au Brésil, sont présentés dans le deuxième chapitre.

La structure des modèles étudiés inclut des paramètres qui permettent l'adaptation du modèle aux caractéristiques particulières de chaque bassin. Ces paramètres n'ont généralement pas de signification physique et ne peuvent pas être, par conséquent, mesurés sur le terrain. Ils doivent être déterminés mathématiquement par un processus d'optimisation. Le choix d'une telle méthode repose sur son efficacité, c'est-à-dire sa capacité à fournir des jeux de paramètres permettant au modèle de donner des simulations satisfaisantes. Nous présentons dans le troisième chapitre les raisons du choix et une évaluation de la méthode d'optimisation sélectionnée pour déterminer les paramètres de tous les modèles testés.

Dans la quatrième partie de ce travail, nous exposons le choix de critères d'évaluation pour juger de la capacité des modèles à reproduire les débits observés. Nous menons notamment une réflexion sur la difficulté de trouver des tests globalement pertinents, ces derniers étant principalement dépendants des objectifs assignés au modèle par l'utilisateur. Certains critères présentent par ailleurs des formulations qui nous ont conduit à discuter de leur pertinence pour le jugement de la qualité des modèles. Partant de cette analyse, nous proposons une nouvelle formulation de certains critères. Nous exposons ensuite les raisons du choix d'un jeu de critères pour l'évaluation des simulations des modèles.

Dans le chapitre suivant, nous donnons les résultats des tests comparatifs, que nous analysons, au travers des critères de qualité précédemment sélectionnés, par une approche thématique. Il nous a paru notamment intéressant de voir quel degré de performance on peut attendre de ce type de modèles et d'évaluer leur robustesse et leur fiabilité. Nous avons aussi voulu mettre en évidence l'influence sur les simulations, du nombre de paramètres permettant dans le modèle de caractériser la relation pluie-débit sur chaque bassin. Enfin, nous explorons quel est le rôle de la structure du modèle sur ses performances et dans quelle mesure le test de ces formulations mathématiques différentes permet d'entrevoir des pistes d'amélioration des structures existantes, notamment en faisant apparaître les complémentarités entre elles.

Dans le sixième chapitre de ce travail, nous nous intéressons à la manière de déterminer les incertitudes sur les paramètres des modèles hydrologiques. Nous évaluons notamment une méthode classique d'analyse d'incertitude, en la comparant à une nouvelle approche permettant de mieux tenir compte de l'effet de la variabilité des données d'entrée sur les incertitudes. Ces deux méthodes, appliquées au modèle GR, donnent par ailleurs l'occasion d'étudier plus en détail l'organisation structurelle du modèle et le rôle des différents paramètres.

En partant de la structure simple du modèle GR à trois paramètres et en utilisant les résultats de l'approche comparative, nous avons tenté d'améliorer la qualité des simulations qu'il est possible d'obtenir par ce type de modélisation à réservoirs. Nous avons pour cela introduit dans la structure du modèle des formulations mathématiques issues de modèles différents, en couplant des outils qui apparaissent complémentaires. Nous rapportons dans le chapitre 7 les résultats des tests de ces différentes versions modifiées, dont les performances en terme de simulation des débits sont comparées à celles des autres structures testées dans la comparaison et des versions antérieures du modèle GR.

Enfin, dans le dernier chapitre de ce travail, nous explorons la possibilité de mettre en œuvre une approche améliorée de recherche de liens entre les valeurs des paramètres du modèle et des descripteurs du bassin. Nous appliquons la méthode au cas du modèle GR. La possibilité d'obtenir de bons estimateurs des paramètres à partir de caractéristiques du bassin facilement mesurables pourrait permettre l'application du modèle sur des bassins non jaugés, c'est-à-dire des bassins où les données nécessaires pour le calage numérique des paramètres ne sont pas

disponibles. Cependant, l'objectif de ce chapitre se limite simplement à évaluer quelle information pertinente pour la modélisation il est possible d'extraire de caractéristiques simples du bassin versant.

Chapitre 1

Chapitre 1

Modélisation hydrologique et modèles pluie-débit - Contexte d'un travail de comparaison

1.1. Introduction

La première partie de ce chapitre est destinée à situer la démarche de modélisation pluie-débit dans le contexte de l'hydrologie et à donner un aperçu des diverses approches développées dans ce domaine. Nous présentons trois classes de modèles permettant de différencier les grands types d'approches de construction de modèles destinés à simuler la transformation de la pluie en débit. Si la proposition de nouveaux modèles reste encore à l'heure actuelle assez fréquente dans les revues scientifiques hydrologiques, l'évaluation et la validation réelles de ces outils semblent cependant être beaucoup plus rares, du fait de la difficulté de tels exercices mais aussi certainement en partie à cause des déceptions que ces tests suscitent. Un moyen de rendre compte des capacités des modèles et de pallier certaines limitations des évaluations individuelles est de conduire des études comparatives dans lesquelles les modèles sont appliqués aux mêmes cas d'études. De telles approches peuvent également faire progresser dans la réflexion sur les modèles existants et suggérer l'émergence de nouveaux.

Partant d'une analyse synthétique des comparaisons de modèles, nous présentons dans un deuxième temps une sélection de modèles réalisée en vue d'une large comparaison. Les versions des modèles retenues et testées dans notre étude sont présentées globalement et leurs structures analysées en fonction de leur paramétrisation et de leur complexité.

Enfin, nous exposons la méthodologie de comparaison que nous avons adoptée dans notre étude.

1.2. Hydrologie et modélisation pluie-débit

Cette partie s'appuie notamment sur les réflexions de Nascimento et Michel (1992) sur la modélisation hydrologique et sur l'utilité des modèles pluie-débit.

1.2.1. Objet de l'hydrologie

Il est assez difficile de définir l'hydrologie car ce n'est pas une science unifiée. Etymologiquement "science qui traite des eaux", la définition change d'une époque ou d'un hydrologue à l'autre. Dans les dictionnaires, elle est parfois définie comme la science qui étudie les eaux, leurs caractéristiques, leurs propriétés, s'appuyant sur des considérations

physiques, météorologiques, géologiques ou chimiques. Le sujet d'intérêt fondamental de l'hydrologie est le cycle de l'eau, illustré à la Figure 1.1 dans sa partie continentale (les chiffres indiquent les flux moyens annuels en notant 100 le volume annuel total des précipitations sur les terres émergées, c'est-à-dire 119.000 km^3 , soit une lame d'eau annuelle d'environ 800 mm sur les $149.400.000 \text{ km}^2$ de terres émergées). Cet immense transfert d'eau naît des variations spatiales et temporelles des flux journaliers d'énergie solaire, des hétérogénéités de la surface du globe et de la différence de mobilité de l'eau, qu'elle soit sous forme de vapeur dans l'atmosphère, liquide à la surface de la Terre ou dans le sous-sol, ou sous forme solide dans les neiges et les glaces.

Dooge (1988) mentionne que 'l'affaire de l'hydrologie est de résoudre l'équation du bilan de l'eau'. L'hydrologie continentale, qui s'intéresse plus particulièrement à la partie du cycle de l'eau sur ou proche des terres émergées, peut être aussi définie comme la science de l'eau qui traite de la circulation, de la distribution, de la dynamique et des propriétés de l'eau sur Terre au travers du cycle hydrologique (Eagleson, 1991). Ses thèmes d'étude sont les précipitations, l'évaporation, l'infiltration, le ruissellement, les écoulements dans les nappes et les cours d'eau, et le transport de substances dissoutes ou en suspension.

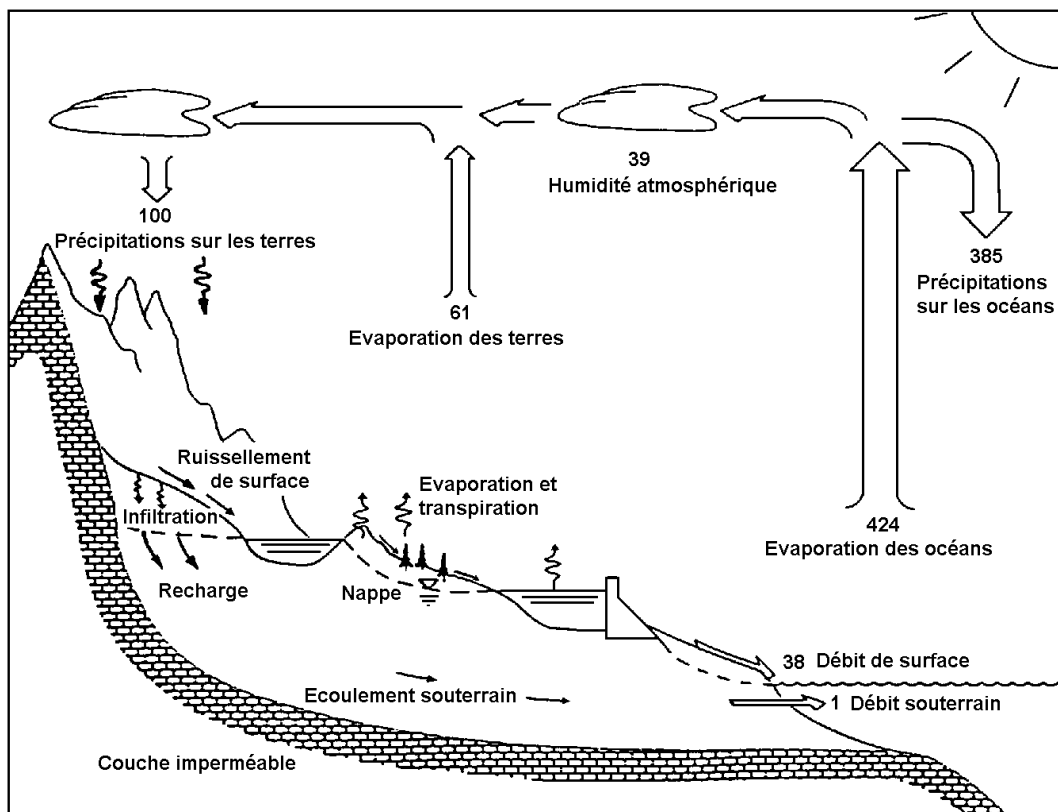


Figure 1.1: Le cycle hydrologique, avec les flux moyens annuels en pourcentage du volume annuel total des précipitations sur les terres émergées (d'après Maidment, 1992)

Du fait de l'étendue de ces centres d'intérêt, l'hydrologie est donc une science pluridisciplinaire comprenant l'hydrologie de surface, la glaciologie, l'hydrogéologie, la nivologie, la physico-chimie, en incluant aussi l'étude de l'érosion ou du transport de sédiments. Mais elle est également plus ou moins directement liée à la météorologie, l'hydraulique, la géographie, la géologie, la biologie ou l'écologie. Il est donc très difficile d'assigner à l'hydrologie un but plus précis que la lourde tâche de décrire et comprendre le cycle de l'eau, dans un environnement éminemment complexe, hétérogène et variable dans le temps. Klemeš (1988) illustre d'ailleurs cette complexité en disant que pour le scientifique,

résoudre l'équation du bilan de l'eau peut être considéré comme l'un des Rubic Cubes les plus difficiles à résoudre de la nature, pour lequel les facettes changent de couleur, de forme, de taille au fur et à mesure qu'elles sont déplacées par différentes forces, et dans lequel même les bases structurales changent au cours du temps.

Bien que l'hydrologie vienne d'être présentée comme une science, cette désignation n'est pas la seule acceptée. En plus de la recherche fondamentale ou appliquée, l'hydrologie inclut aussi l'ingénierie hydrologique ou l'hydrologie opérationnelle. Pendant les années 80, certains hydrologues se sont d'ailleurs interrogés sur la nature de l'hydrologie en tant que science. Klemeš (1986a), par exemple, a alarmé la communauté scientifique sur ce qu'il a appelé du 'dilettantisme en hydrologie', mettant en garde les hydrologues sur la dangereuse tendance que suivait l'hydrologie à devenir seulement une technologie mal appliquée.

Le moteur de cette science réside principalement dans le lien étroit qui existe entre l'homme et son environnement, et en particulier dans la dépendance qu'il a vis-à-vis de la ressource en eau, pour ses besoins d'alimentation ou dans ses activités. La variabilité de cette ressource dans le temps est source de multiples problèmes de gestion, auxquels l'hydrologie peut contribuer à apporter des solutions (Michel, 1989).

1.2.2. Qu'est qu'un modèle ?

Parce qu'une très grande complexité caractérise les processus impliqués dans le cycle de l'eau, les hydrologues conviennent aujourd'hui de leurs connaissances partielles. En effet, la représentation de tout ou partie de ce cycle (démarche de modélisation), à des fins variées, se heurte à la difficulté d'appréhension, de description ou de compréhension des phénomènes et des systèmes étudiés. Par conséquent, les représentations de ces systèmes, c'est-à-dire les modèles, sont nécessairement simplificatrices, réductrices de la complexité naturelle, et donc grossièrement inexactes. Le modélisateur opère des choix de représentation du système en ne retenant que les aspects qui lui semblent les plus pertinents, pour tendre vers une solution la moins inexacte possible.

Le développement du modèle repose généralement sur trois éléments:

1. le système observé et sa discrétisation spatiale et temporelle, qui en définissent l'objet et ses limites (spatiales ou temporelles). La connaissance du système est conditionnée par la mesure de ses caractéristiques et l'acquisition de données sur les flux, les stocks et les transformations de phase. Dans notre cas, le système est typiquement le bassin versant défini plus loin;
2. l'objectif de modélisation, pour lequel le modèle est développé. Le modèle est construit pour répondre à des questions et peut ainsi, au delà d'un outil de représentation, être également un instrument de connaissance. L'étude du cycle de l'eau est l'objectif général de l'hydrologie. Nous nous intéresserons plus particulièrement dans ce travail de recherche à la représentation de la transformation de la pluie en débit;
3. le choix d'une formulation de la réalité. Après la définition des deux points précédents, l'essentiel de la démarche de modélisation consiste à trouver la formulation de la réalité la plus satisfaisante relativement aux objectifs fixés. En fait, le modèle est la plupart du temps le résultat d'un compromis entre généralité, réalisme et précision (Kauark-Leite et Nascimento, 1993). La formulation d'un modèle est conditionnée par la connaissance antérieure des processus ou des systèmes considérés, par les idées et l'imagination du modélisateur et par les hypothèses qui sont formulées. Notre travail ne porte que sur des modèles mathématiques, c'est-à-dire qui proposent une mise en équations du système observé. Les modèles réduits du système naturel ne sont guère utilisés en hydrologie car

ils sont souvent difficiles à construire, trop coûteux et laissent beaucoup moins de possibilités d'utilisation.

S'il est souvent possible de construire un très grand nombre de modèles sur une réalité particulière, les points 1 et 2 fournissent des balises généralement capables de restreindre les choix du modélisateur (Thom, 1979, cité par Nascimento, 1995). Le travail de vérification de ces modèles, permettant notamment de conforter les hypothèses avancées, est indispensable après la phase de construction pour assurer une certaine crédibilité au modèle, comme nous le verrons par la suite.

1.2.3. Objet et enjeux de la modélisation pluie-débit

L'étude de la transformation de la pluie en débit est une des nombreuses disciplines hydrologiques, à laquelle nous nous intéressons plus particulièrement dans ce travail. Elle est assez naturelle puisqu'elle essaie de trouver un lien entre les débits et les phénomènes qui en sont la cause directe, les pluies (le traitement de la neige ne sera pas abordé dans notre étude). Nous nous proposons ici de définir brièvement l'objet d'étude et les objectifs de la construction de modèles pluie-débit.

Excepté pour les modèles de calcul des flux d'eau à une échelle continentale souvent couplés à des modèles climatiques de circulation générale, l'objet d'étude des modèles pluie-débit est le bassin versant. Celui-ci est défini relativement à un point sur un cours d'eau, comme l'ensemble des terres drainées par le réseau hydrographique situé en amont de ce point. Ainsi, une goutte de pluie tombant sur cette surface finira par passer au point considéré de la rivière, à moins qu'elle ne soit évaporée ou qu'elle ne s'infilte vers des couches trop profondes. Lorsque l'on parcourt une rivière de l'amont vers l'aval, la surface du bassin croît progressivement. Le bassin est un système ouvert, avec des échanges d'eau et d'énergie avec l'atmosphère, le sous-sol et l'aval du cours d'eau ou la mer.

La définition des contours du bassin est essentiellement topographique. Elle peut donc être parfois difficile dans les zones au relief peu contrasté. Par ailleurs, ces limites ne correspondent pas toujours aux contours définis par la structure géologique sous-jacente. Ainsi des terres situées en dehors des limites topographiques peuvent être drainées par des couches géologiques qui alimentent le bassin. La définition du bassin versant n'est donc pas toujours facile et peut être source d'incertitude dans la démarche de modélisation. A titre d'exemple, les formations karstiques induisent souvent de très fortes différences entre les limites topographique et géologique.

Par ailleurs, il faut souligner que le bassin versant est un système vivant, composition complexe hautement hétérogène et en constante évolution. A ce titre, on pourrait le comparer à un corps humain par exemple: chaque partie a un lien avec les autres, et les réactions à des sollicitations extérieures sont conditionnées par un grand nombre de paramètres et de variables d'état du système (bien entendu, les interactions, rétroactions ou régulations qui existent dans un organisme vivant sont plus fortes que dans un bassin versant). Son évolution au cours du temps dépend à la fois de ses caractéristiques intrinsèques mais également de son environnement qui le conditionne en partie. De cette image, il faut certainement retenir que le bassin est avant tout une entité. Son analyse fine est éminemment délicate mais source d'une meilleure connaissance. Son analyse globale permet d'en tirer des traits et caractéristiques tout à fait pertinents pour en comprendre le fonctionnement d'ensemble.

Pourquoi s'intéresser à la manière dont l'eau de pluie rejoint les cours d'eau ? Les enjeux de ce questionnement sont nombreux. Nous pouvons d'abord identifier les objectifs de

connaissance interne du système. Il s'agit d'arriver à comprendre, par une approche dynamique, le fonctionnement du bassin, la nature et le rôle des processus mis en jeu, leur importance relative par rapport au phénomène étudié ou encore les liens spatiaux ou temporels entre ces processus. Ces études permettent de renforcer, d'infirmer ou de proposer des concepts ou théories sur les écoulements de l'eau en milieu naturel. Les modèles correspondants constituent alors une formalisation des connaissances. Ils contribuent à l'accroissement de ces dernières, notamment au travers de scénarii de simulation.

C'est par cet apport à la connaissance de notre environnement, de son fonctionnement, du comportement du bassin versant que la modélisation pluie-débit peut réussir à répondre à de nombreuses questions centrées sur l'eau, gestion des risques et de la ressource notamment. Michel (1989) et Refsgaard et Abbott (1996) répertorient l'essentiel de ces problématiques. Si des aspects de qualité des eaux peuvent être étroitement liés à des aspects de quantité, nous nous intéresserons ici uniquement à une modélisation pluie-débit quantitative. Nous donnons quelques-unes de ces problématiques:

- Simulation de débits, pour le comblement de lacunes dans des séries de données, la reconstitution de débits historiques (les données de pluie étant souvent disponibles sur des périodes beaucoup plus longues que les débits) ou pour permettre des traitements statistiques;
- Prédétermination des débits de crue ou d'étiage: on désire savoir avec quelle fréquence des débits de crue (supérieurs à un seuil de risque par exemple) ou des faibles débits (en deçà d'un débit réservé par exemple) risquent de se produire, et sur quelle durée. On se place ici dans une démarche d'analyse fréquentielle. Cette connaissance peut permettre le dimensionnement d'ouvrages et de réservoirs ou d'aménagements dans le lit (mineur à majeur) du cours d'eau;
- Préviation des crues et des étiages: il s'agit d'évaluer par avance (avec un délai de quelques heures à quelques jours), connaissant l'état du bassin, les débits de crues susceptibles de présenter des risques (inondation) ou les débits d'étiages pouvant demander de mettre en place une gestion particulière de la ressource (par des barrages-réservoirs par exemple) pour assurer l'approvisionnement en eau ou la préservation de la vie halieutique. On s'inscrit ici dans une démarche d'analyse en continu du bassin;
- Influence d'aménagements sur l'hydrologie: on désire pouvoir prédire les changements de la réponse du bassin suite à des modifications des caractéristiques du bassin d'origine humaine ou à des changements environnementaux.

Ces problématiques font ressortir deux aspects importants, celui de l'évaluation du risque et celui de la gestion de la ressource. La pertinence des réponses que l'on peut leur apporter est conditionnée par celle du modèle dans sa représentation du bassin relativement aux objectifs fixés. Si d'autres approches hydrologiques que la modélisation pluie-débit proposent des réponses à certaines de ces problématiques (analyse fréquentielle statistique sur les débits pour la prédétermination), cette dernière paraît cependant très profitable pour des questions supposant un traitement temporel continu, comme la préviation des débits. Dans ce cas, le fait de remonter à l'origine des débits (la pluie) permet en plus de profiter d'un délai supplémentaire par rapport à des méthodes n'exploitant que l'information sur les débits.

De nombreux auteurs ont étudié ces problèmes en utilisant des modèles pluie-débit. A titre d'exemple, nous pouvons citer les travaux de Cameron *et al.* (1999) et Uhlenbrook *et al.* (1999) sur la prédétermination des crues, ceux de Xu et Vandewiele (1995) et Yang *et al.* (1995) respectivement sur le dimensionnement et la gestion de réservoirs, ou ceux de Da Ros et Borga (1997) et Yang et Michel (2000) sur la préviation des débits. Les modèles pluie-débit,

quels qu'ils soient, ne sont pas encore capables de prédire les effets de changements environnementaux ou d'origine humaine sur le bassin. En revanche, ils sont aujourd'hui capables de détecter a posteriori les effets de tels changements (voir par exemple Lavabre *et al.*, 1993; Nascimento, 1995; Brandt *et al.*, 1988 ou Lørup *et al.*, 1998). Des travaux sont d'ailleurs en cours par Andréassian (thèse en cours, 2001) pour étudier l'effet de l'évolution du couvert forestier sur le comportement hydrologique des bassins.

1.2.4. Une classification des modèles pluie-débit

Si la modélisation pluie-débit représente un étroit domaine de l'hydrologie, elle n'en est pas moins féconde en terme de production de modèles. Le développement des moyens de calculs informatiques a certainement été l'un des facteurs majeurs ayant favorisé l'essor de ce secteur de recherche depuis le début des années 60, avec la création de dizaines de modèles. Encore aujourd'hui, de nouveaux modèles sont très souvent proposés dans la littérature. Ce foisonnement tient d'une part de la diversité des approches possibles, qui font appel à des concepts et des points de vue différents sur la manière de représenter la réponse d'un bassin versant à des événements de pluie. Il tient certainement d'autre part d'un certain degré d'ignorance sur la meilleure façon de modéliser la relation pluie-débit. Devant la complexité et la diversité des systèmes observés, il est probable cependant que la solution miracle n'existe pas.

Beaucoup de classifications des modèles ont été proposées et il y a presque autant de classifications que d'hydrologues ! La difficulté de trouver une classification unifiée provient du fait que la grande diversité des approches entraîne une non moins grande diversité des caractéristiques des modèles. Il est alors presque impossible de distinguer des catégories de modèles nettement distinctes, et ce d'autant plus que la terminologie employée est encore fluctuante. Clarke (1973) et Ambroise (1998) ont proposé quelques clés de différenciation des modèles:

- déterministes ou stochastiques, suivant la nature des variables, des paramètres et/ou des relations entre eux,
- globaux, semi-distribués ou spatialisés, suivant que le bassin versant est considéré dans l'espace comme une entité homogène, qu'il est divisé en sous-unités supposées homogènes (éventuellement des sous-bassins) ou qu'il est finement discrétisé en mailles,
- cinématiques (descriptifs) ou dynamiques (explicatifs), suivant que l'évolution temporelle du système est simplement décrite ou mise en relation avec les forces qui en sont la cause,
- empiriques, conceptuels ou théoriques (fondés sur la physique), suivant les relations utilisées pour modéliser la transformation de la pluie en débit et suivant la représentation du système modélisé.

On pourrait rajouter à ces éléments le type de fonctionnement du modèle au cours du temps, continu ou par événement (Linsley, 1982), le type et la quantité de données requises (Todini, 1988), le but du modèle (Roche, 1988) ou le pas de discrétisation temporel concerné (heure, jour, mois, année). Les classifications proposées emploient souvent des combinaisons de ces caractéristiques. La dernière catégorie de signes distinctifs est probablement la plus ambiguë, car les limites entre modèles empiriques, conceptuels ou théoriques sont très floues. Cependant, ces termes réfèrent en partie à l'approche de développement du modèle et nous nous y référons dans la suite. La classification que nous donnons ici est grossière (et insatisfaisante comme les autres classifications), mais elle a été choisie pour isoler et mieux situer les modèles auxquels nous allons nous intéresser plus particulièrement dans notre étude.

Nous distinguons trois catégories de modèles, les modèles ‘boîte noire’, les modèles ‘à réservoirs’ conceptuels ou empiriques et les modèles fondés sur la physique. Dans chaque cas, nous donnons quelques exemples pour mieux comprendre la nature des modèles désignés.

◆ les modèles ‘boîte noire’

Nous classons dans cette catégorie les modèles ou sous-modèles qui établissent un pur lien mathématique entre les variables d’entrée et de sortie du système. Ces modèles sont globaux. Nous en présentons succinctement quelques exemples.

Les modèles ARMAX (Auto-Regressive Moving Average with eXogenous inputs), autorégressifs à moyenne mobile, sont parmi les plus simples. Initialement développés par Box et Jenkins (1976), ils ont suscité beaucoup d’intérêt, principalement du fait de leur solide fondement mathématique, qui s’appuie sur la théorie des modèles linéaires (modèles de régression et de corrélation). La forme générale du modèle est donnée par l’équation (Salas, 1993):

$$y_t = \mu + \sum_{j=1}^p \phi_j (y_{t-j} - \mu) + \sum_{j=1}^r \psi_j z_{t-j} + \varepsilon_t - \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad \text{Eq. (1.1)}$$

où μ et ψ_1, \dots, ψ_r sont des paramètres du modèle, y_t et z_t respectivement la sortie (débit) et l’entrée exogène (pluie) à l’instant t , ϕ_1, \dots, ϕ_p sont des p paramètres autorégressifs, $\theta_1, \dots, \theta_q$, sont q paramètres de moyenne mobile et le bruit ε_t est un processus normal non autocorrélé, de moyenne nulle et non corrélé à y_{t-1}, \dots, y_{t-p} . Bien qu’ils aient donné des résultats satisfaisants dans un certain nombre de cas (voir par exemple Szöllosi-Nagy, 1976), ces modèles trouvent leurs limites dans le fait que leur linéarité n’est pas adaptée pour représenter la transformation pluie-débit non linéaire. Ils doivent donc être utilisés en combinaison avec des opérateurs de transformation non linéaires (Weeks et Boughton, 1987).

Contrairement aux modèles ARMAX, les réseaux de neurones artificiels sont non linéaires et ne rencontrent donc pas les problèmes mentionnés précédemment. Peut-être du fait de leur lien séduisant à l’intelligence artificielle, ils ont été utilisés par de très nombreux auteurs depuis le milieu des années 90 dans le contexte de la modélisation pluie-débit (Hsu *et al.*, 1995; Lek *et al.*, 1996; Shamseldin, 1997; Dawson et Wilby, 1998; Zealand *et al.*, 1999; Coulibaly *et al.*, 2000). Les réseaux de neurones ont une structure mathématique flexible, capable d’identifier des relations non linéaires complexes entre les entrées et les sorties d’un système. Ils imitent, de façon simplifiée, la manière dont les neurones d’un cerveau se comportent. Ils sont organisés en couches successives de nœuds (neurones) et chaque nœud d’une couche est relié à tous les nœuds de la couche suivante par une relation paramétrée. On comprend alors aisément que le nombre de paramètres augmente très rapidement lorsque le nombre de nœuds augmente, rendant le système sur-paramétré. Hsu *et al.* (1995) ont par exemple appliqué des réseaux de neurones comprenant de 33 à 64 paramètres. Par ailleurs, l’utilisation du réseau de neurones suppose que les données soient normées, ce qui pose des problèmes lorsque l’on veut appliquer le modèle sur des données sortant de l’intervalle de variation de celles utilisées pour le calage des paramètres.

Nous avons également regroupé sous l’appellation ‘boîte noire’ des composants de modèle utilisés notamment dans la fonction de transfert. C’est le cas par exemple de l’hydrogramme unitaire (HU), qui a été proposé pour la première fois par Sherman (1932). L’HU tente de représenter l’effet des caractéristiques du bassin sur une entrée unitaire de pluie nette. Simple dans sa conception (il découle de l’interprétation physique des courbes isochrones) et basé comme les modèles ARMAX sur les mathématiques linéaires, il est confronté aux mêmes limitations. Essentiellement outil de transfert, il doit généralement être couplé à des modules

non linéaires qui permettent de convertir la pluie brute en pluie nette (voir Jakeman et Whitehead, 1996), c'est-à-dire de déterminer la proportion de la pluie en entrée qui contribuera finalement à l'écoulement. Par ailleurs, l'application de l'HU demande la séparation préalable du débit de base, dont la détermination est souvent arbitraire.

La méthode DPFT-ERHUDIT (Différentielle Première de la Fonction de Transfert - Excess Rainfall and Unit Hydrograph by a Deconvolution Iterative Identification Technique), inspirée des travaux de Newton et Vinyard (1967), a pour but d'essayer de limiter les problèmes liés à l'utilisation des HU, en évitant un choix arbitraire dans la sélection et le calage d'une fonction de production et d'une technique de séparation du débit de base (Rodriguez *et al.*, 1989, 1991; Duband *et al.*, 1993). Il s'agit d'une procédure itérative, dans laquelle la première estimation de la pluie nette correspond à la pluie brute. La première fonction de transfert est alors calculée, puis rectifiée par des contraintes physiques. A partir de cette nouvelle fonction de transfert, on résout alors le problème inverse, c'est-à-dire l'estimation de la pluie nette par déconvolution. La procédure est répétée jusqu'à ce qu'il y ait convergence. Ce modèle, dont le calage repose sur la sélection d'événements de crues, a donné des résultats satisfaisants dans un contexte de prévision de ces événements. Cependant, il doit être considéré comme un modèle de crue puisqu'il est calé sur des événements où le système hydrologique est en état de répondre par une crue à une entrée de pluie. Sa faiblesse dans l'extension à d'autres périodes réside dans l'absence de procédure de suivi d'humidité et Duband *et al.* (1993) ont remarqué qu'il subsistait des ambiguïtés sur la question du débit de base. D'autres auteurs tels que Chapman (1996a, 1996b) ont proposé des approches très similaires.

Enfin, dans cette classe de modèles 'boîte noire', nous citerons ceux reposant sur une simple équation non linéaire qui donne directement le débit en fonction des pluies antérieures. Un exemple est l'équation de Tsykin (1985) que nous verrons plus en détail dans la suite. Nous pouvons également citer la méthodologie développée par Pinault *et al.* (1997) qui s'appuie sur des techniques de traitement du signal.

La plupart de ces approches reposent sur un solide environnement mathématique. Le caractère linéaire de certaines d'entre elles en limite l'utilisation dans le contexte de la modélisation pluie-débit. Les méthodes par événements sont essentiellement destinées à s'intéresser à des périodes de crue et donc peu ou pas adaptées à un suivi en continu du bassin auquel nous nous intéressons ici. Les réseaux de neurones sont des méthodes largement sur-paramétrées et leur utilisation reste délicate mais connaît de constants développements.

◆ Les modèles 'à réservoirs', conceptuels ou empiriques

Les modèles de cette classe diffèrent de ceux du type 'boîte noire' en ce qu'ils décomposent la transformation de la pluie en débit en sous-processus qui seront traités de façon globale (contrairement aux modèles fondés sur la physique par référence aux lois acceptées de la physique, comme nous le verrons dans la suite). Ces procédures prennent souvent la forme de réservoirs qui se remplissent et se vident au cours du temps. Elles permettent donc au système d'avoir une mémoire des conditions antérieures, qui ne se limitent pas comme dans les modèles ARMAX aux débits et pluies antérieures. Ces modèles sont schématiquement constitués de réservoirs interconnectés qui assurent la transformation de la pluie en débit. Ils sont donc moins abstraits que les modèles 'boîte noire' dans leur représentation de la transformation pluie-débit, même si leur interprétation physique n'est généralement pas immédiate. Ces modèles sont structurés à partir de deux composants majeurs:

- un module de production responsable de la détermination des bilans en eau, c'est-à-dire de la répartition de la pluie brute en pluie nette (alimentant le débit), en quantité évaporée et

en quantité stockée. Parfois présentes dans les modèles, les fonctions d'échanges en eau, permettant de simuler des pertes et/ou des apports vers ou de l'extérieur, font également partie de ce module;

- un module de routage ou transfert permettant d'assurer la répartition temporelle de la quantité d'eau transitant dans le cours d'eau.

Lorsque le pas de temps d'étude augmente, l'importance du transfert diminue et le modèle tend à se réduire à une fonction de production destinée à gérer les bilans en eau. Lorsque le pas de temps diminue, le transfert a tendance à être de plus en plus sophistiqué. Production et routage ne sont pas toujours clairement dissociables dans le modèle: un réservoir peut jouer un rôle pour chacune de ces fonctions. Ces dernières dépendent de paramètres dont les valeurs doivent être déterminées sur chaque bassin. Rarement directement mesurables, la plupart des paramètres doivent être optimisés par une procédure de calage numérique.

Les premiers modèles à réservoirs ont été proposés au début des années 60. Ils correspondaient en fait à une simplification des équations de la physique, alors trop complexes pour être utilisées telles quelles dans des modèles numériques. Ils essayaient de représenter de façon assez exhaustive tous les processus pouvant influencer les débits et étaient donc très complexes, avec des structures faisant intervenir de nombreux réservoirs et de très nombreux paramètres: le modèle de Stanford (Crawford et Linsley, 1963) comprenait plus de 30 paramètres optimisables, le modèle SSARR (Schermerhorn et Kuehl, 1968) plus de 25 paramètres, celui proposé par Girard (1970) plus de 20 paramètres. Devant les problèmes de calage et d'utilisation de tels modèles, ce sont ensuite des modèles moins complexes (moins de paramètres) qui ont été proposés. Nous les étudions plus en détail par la suite.

Ces modèles sont la plupart du temps globaux, ou dans quelques cas semi-distribués ou distribués. Dans ces deux dernières situations, le modèle est appliqué sur chaque sous-bassin ou maille et une méthode de propagation permet d'assurer le transfert des flux entre sous-unités vers l'exutoire.

A titre d'exemple, nous pouvons citer le cas du TOPMODEL (TOPography-based hydrological MODEL) qui est le modèle autour duquel le plus de travaux sont réalisés actuellement. Développé à l'Université de Lancaster au Royaume-Uni, ce modèle a été construit sur le concept d'aires contributives variables à l'écoulement (Beven et Kirkby, 1979). Prenant explicitement en compte la topographie du bassin étudié, il utilise le fait que les sols des zones de plus faible pente (fond de vallée) ont une plus grande propension à atteindre un état de saturation que les zones de forte pente, et donc à générer des écoulements rapides lors d'événements pluvieux. Une représentation schématique du modèle est donnée à la Figure 1.2.

TOPMODEL est un modèle que l'on peut qualifier de conceptuel. La différence entre les deux adjectifs conceptuel et empirique que nous avons utilisés pour qualifier les modèles à réservoirs fait ici exclusivement référence à la manière dont les modèles ont été développés:

- la grande majorité des modèles à réservoirs existants sont des modèles conceptuels mis au point à partir de concepts sur la manière dont se passe la transformation pluie-débit. On représente alors dans le modèle de façon simplifiée les processus jugés pertinents pour la modélisation. Le concepteur du modèle introduit des idées a priori, généralement fondées sur les connaissances physiques et l'observation du bassin. Beven (1987) qualifie d'ailleurs cette approche de 'perceptuelle', soulignant qu'elle reste personnelle et dépendante de la façon dont le modélisateur perçoit la réalité;
- beaucoup plus marginaux que les précédents, les modèles empiriques, eux, sont construits seulement à partir des observations hydro-météorologiques (correspondant aux entrées et

sorties du modèle), sans avancer d'hypothèses a priori sur la nature des processus dominants ou d'idées préconçues sur la structuration du modèle. Le modèle est alors construit à partir d'une structure élémentaire, complexifiée progressivement si nécessaire pour une amélioration de l'efficacité. Le modélisateur cherche en fait à découvrir quels éléments doivent intervenir dans cette structure et comment ils sont reliés, non pas en fonction de la chaîne de phénomènes dont il a connaissance dans le milieu naturel, mais en fonction du plus ou moins grand succès du modèle à reproduire les débits observés.

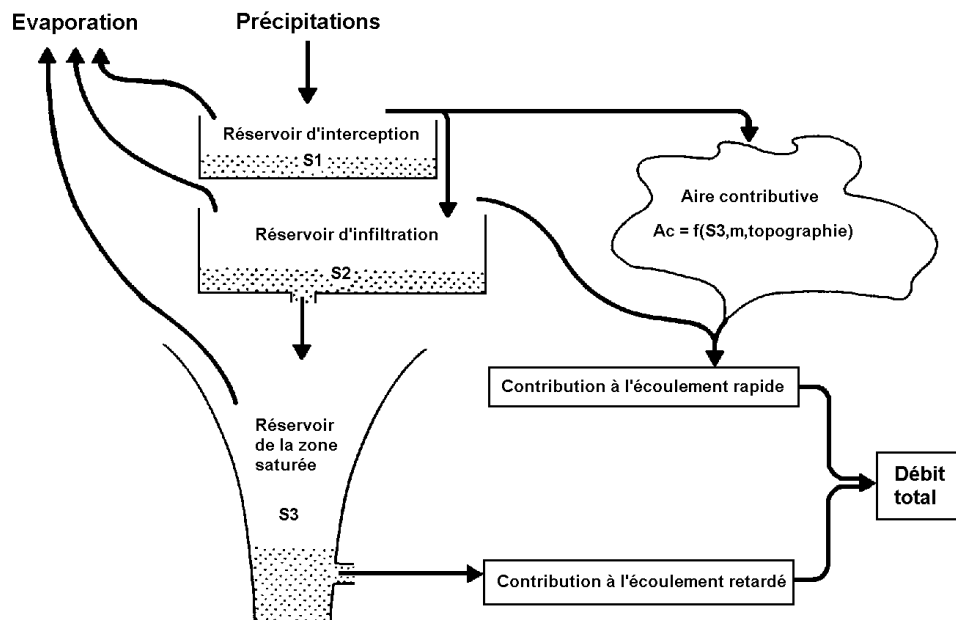


Figure 1.2: Représentation schématique de TOPMODEL (S1, S2, S3: niveaux dans les réservoirs; m: un des paramètres du modèle) (d'après Beven et Kirkby, 1979)

Wheater *et al.* (1993) discutent du problème de la représentation de la complexité du bassin versant et des processus de génération des débits dans les modèles à réservoirs. Ils soulignent notamment que la représentation opérée dans certains modèles conceptuels, en privilégiant a priori un type de processus dominant, pourrait s'avérer trop simpliste à l'échelle du bassin versant. Ils citent notamment des travaux qui ont montré qu'un très grand nombre de processus sont généralement observables sur le bassin. Ambroise (1998) identifie notamment comme processus possibles de génération des débits, le ruissellement de surface (par dépassement de l'infiltrabilité ou par saturation du sol), l'écoulement hypodermique en zone saturée ou non saturée, l'écoulement en macropores ou le drainage de la nappe. La réponse globale en est une agrégation complexe, très différente d'une simple juxtaposition telle qu'elle est réalisée dans certains modèles (modèle de Stanford par exemple). Wheeler *et al.* (1993) précisent que la détermination a priori des processus dominants est quasiment inenvisageable sans campagne de terrain très détaillée. Ils concluent qu'il existe un problème fondamental d'observabilité des systèmes hydrologiques et qu'il est aujourd'hui impossible d'observer la 'vérité' hydrologique à l'échelle du bassin.

On peut déduire de ces réflexions que, d'un bassin à l'autre, il faut probablement considérer que c'est l'importance relative des différents processus de génération des débits qui change, mais que beaucoup restent présents. Ainsi, la construction d'un modèle sur la représentation d'un processus particulier rend le modèle très spécifique aux bassins où la prédominance de ce processus est avérée. Du fait de la difficulté d'identification de cette prédominance des processus sur le terrain, l'utilisation du modèle peut alors devenir peu fiable dans le cas d'un bassin choisi au hasard. Un modèle pouvant simuler une large gamme de conditions sur des

bassins très variés, répondant ainsi à une qualité de ‘généralité’ prônée par Linsley (1982), paraît plus souhaitable.

Nascimento et Michel (1992) discutent en détail les différences entre les approches de modélisation conceptuelle et empirique. Leurs réflexions vont dans le sens du point de vue exposé précédemment. Il ressort notamment que les hypothèses avancées dans la représentation des sous-processus dans les modèles conceptuels sont souvent difficiles à valider avec les données disponibles. Ces dernières sont souvent en trop faibles quantités pour justifier la description détaillée des sous-processus. Comme elles reposent sur une perception de la réalité du modélisateur, les structures des modèles conceptuels incluent une certaine part de subjectivité, avec une tendance à favoriser a priori certains aspects de la représentation au détriment d’autres moins bien connus ou moins visibles. Ainsi le rôle de la perception du bassin et des processus dominants par le modélisateur reste un moteur essentiel de la démarche conceptuelle.

Dans l’approche empirique, la perception du modélisateur du monde réel est pratiquement absente. Il y a de ce fait, pour les modèles empiriques, une interaction relativement faible avec l’expérimentation et ce, d’autant plus que leur développement repose essentiellement sur des séries hydrologiques nombreuses. Ils sont peu propices à l’augmentation des connaissances sur les processus détaillés. En revanche, ils sont des outils pertinents d’analyse comportementale des bassins. L’absence d’idées préconçues dans leur construction, c’est-à-dire l’absence d’un cadre rigide de développement et d’hypothèses contraignantes, pourrait leur permettre une évolution plus rapide vers des réponses efficaces à des problématiques données. Leur développement reste cependant très dépendant de la disponibilité des données.

◆ Modèles fondés sur la physique

Contrairement aux modèles précédents qui mettent l’accent sur la représentation du comportement hydrologique final (à l’exutoire), les modèles fondés sur la physique tentent d’utiliser des explications physiques à ce comportement. Ils utilisent le cadre théorique des équations de la physique (équations aux dérivées partielles), avec par exemple les équations de Saint-Venant pour les écoulements en rivière ou celles de Boussinesq ou Richards pour les écoulements dans les sols saturés ou non saturés. Ils tiennent compte de la variabilité spatiale du bassin par une discrétisation fine à base de mailles sur lesquelles sont appliquées les lois concernées, ce qui permet de donner une représentation des flux et stocks au sein du bassin. Chacune est caractérisée par un ensemble de paramètres, ayant en principe une signification physique et donc mesurables sur le terrain.

Le niveau de raffinement de ces modèles devrait leur permettre de pouvoir simuler simultanément d’autres variables hydrologiques que le débit (niveau des nappes, évaporation, etc.). Théoriquement, ces modèles, par leur approche distribuée et physique, se mettent en position de pouvoir répondre à des questions d’effet de changement d’occupation du sol et de pouvoir être appliqués sur des bassins non jaugés. La valeur de ces atouts potentiels a contribué à l’engouement qu’a connu la modélisation fondée sur la physique dans les années 80, avec la perspective de répondre à des questions devant lesquelles les autres approches de modélisation citées précédemment étaient sans réponses satisfaisantes. Ainsi on a assisté au développement des modèles SHE (Abbott *et al.*, 1986), IHDM (Beven *et al.*, 1987), SWATC (Morel-Seytoux et Al Hassoun, 1989) ou plus récemment du modèle proposé par Kuchment *et al.* (1996).

Cependant, les discussions passionnées qu’ont suscité le développement et l’application de ces modèles (Klemeš, 1982; Dooge, 1986; Beven, 1989; Grayson *et al.*, 1992) tendent à indiquer que leurs avantages théoriques sur des modèles plus simples sont en pratique loin

d'être vérifiés, voire inexistantes. Du fait de la complexité des équations, chaque maille dépend d'une grande quantité de paramètres, dont la mesure est, économiquement et techniquement, souvent impossible à l'échelle de discrétisation. Par ailleurs, il est difficile de savoir dans quelle mesure l'hétérogénéité à l'échelle de la maille peut être représentée par des propriétés effectives d'un élément uniforme (Wheater *et al.*, 1993). Les modèles doivent alors recourir à un calage, qui est toujours très aléatoire étant donné le nombre de degrés de liberté. Les valeurs effectives des paramètres à l'échelle de la maille ne correspondent alors plus par conséquent à aucune réalité physique du terrain. Les modèles perdent de ce fait leur applicabilité théorique à des bassins non jaugés et la possibilité de prédire les effets de changements d'occupation du sol.

Par ailleurs, la validation du fonctionnement interne complexe de ces modèles est quasiment impossible, notamment du fait de l'absence de données suffisantes: seules les données de débit sont généralement disponibles, rarement celles correspondant à d'autres variables internes du modèle. Les paramètres nécessaires à la caractérisation des sols et des couches sous-jacentes sont souvent difficiles à acquérir, et les conditions aux limites dans ces zones ne sont connues qu'exceptionnellement (Bronstert *et al.*, 1998). D'un point de vue pratique, l'utilisation de tels modèles reste du champ de la recherche, avec des temps de calcul trop longs et des besoins en données trop importants pour pouvoir envisager, pour l'instant, une application opérationnelle.

Cette présentation des différents types de modèles nous a permis de mieux souligner les différences de points de vue dans les démarches de modélisation, et la façon dont se traduit cette différence dans la formalisation du modèle. Pour évaluer l'intérêt de ces approches variées, il faut d'abord définir les qualités recherchées pour ces modèles. Lindström *et al.* (1997) présentent les qualités suivantes comme règles de modélisation:

- le modèle doit avoir des fondements scientifiques sains,
- les besoins en données doivent être en accord avec la disponibilité de celles-ci sur la majorité des bassins,
- la complexité du modèle doit être justifiée en termes de performances,
- le modèle doit être convenablement validé,
- le modèle doit être compréhensible par les utilisateurs.

On peut ajouter à cela que le modèle doit apporter des réponses satisfaisantes aux questions posées et doit présenter des qualités de généralité et de robustesse. Si certaines de ces qualités sont plutôt qualitatives, d'autres en revanche peuvent être mesurées quantitativement et objectivement. Un moyen d'évaluer le degré de satisfaction que ces modèles apportent dans leurs réponses aux questions posées est alors de conduire des études comparatives. Si les comparaisons sont rarement le point central de travaux de recherche, elles demeurent cependant des sources de réflexion très intéressantes sur la crédibilité et la validité des différentes approches. Dans le paragraphe suivant, nous présentons une analyse des comparaisons de modèles pluie-débit déjà réalisées.

1.2.5. Etudes comparatives: une évaluation objective des modèles ?

Nous avons vu dans la présentation précédente que les modèles 'boîte noire' ou les modèles fondés sur la physique présentent des caractéristiques pouvant limiter leur fiabilité et/ou leur utilisation pratique pour un suivi en continu du bassin versant. Si ces remarques s'appuient sur une analyse structurelle, le jugement des modèles se doit, pour être complet, de comporter

l'analyse de leurs capacités à fournir des réponses adéquates à la simulation de la transformation pluie-débit. La multitude des modèles existants a donné lieu à beaucoup d'évaluations individuelles, et relativement peu d'analyses comparatives. Or ce sont bien ces dernières qui permettent de juger de la valeur d'un modèle. L'évaluation individuelle, dans l'absolu, sans aucune autre référence que la valeur indicative d'un critère de qualité, est peu révélatrice de la confiance que l'on peut accorder au modèle: un modèle satisfaisant à 80 % peut paraître bon, mais il devient peu intéressant s'il en existe un autre satisfaisant à 90 %. Cette évaluation comparative est d'autant plus importante lorsque l'on se place dans un contexte opérationnel, où l'utilisateur est désireux d'appliquer le modèle qui donnera les réponses les plus fiables à son problème, avec des enjeux économiques, sociaux ou environnementaux.

C'est donc dans le but d'essayer d'apporter des conseils aux utilisateurs qu'ont été menés les premiers travaux comparatifs, comme l'intercomparaison de l'Organisation Mondiale de la Météorologie (WMO, 1975). Cette étude, qui a porté sur des modèles conceptuels testés sur des bassins variés, n'a cependant pas permis de dégager des conclusions très nettes sur les performances relatives des modèles testés. Par la suite, de nombreuses autres études comparatives ont été menées, testant différents types de modèles, à différents pas de temps, sous des conditions climatiques variées, sur des nombres variables de bassins et sur différents modes de fonctionnement (simulation, prévision, transposition). La revue chronologique complète de ces comparaisons serait un peu longue dans le cadre de ce mémoire et nous renvoyons le lecteur à la liste relativement complète des comparaisons établie par Michaud et Sorooshian (1994) et Refsgaard et Knudsen (1996). Nous adoptons ici une approche plus thématique et nous ne citerons que quelques travaux dont les conclusions nous semblent intéressantes pour notre démarche de recherche. Un point commun à ces travaux est que leur évaluation porte sur la capacité des modèles à reproduire les débits observés. Le débit est en effet la sortie présentant le principal intérêt dans ces modèles. C'est uniquement sur cette variable que portera notre évaluation. Par ailleurs, nous ne nous intéresserons ici qu'aux résultats obtenus en dehors des phases de calage, plus représentatifs de la valeur intrinsèque des modèles.

Les travaux comparant des modèles 'boîte noire' à d'autres types de modèles sont peu nombreux. Weeks et Hebbert (1980) montrent qu'un modèle de type ARMA donne des résultats aussi satisfaisants que des modèles conceptuels complexes (Sacramento, Stanford, Monash) sur trois bassins du sud-ouest de l'Australie sous un climat de type méditerranéen. En revanche, Chiew *et al.* (1993) montrent sur huit bassins australiens aux conditions climatiques variées que de simples équations non-linéaires ne réussissent pas à donner de résultats satisfaisants aux pas de temps journalier ou mensuel par rapport à des modèles conceptuels simples ou complexes. De même, Loague et Freeze (1985) concluent à un constat d'échec dans les tests d'un modèle régressif linéaire et d'un modèle d'hydrogramme unitaire en prévision, dans une évaluation par événement sur trois bassins versants expérimentaux aux Etats-Unis. Ces modèles sont un peu plus satisfaisants en simulation, et en tous cas meilleurs qu'un modèle à fondements physiques. Hsu *et al.* (1995), eux, indiquent qu'un réseau de neurones appliqué sur un bassin américain, obtient de meilleurs résultats qu'un modèle ARMAX et que le modèle conceptuel complexe de Sacramento. Ce sont également les conclusions auxquelles aboutissent Coulibaly *et al.* (2000) dans un contexte de prévision des débits sur un bassin canadien: le réseau de neurones donne sur des prévisions de un à sept jours des résultats plus satisfaisants qu'un modèle conceptuel et un modèle ARMAX.

De ces études, il est donc difficile de tirer des conclusions claires quant à la valeur relative des approches que nous avons qualifiées de 'boîte noire'. Ce manque de tendance nette s'explique selon nous par le fait que ces comparaisons, exceptée celle de Chiew *et al.* (1993), ont été

menées sur un trop petit nombre de bassins ou dans des conditions climatiques à chaque fois trop homogènes. Leurs conclusions se trouvent de ce fait trop dépendantes des caractéristiques des bassins tests.

Quelques études se sont attachées à comparer des modèles fondés sur la physique avec des modèles 'boîte noire' ou des modèles à réservoirs. Une conclusion de l'étude de Loague et Freeze (1985) citée précédemment est que l'application des modèles fondés physiquement pose le problème de mesure de la variabilité spatiale des entrées du modèle (pluie) et surtout des caractéristiques du bassin (propriétés des sols par exemple). N'obtenant pas de meilleurs résultats que des modèles très simples, ils tirent un constat d'échec de cette approche complexe. Michaud et Sorooshian (1994) ont appliqué, sur un bassin expérimental américain sous climat semi-aride, un modèle fondé sur la physique et deux versions, l'une distribuée et l'autre globale, du modèle du SCS couplé à un hydrogramme unitaire. Leurs résultats (évaluation par événement) montrent que lorsque l'on dispose de données permettant de réaliser un calage, le modèle le plus complexe présente un avantage très limité par rapport au modèle simple mais cependant distribué.

Ces conclusions rejoignent celles de Wilcox *et al.* (1990) qui montrent que la sophistication du module d'infiltration d'un modèle simple, remplacé par un module à base physique, n'apporte qu'une amélioration marginale dans les résultats du modèle (évaluation par événement sur six bassins américains sans calage). Enfin, Refsgaard et Knudsen (1996) ont mené une comparaison détaillée de trois modèles allant du conceptuel global au fondé physiquement et distribué, sur trois bassins au Zimbabwe. Cette étude a porté cette fois-ci sur des simulations en continu et non par événements. Lorsque des données sont disponibles pour le calage, aucune supériorité n'apparaît pour le modèle le plus complexe, et l'avantage de ce dernier est marginal lorsque l'on se place en situation de bassins non jaugés.

Il ressort donc assez clairement que les modèles fondés sur la physique sont en état de relatif échec. Le degré de réalisme qu'ils visent et les efforts consentis pour les appliquer (collecte de données par exemple) devraient leur permettre de faire mieux que des modèles beaucoup plus simples, ce qui n'a pu être mis en évidence dans aucune étude comparative. Par ailleurs, si l'avantage d'une approche distribuée par rapport à une approche globale est de pouvoir fournir une représentation spatialisée du fonctionnement du bassin, la validation de cette représentation reste délicate et ce qu'elle apporte dans un contexte de simulation pluie-débit ne ressort pas non plus clairement. Même si la représentation explicite de l'hétérogénéité spatiale paraît souhaitable (en permettant notamment de simuler les variations spatio-temporelles d'autres variables d'intérêt, comme par exemple l'état de saturation des sols), elle ne semble pas trouver jusqu'à présent de réelle justification. Enfin, les modèles fondés physiquement, du fait de leur complexité, restent encore très lourds d'utilisation et très exigeants en capacité de calcul (même si cet aspect devient de moins en moins un facteur limitant). Yang *et al.* (2000) indiquent par exemple que la simulation de deux ans de données au pas de temps horaire nécessite 72 heures de calcul pour le modèle fondé physiquement MIKE SHE et seulement deux minutes pour le modèle conceptuel distribué TOPMODEL.

Nous allons maintenant analyser les études comparatives qui ont impliqué plusieurs modèles globaux à réservoirs, parmi lesquels nous pouvons citer les travaux de Kite (1975,1978), Moore et Mein (1975), WMO (1975, 1986, 1992), Weeks et Hebbert (1980), Franchini et Pacciani (1991), Chiew *et al.* (1993), Franchini *et al.* (1996), Zhang et Lindström (1996) Gan *et al.* (1997), Ye *et al.* (1997) et Perrin et Littlewood (2000). Un premier point à noter est qu'il est difficile de tirer de toutes ces études de réelles conclusions quant à la supériorité de tel ou tel modèle. Aucun ne semble être en mesure de fournir dans tous les cas de figure les résultats les plus satisfaisants. Les enseignements de ces travaux comparatifs restent la plupart

du temps conditionnés par leurs objectifs, les caractéristiques des bassins employés ou la méthodologie de comparaison (calage, critères...).

Toutes ces études ont été réalisées sur un nombre limité de bassins, généralement moins d'une dizaine. Si certaines regroupent des bassins situés sous des conditions climatiques variées (WMO, 1975; Chiew *et al.*, 1993), d'autres en revanche se sont intéressées à des conditions particulières, notamment à dominante méditerranéenne ou semi-aride (Weeks et Hebbert, 1980; Gan *et al.*, 1997; Ye *et al.*, 1997). Si ces études centrées sur des conditions très spécifiques affichent des objectifs précis, nous pensons que la qualité et la fiabilité d'un modèle sont bien davantage mises en valeur lorsque l'on s'intéresse à des conditions variées qui mettent bien plus à l'épreuve ses capacités et limites.

Certaines de ces études comparatives ont été appliquées sur des méthodologies complètes de modélisation, c'est-à-dire que chaque modèle est appliqué dans son propre environnement de développement et suivant les procédures préconisées par le concepteur (WMO, 1975, 1986, 1992; Perrin et Littlewood, 2000). Il est alors difficile d'évaluer le pouvoir réel de simulation des structures elles-mêmes, c'est-à-dire la nature et l'agencement des relations au sein du modèle, qui sont au cœur même de la réussite du modèle. Le choix de différentes procédures de calage par exemple peut venir 'parasiter' le jugement que l'on porte sur les capacités des structures. Nous pensons donc qu'une étude comparative de modèles du même type doit s'attacher à les mettre dans une même situation. C'est à cette condition que l'on peut se faire une idée objective de la valeur de leurs structures dans des travaux de simulation.

Le nombre trop réduit de bassins et la non-homogénéité des cadres comparatifs rendent souvent obscures les raisons de supériorité ou d'échec des modèles. Elles sont souvent davantage avancées comme des hypothèses que réellement démontrées. Par ailleurs, les différences qui apparaissent entre modèles sont souvent trop peu significatives. Nombre d'études comparatives citées précédemment tendent d'ailleurs à indiquer une relative équivalence entre les modèles, comme le montrent par exemple les travaux de Franchini et Pacciani (1991) complétés par Franchini *et al.* (1996): huit modèles conceptuels très différents donnent des résultats très proches dans le cas d'un bassin versant italien. Il apparaît également que d'une étude à l'autre, le classement de certains modèles change, ceci étant soit le signe des problèmes des approches comparatives mentionnés précédemment, soit l'indication que ces modèles ont des qualités très voisines.

Concernant l'influence de la complexité (nombre de paramètres optimisés), Ye *et al.* (1997) montrent qu'un modèle à six paramètres semble aussi satisfaisant qu'un modèle à 22 paramètres. Pour des modèles comprenant entre 9 et 21 paramètres, Gan *et al.* (1997) concluent que c'est davantage la structure plutôt que la complexité qui joue un rôle dans les performances du modèle. Cependant, leur étude n'inclut pas de modèles relativement simples (moins de neuf paramètres). Enfin, les conclusions de Chiew *et al.* (1993) vont dans le sens inverse de celles de Ye *et al.* (1997): ils concluent qu'un modèle complexe à 17 paramètres donne des résultats nettement supérieurs aux modèles plus simples (trois et sept paramètres).

De toutes ces remarques ressortent des conclusions assez floues sur les comparaisons de modèles conceptuels globaux.

1.2.6. Quels objectifs pour une nouvelle étude comparative ?

Les méthodologies des travaux comparatifs que nous venons d'analyser présentent des limites ne permettant pas de tirer des conclusions toujours très nettes quant aux valeurs respectives des modèles testés. Nous nous proposons donc ici de réaliser un travail comparatif pour

évaluer des structures de modèles, qui présentera les trois caractéristiques principales suivantes:

- sélection d'un grand nombre de structures faisant appel à des concepts et des formulations mathématiques variées,
- vaste échantillon de données regroupant un grand nombre de bassins versants,
- cadre de comparaison homogène pour toutes les structures.

Comme nous l'avons mentionné en introduction, la structure est entendue comme l'assemblage des composants mathématiques du modèle. Elle constitue le noyau de méthodologies de modélisation pouvant inclure, dans un sens plus large, des méthodes de calage, de sélection de données etc.. Dans la suite cependant, nous utiliserons parfois abusivement le terme modèle pour en fait désigner sa structure.

En se restreignant à des modèles conceptuels ou empiriques globaux, nous ne confronterons pas des modèles fondamentalement différents. Les objectifs de la comparaison seront donc centrés sur les aspects principaux suivants:

- évaluation des qualités des modèles,
- rôle de la complexité sur les performances, la robustesse et la fiabilité des modèles,
- rôle de la structure sur les performances des modèles,
- existence de complémentarités entre modèles,
- possibilité de construire des structures améliorées.

Les tests des modèles seront réalisés au pas de temps journalier. Ce pas de discrétisation temporelle est bien adapté pour étudier le comportement d'un grand nombre de bassins versants qui ont un temps de réponse de l'ordre de ou supérieur à la journée. Nous verrons néanmoins au chapitre 2 que notre échantillon test comporte des bassins de faible superficie (moins de quelques km²) et vraisemblablement des bassins à réponse rapide, ayant des temps de concentration caractéristiques inférieurs à 24 h. Le pas de temps journalier correspond alors à un niveau d'agrégation (de globalisation) temporelle ne permettant pas de décrire toute la richesse de la dynamique du bassin, en particulier lors d'événements rapides de crue. Les implications de la variété des superficies des bassins sur la performance des modèles sont commentées au chapitre 5.

Notons par ailleurs que les données hydrométriques disponibles dans les bases de données sont majoritairement au pas de temps de la journées, ce qui rend leur collecte plus aisée. Dans la partie suivante, nous exposons la démarche de choix des modèles et d'élaboration des structures testées.

1.3. Sélection de structures de modèles

1.3.1. Choix des modèles

La modélisation pluie-débit a suscité un intérêt croissant depuis quatre décennies dans la communauté des hydrologues. Parmi les très nombreux modèles proposés, des dizaines font partie de la classe des modèles à réservoirs conceptuels ou empiriques, qui nous intéresse particulièrement ici.

La sélection des modèles s'est appuyée sur une vaste recherche bibliographique, qui a permis d'identifier un grand nombre de modèles. L'échantillon des modèles pouvant être testés a été restreint en excluant ceux trop complexes en terme de nombre de paramètres optimisables tels que les modèles SSARR (Speers, 1995), Stanford (Crawford et Linsley, 1963), LASCAM (Sivapalan *et al.*, 1996), IRMB (Bultot et Dupriez, 1976), HYSIM (Manley, 1975) ou encore CREAMS (Knisel et Williams, 1995).

Les modèles retenus, qui incluent au plus 18 paramètres dans la version originale, ont souvent été développés dans des contextes très différents, avec des objectifs variés, reposant sur de nombreux concepts. Pour la majorité, ils ont déjà été testés de façon satisfaisante sur des cas réels. Certains ont été très intensivement utilisés depuis leur création. C'est par exemple le cas du modèle HBV qui, initialement développé pour les climats du nord de l'Europe, a ensuite été appliqué dans plus de 30 pays à travers le monde (Bergström, 1995). Ce sont finalement 38 structures de modèles que nous avons choisi de tester.

On pourrait penser que la variété des objectifs de développement ne met pas tous les modèles sur un pied d'égalité. Cependant, l'idée de ce vaste regroupement de modèles était avant tout de rassembler un nombre aussi grand que possible de structures, c'est-à-dire d'idées sur la façon d'agencer des outils mathématiques permettant de simuler la transformation pluie-débit.

Pour situer la place de ces modèles conceptuels par rapport à des modèles de type 'boîte noire', nous avons choisi de tester également un modèle simple sous forme d'équation. Il sera utilisé comme modèle élémentaire. Sélectionné arbitrairement, il n'entend pas être représentatif de l'ensemble des modèles du type 'boîte noire' que nous avons présentés précédemment. Il donne seulement une alternative simple de modélisation. D'autres modèles que nous avons qualifiés de 'boîte noire' tels que les réseaux de neurones par exemple auraient sans doute conduit à des résultats différents. La simplicité de calage et la non-linéarité de l'équation proposée par Tsykin (1985) ont paru intéressantes. Ce modèle a déjà été testé par Chiew *et al.* (1993) dans leur étude comparative. L'équation donne la valeur du débit en fonction des pluies antérieures d'après:

$$RUN_i = a + b \cdot RAIN_i^c \cdot RAIN_{i-1}^d \cdot RAIN_{i-2}^e \dots \quad \text{Eq. (1.2)}$$

où RUN_i est le débit au pas de temps i , $RAIN_i$, $RAIN_{i-1}$, $RAIN_{i-2} \dots$ sont les pluies au pas de temps i , $i-1$, $i-2 \dots$ respectivement, et a , b , c , d , $e \dots$ sont les paramètres du modèle. Pour retenir une version simple et introduire une composante de saisonnalité, nous proposons la formulation voisine suivante:

$$RUN_i = \frac{X3}{1 + \frac{PE_i}{X2}} \prod_{j=0}^n \left(1 + \frac{RAIN_{i-j}}{X1} \right)^{X4^{1-j}} \quad \text{Eq. (1.3)}$$

où PE_i est l'évapotranspiration potentielle au pas de temps i , n la mémoire du système en nombre de pas de temps et $X1$, $X2$, $X3$ et $X4$ sont quatre paramètres. La prise en compte de l'évapotranspiration permet de réduire l'importance des précipitations antérieures lors de la saison chaude. La progression géométrique des exposants en fonction de $X4$ diminue le nombre de paramètres et donne des résultats relativement satisfaisants.

1.3.2. Les enjeux de la parcimonie

L'étude du rôle de la complexité est un point important de notre étude. Nous avons voulu voir dans un premier temps quels enjeux cette complexité représente et si l'on peut restreindre la plage du nombre de paramètres optimisables davantage que ce que nous l'avons fait dans une

première sélection (moins de 18 paramètres, alors que certains modèles dépassent la trentaine de paramètres).

Dans un article de référence, Nash et Sutcliffe (1970) ont présenté quelques principes de construction de modèles comportant des paramètres optimisables. Ils ont exprimé notamment le nécessité d'avoir des structures simples et sans redondances. Ils y ont ajouté également l'exigence d'une capacité d'adaptation, dans laquelle l'ajout de composantes nouvelles dans la structure du modèle ne peut être acceptable que dans la mesure où cela augmente de façon significative la pertinence du modèle et sa robustesse. Cependant, conscients du besoin de traiter des problèmes tels que les impacts de changements environnementaux, de nombreux modélisateurs ont choisi de développer des modèles plus ambitieux (et donc avec plus de paramètres), sans suivre les recommandations de Nash et Sutcliffe (1970). On se retrouve donc aujourd'hui dans une situation où beaucoup de modèles sont sur-paramétrés, et de ce fait confrontés au problème d'équifinalité exposé par Beven (1993): des jeux de paramètres différents peuvent conduire à des résultats équivalents du modèle. Ceci signifie qu'une large incertitude pèse sur la détermination de ces paramètres, avec notamment de gros problèmes d'identification au cours de la procédure de calage (voir par exemple Gupta et Sorooshian, 1983). Ce problème ne serait pas gênant s'il ne remettait en cause certaines applications du modèle où une détermination fiable des paramètres est fortement souhaitable, comme la régionalisation dont nous reparlerons au chapitre 8 ou la détection des effets de changements d'occupation des sols (Mein et Brown, 1978; Nascimento, 1995; Nandakumar et Mein, 1997).

La question de l'incertitude sur les paramètres a attiré un intérêt croissant depuis une vingtaine d'année, en partie du fait de l'influence qu'elle est susceptible d'avoir sur l'utilisation du modèle. Pour contraindre davantage les valeurs des paramètres optimisés, certains auteurs ont proposé d'utiliser des informations supplémentaires sur le bassin, en explorant la possibilité de valider certaines variables internes des modèles conceptuels. Il s'agit par exemple d'essayer de voir si le niveau dans un réservoir 'souterrain' du modèle correspond effectivement à des niveaux piézométriques. Les recherches dans ce sens menées par Kuczera et Mroczkowski (1998) sur le modèle CATPRO et par Lamb *et al.* (1997) sur TOPMODEL ont montré qu'une nette amélioration dans la détermination des paramètres pouvait être obtenue. En revanche, parallèlement aucune amélioration n'est constatée dans la simulation des débits. Ceci indique qu'il est difficile d'avoir un modèle à plusieurs sorties et que la représentation conceptuelle de sous-processus sur le bassin est rarement exactement en accord avec la réalité.

Bien que la sur-paramétrisation soit un problème bien connu, de façon assez surprenante seuls quelques modélisateurs adhèrent au principe de parcimonie, de minimalité recommandé par Nash et Sutcliffe (1970) et plus récemment par Jakeman et Hornberger (1993) et Wheeler *et al.* (1993), sorte de rasoir d'Occam, garde-fou à toute complication superflue du modèle. Cependant, de nombreux exemples tirés de la littérature semblent converger vers ce point de vue. Mein et Brown (1978) dans le cas du modèle SFB modifié à 13 paramètres montrent qu'une réduction drastique du nombre de paramètres ne fait baisser que légèrement les performances du modèle. Chiew et McMahon (1994) indiquent dans le cas de MODHYDROLOG que seuls 9 des 19 paramètres du modèle ont généralement besoin d'être optimisés pour obtenir des résultats satisfaisants. Zhao et Liu (1995) remarquent que les sorties du modèle Xinanjiang ne sont généralement sensibles qu'à sept des quinze paramètres optimisables. Dans le cas de la famille de modèles SMAR, Tan et O'Connor (1996) montrent que la version à huit paramètres est plus robuste que la version SMARG à neuf paramètres. Récemment, Abdulla *et al.* (1999) observent que, parmi les quatre paramètres contrôlant le débit de base du modèle ARNO, certains ne semblent pas utiles et qu'une formulation simplifiée faisant intervenir moins de paramètres pourrait donner des résultats aussi

satisfaisants. Enfin, Uhlenbrook *et al.* (1999) rapportent aussi que de bonnes simulations ont pu être obtenues avec le modèle HBV avec des jeux de paramètres très différents (même dans le cas des paramètres les plus sensibles), et que l'amélioration de la qualité des simulations était faible lorsque des versions plus complexes étaient testées.

La parcimonie semble donc un enjeu bien réel en modélisation. Les quelques exemples cités précédemment tendraient à indiquer que la modélisation conceptuelle pluie-débit pourrait donner des résultats satisfaisants avec moins d'une dizaine de paramètres. Beven (1989) indique d'ailleurs que seulement trois à cinq paramètres devraient suffire pour reproduire l'essentiel de l'information contenue dans les séries hydrologiques. Jakeman et Hornberger (1993) abondent en ce sens, indiquant que six paramètres semblent suffisants pour modéliser la relation pluie-débit sur la plupart des bassins.

1.3.3. Modifications des modèles et présentation des versions retenues

L'évaluation des modèles originaux complets n'est pas le but de notre travail. Notre objectif est de voir si certaines structures sont plus efficaces que d'autres et si cette efficacité est liée à la complexité de la structure reflétée par le nombre de paramètres libres. Nous avons choisi de limiter la complexité des structures testées à neuf paramètres optimisables, ce qui nous semble une limite supérieure acceptable étant donné les commentaires de la partie précédente.

Les structures originales ayant plus de neuf paramètres ont dû être simplifiées. Pour cela, nous avons fixé certains paramètres de faible sensibilité, suivant les études disponibles d'analyse de sensibilité et les commentaires des concepteurs de modèles. Dans certains cas, des relations ou parties de la structure ont été simplifiées.

Certaines modifications ont été faites dans les structures car elles permettaient d'obtenir de meilleurs résultats. Par exemple, nous avons renforcé la structure du module de routage des modèles mensuels retenus. Des modifications *ad hoc* ont aussi parfois été faites. Pour tenir compte du temps de transit excédant souvent la journée sur bon nombre de bassins, nous avons introduit de façon systématique une fonction de délai dans le temps (pure translation temporelle d'un nombre non entier de jours) dans les structures qui initialement ne comportaient pas de fonctions équivalentes. Ainsi, tous les modèles sont aisément applicables sur les grands bassins.

Les modèles à réservoirs ont majoritairement été conçus pour avoir un minimum d'exigence en données. Ils se contentent généralement de séries de pluie et d'évapotranspiration potentielle (ETP) en entrée. Des séries concomitantes de débit sont nécessaires pour le calage et l'évaluation du modèle. Dans le but de donner une quantité équivalente d'information sur le bassin à chacune des structures, les modèles conçus initialement pour utiliser d'autres caractéristiques du bassin que ces séries hydro-météorologiques ont été modifiés. C'est le cas de TOPMODEL, pour lequel la courbe de distribution des indices topographiques, généralement calculée à partir d'un modèle numérique de terrain du bassin, a été ici paramétrée par une fonction logistique en utilisant deux paramètres supplémentaires. C'est le cas également du modèle IHACRES qui utilise généralement des données de température au lieu de données d'ETP.

Remarquons que certaines de ces modifications éloignent parfois la version testée des représentations conceptuelles initiales de développement du modèle. Cependant, nous voulions ici tester avant tout la pertinence d'outils mathématiques vis-à-vis de la simulation des débits. Certains modèles seront de ce fait appliqués en dehors du domaine de validité préconisé initialement par les concepteurs. Nous verrons au chapitre 5 les conséquences d'un tel traitement des modèles.

Le Tableau 1.1 donne la liste des structures retenues, avec le nombre de paramètres correspondants. Un code de quatre lettres, choisi pour chacune d'elles, sera utilisé dans la suite pour désigner ces structures. Tout en rappelant le modèle d'origine, ce code a pour but de souligner de façon claire que ce ne sont pas les versions originales des modèles qui seront étudiées.

N°	Modèle	Code structure	Paramètres
0	Tsykin (1985)	TSYK	5
1	GR3J (Edijatno <i>et al.</i> , 1999)	GR3J	3
2	GR4J (Nascimento, 1995)	GR4J	4
3	GR3J à 4 paramètres (Edijatno <i>et al.</i> , 1999)	GR4K	4
4	MHR (Leviandier, 1993)	MHR0	4
5	Modèle C (Bonvoisin et Boorman, 1992)	BOOC	5
6	ABCD (Thomas, 1981)	ABCD	6
7	Modèle B (Bonvoisin et Boorman, 1992)	BOOB	6
8	Bucket (Thornthwaite et Mather, 1955)	BUCK	6
9	CREC (Cormary et Guilbot, 1973)	CREC	6
10	Gardenia (Thiery, 1982)	GARD	6
11	GR5J (Ma <i>et al.</i> , 1990)	GR5J	6
12	PDM (Moore et Clarke, 1981)	PDM0	6
13	IHACRES (Jakeman <i>et al.</i> , 1990)	IHAC	7
14	Martine (Mazenc <i>et al.</i> , 1984)	MART	7
15	MODALP (Arikan, 1988)	MODA	7
16	TANK (Sugawara, 1979)	TANK	7
17	TOPMODEL (Beven et Kirkby, 1979)	TOPM	7
18	mSFB (Summer <i>et al.</i> , 1997)	BOUG	8
19	CATPRO (Raper et Kuczera, 1991)	CATP	8
20	Haan (1972)	HAAN	8
21	MODGLO (Servat, 1986)	MODG	8
22	SIXPAR (Gupta et Sorooshian, 1983)	SIXP	8
23	SMAR (O'Connell <i>et al.</i> , 1970)	SMAR	8
24	TMWAM (Bobba et Lam, 1985)	TMWA	8
25	Wageningen (Warmerdam <i>et al.</i> , 1997)	WAGE	8
26	Xinanjiaog (Zhao <i>et al.</i> , 1980)	XINA	8
27	Arno (Todini, 1996)	ARNO	9
28	Cequeau (Girard <i>et al.</i> , 1972)	CEQU	9
29	Georgakakos et Baumer (1996)	GEOR	9
30	GRHUM (Loumagne <i>et al.</i> , 1996)	GRHU	9
31	HBV (Bergström et Forsman, 1973)	HBV0	9
32	HMS (Morel-Seytoux, 1999)	HMS0	9
33	Institute of Hydrology Lumped Model (Blackie et Eeles, 1985)	IHLM	9
34	MODHYDROLOG (Porter et McMahon, 1971)	MODH	9
35	NAM (Nielsen et Hansen, 1973)	NAM0	9
36	Dawdy et O'Donnell (1965)	ODON	9
37	Sacramento (Burnash <i>et al.</i> , 1973)	SACR	9
38	SDI (Langford et O'Shaughnessy, 1977)	SDI0	9

Tableau 1.1: Liste des modèles sélectionnés, avec le code et le nombre de paramètres de la version retenue

Nous avons fait, au cours de la présentation des différents types de modèles, la distinction entre modèles conceptuels et empiriques. D'après les informations que nous avons pu collecter, il semble que la majorité des 38 modèles soient conceptuels. Les premiers articles de présentation des modèles font très souvent référence à des interprétations conceptuelles de l'architecture. Le modèle GR3J présenté par Edijatno *et al.* (1999) est l'un des seuls modèles empiriques de notre échantillon. Sa structure a été développée à partir d'une architecture très simple, complexifiée progressivement dans le seul but d'améliorer les performances du

modèle (Michel, 1983; Edijatno, 1991; Nascimento, 1995). Le modèle IHACRES, dont le module de routage est fondé sur l'utilisation de l'hydrogramme unitaire, est qualifié par ses concepteurs de modèle hybride métrique-conceptuel (Wheater *et al.*, 1993).

Nous ne réaliserons pas ici une analyse détaillée de chacune des structures, comme ont pu le faire par exemple Franchini et Pacciani (1991) et Franchini *et al.* (1996). Pour chaque modèle, nous avons réalisé une fiche analytique regroupant un certain nombre de renseignements que nous avons pu collecter dans notre analyse bibliographique et une description de la version retenue. Ces fiches sont regroupées en Annexe 1.

A ce stade de la réflexion, nous préférons ne donner que quelques commentaires généraux sur ces modèles. Entre trois et neuf paramètres sont optimisables dans les structures. Elles comprennent de deux à cinq réservoirs. Certains paramètres ont des rôles à la fois de rendement et de transfert. Le nombre de paramètres agissant sur le routage est généralement supérieur au nombre de paramètres agissant sur le rendement, indiquant l'importance accordée au routage dans les modèles au pas de temps journalier. Chaque structure comporte une procédure de suivi d'humidité du bassin, mémoire des conditions antérieures, la plupart du temps par l'intermédiaire d'un ou plusieurs réservoirs.

Les modules de production, qui gèrent la proportion des différentes allocations de l'eau, impliquent de un à cinq paramètres. La gestion de l'eau à son entrée dans le modèle adopte des dispositifs variés. Certains adoptent une séparation de la pluie brute en une composante destinée à remplir un réservoir d'humidité et une autre destinée à être routée, cette séparation s'effectuant ou non en fonction d'états internes du système. D'autres font pénétrer la totalité de cette pluie directement dans un réservoir d'humidité, dont elle ressort ensuite sous forme de différentes vidanges. La détermination de l'évapotranspiration réelle a également des formulations diverses. Elle peut impliquer plusieurs réservoirs de suivi d'humidité. La plupart des structures font d'abord agir l'évapotranspiration à un taux potentiel soit sur la pluie brute, soit dans un réservoir d'interception ou encore dans un réservoir d'humidité au dessus d'un seuil. Ensuite, le restant éventuel d'évapotranspiration non satisfaite agit à un taux réel, la réduction de l'évapotranspiration potentielle étant généralement dépendante de la disponibilité en eau dans un ou plusieurs réservoirs du système. Certains modèles considèrent le bassin comme globalement imperméable: les sorties d'eau autres que le débit se résument à l'évaporation. D'autres, au contraire, permettent des pertes ou des échanges en eau pouvant représenter des interactions avec des nappes profondes.

Les procédures de routage, responsables de la distribution temporelle des volumes écoulés, dépendent de deux à sept paramètres. Elles simulent au moins deux composantes d'écoulement, et sont, selon les cas, globalement linéaires ou non linéaires. Beaucoup de modèle utilisent des réservoirs de routage à vidange linéaire, permettant une décroissance exponentielle des hydrogrammes. Les réservoirs à vidanges non linéaires sont moins fréquents mais adoptent des formes variées (puissance ou exponentielle par exemple). Quelques modèles simulent des remontées 'capillaires' au sein du modèle, permettant à de l'eau contenue dans un réservoir de remonter vers un autre sus-jacent. L'utilisation d'hydrogrammes unitaires existe également dans plusieurs modèles, ceux-ci adoptant des formulations mathématiques variées.

Notons enfin qu'il existe des ressemblances parfois fortes entre certains modèles. C'est le cas des six structures apparentées à la famille GR (GR3J, GR4J, GR4K, GR5J, GRHU et MHR0) qui reposent sur des formulations voisines. Quelques modèles utilisent des concepts voisins, comme par exemple celui de distribution de probabilité de capacité de réservoir que l'on retrouve dans les structures ARNO, PDM0 ou XINA.

1.4. Méthodologie de comparaison

1.4.1. Choix d'une démarche d'évaluation

La démarche de comparaison de modèles hydrologiques est assez complexe. Nous avons vu, dans les parties précédentes, que le modèle est construit à partir d'un objectif, sur une représentation du système réel. On est alors en droit de se demander s'il doit être jugé suivant le degré de satisfaction qu'il apporte aux questions qui sous-tendent son développement, ou suivant la fidélité de sa représentation de la réalité. Devant la complexité du système réel (le bassin versant), nous avons vu que même les modèles les plus complexes n'en sont que de subjectives et partielles imitations. Ils ne peuvent en effet prendre en compte qu'une petite partie de la réalité, celle jugée la plus pertinente dans l'explication ou la reproduction des phénomènes étudiés. Il semble donc difficile, en modélisation pluie-débit, de prendre la fidélité de représentation du bassin réel comme critère valide de comparaison. Elle n'en est pas moins souhaitable, mais reste, à l'heure actuelle, difficilement envisageable. L'adéquation entre question posée et réponse donnée par le modèle semble beaucoup plus instructive et primordiale dans l'utilisation que l'on fait du modèle. C'est avant tout suivant cet aspect que la crédibilité et l'utilité du modèle doivent être appréciées. C'est d'autant plus vrai que l'on se place dans un contexte d'hydrologie à vocation opérationnelle.

Partant de ce point de vue, il faut arriver à cibler l'évaluation du modèle. Il en existe en effet plusieurs modes d'utilisation, dont les plus communs sont la simulation, la prévision ou la transposition (application aux bassins non jaugés). Le modèle peut aussi être testé en conditions stationnaires ou non. Par ailleurs, il n'existe pas d'unique critère permettant d'apprécier les performances des modèles, ces derniers étant souvent dépendants des objectifs fixés. Klemeš (1986b) a voulu systématiser l'évaluation des modèles en proposant une procédure hiérarchique, comprenant quatre types de tests:

1. 'split-sample test': séparation de la période disponible en deux sous-périodes indépendantes; calage du modèle sur la première période et test en simulation sur la deuxième; puis échange des rôles des deux sous-périodes,
2. 'proxy-basin test': calage du modèle sur un bassin A et application à un bassin B; on échange ensuite le rôle des bassins A et B. Si les résultats sont satisfaisants dans les deux cas, le modèle est applicable en transposition sur un bassin C non-jaugé,
3. 'differential split-sample test': application du test n°1 mais avec non-stationnarité des conditions (caractéristiques climatiques différentes d'une période à l'autre),
4. 'proxy-basin differential split-sample test': application du test n°2 mais avec non-stationnarité des conditions climatiques.

Cette procédure hiérarchique de Klemeš (1986b) est très exigeante. Si elle paraît souhaitable et capable de mettre en évidence les qualités des modèles en les mettant à rude épreuve, elle est assez lourde à mettre en place sur un grand nombre de bassins versants. Seules deux études ont, à notre connaissance, appliqué dans sa totalité ce cadre de comparaison: Refsgaard et Knudsen (1996) l'ont utilisé pour tester trois modèles (NAM, MIKE SHE et WATBAL) sur trois bassins au Zimbabwe; Donnelly-Makowecki et Moore (1999) s'en sont servi pour tester TOPMODEL et des modèles simples à réservoirs sur deux bassins de Colombie Britannique au Canada.

La plupart des études comparatives existantes ont adopté le seul ‘split-sample test’ (avec dans certains cas le test en conditions non stationnaires, c’est-à-dire le test n°3). Nous adopterons ici une procédure identique (test n°1). Il nous semble en effet que ce test est assez robuste pour évaluer les qualités premières requises pour le modèle (souhaitables pour une grande partie des applications). Etant donné la taille de notre échantillon et la procédure de découpage des séries de données (voir chapitre 2), il paraît probable que l’on se placera implicitement dans un certain nombre de cas dans la situation du test n°3, qui sera ainsi partiellement appliqué.

1.4.2. Stratégie de comparaison

Une première étape a été la construction du cadre de comparaison. Après la sélection des modèles que nous avons exposée précédemment, chaque structure a été reprogrammée dans un environnement homogène. Un large échantillon de données a été constitué et pour chaque bassin, un découpage des enregistrements disponibles en sous-périodes a été réalisé. Cela est détaillé au prochain chapitre.

Nous avons ensuite choisi une procédure de calage, dont une évaluation est présentée au chapitre 3. Un critère d’optimisation et des critères d’évaluation ont été sélectionnés, comme cela est exposé au chapitre 4.

Une procédure identique a été utilisée avec chacune des structures sélectionnées. L’évaluation a été opérée de la façon suivante. Sur chaque bassin, la structure a été successivement calée sur chacune des périodes et testée en contrôle sur les périodes restantes. Par exemple, dans le cas d’un bassin avec six sous-périodes, au total six calages et 30 tests en contrôle ont été effectués. Les résultats ont pu ensuite être analysés statistiquement. Cela fait l’objet du chapitre 5.

1.5. Conclusion

Nous avons présenté dans ce chapitre le contexte général dans lequel s’inscrivent nos recherches. La modélisation pluie-débit, discipline de l’hydrologie, s’intéresse à la compréhension et la représentation de la transformation de la pluie en débit à l’échelle du bassin versant. Transcriptions mathématiques de cette transformation, les modèles qui ont été développés depuis une quarantaine d’années sont très variés, faisant appel à différentes perceptions du monde réel et répondant à différents objectifs. Notre analyse nous a conduit à identifier trois grandes catégories de modèles, à savoir les modèles ‘boîte noire’, les modèles ‘à réservoirs’ et ceux fondés sur la physique. Parmi eux, les modèles conceptuels ou empiriques ‘à réservoirs’ semblent les plus en mesure de faire avancer la compréhension des systèmes réels. Cette dernière est souvent entendue par rapport à un référentiel de connaissance bien établi, ici par exemple l’ensemble des équations aux dérivées partielles régissant le mouvement de l’eau dans les sols et les cours d’eau. Le modèle est alors jugé comme contribuant à la connaissance si ce noyau de connaissance y trouve un sens, c’est-à-dire en fait si le modèle dérive de ce noyau.

Or précisément dans le cas du bassin versant, les équations de la physique ne nous semblent pas être le moyen le plus adapté de comprendre son comportement et sa dynamique, du fait notamment de l’incapacité où l’on se trouve à décrire le bassin de façon fine (en particulier dans sa partie souterraine), rendant tout le référentiel de connaissance précédemment cité peu ou pas approprié. Partant de ce point de vue, il faut alors adopter une approche alternative, elle-même source de compréhension, qui proposerait d’autres descripteurs non liés à un noyau

monolithique de connaissance. Dans le cas du transport sédimentaire par exemple, les résultats des travaux de Hérouin (1998) indiquent qu'il semble préférable de développer des modèles basés sur des descripteurs synthétiques de la géométrie du lit et du régime hydrologique à l'échelle du tronçon, plutôt que de poursuivre dans une approche déterministe à l'échelle des processus élémentaires.

De ce fait, les modèles conceptuels ou empiriques globaux nous apparaissent comme étant les plus en adéquation avec le niveau actuel de compréhension du fonctionnement hydrologique du bassin, ce qui leur confère une bonne crédibilité dans un contexte d'hydrologie appliquée.

Il existe aujourd'hui un grand nombre de tels modèles, dont la multitude est le reflet des difficultés que rencontre la modélisation dans la représentation du comportement du bassin versant. Ce foisonnement tient également en partie de l'attitude commune qui consiste à croire que développer un modèle pour un cas spécifique est un gage de meilleure qualité de simulation et que chaque modèle n'est applicable que dans les conditions ou sur les bassins pour lesquels il a été développé. Face à cette production prolifique, l'évaluation et la validation approfondie de ces outils restent rares et difficiles. L'utilisateur est alors confronté à un véritable dilemme pour choisir le modèle qui pourra répondre de la manière la plus fiable aux problèmes qu'il se pose. Les évaluations comparatives sont, à cet égard, riches d'enseignements. Les conclusions tirées de celles réalisées jusqu'à présent manquent parfois de clarté, probablement en raison du faible nombre de bassins versants testés à chaque fois et de la non-homogénéité des cadres comparatifs qui masque la valeur intrinsèque des structures des modèles. Ce manque de conclusions nettes pourrait aussi être l'indication d'une certaine équivalence entre ces modèles.

Nous nous proposons donc dans ce travail de mettre en place une vaste comparaison de modèles à réservoirs conceptuels ou empiriques, impliquant un grand nombre de structures qui seront testées de façon homogène sur un large échantillon de bassins. Les objectifs premiers de cette étude sont d'explorer l'influence de la complexité (entendue comme le nombre de paramètres optimisés) et de la formulation mathématique de ces modèles sur leurs performances, leur robustesse et leur fiabilité. Cela devrait permettre de dégager des pistes pour construire ou améliorer des modèles, et apporter ainsi des réponses plus satisfaisantes à la modélisation de la transformation pluie-débit.

Les 38 structures retenues, dont nous avons limité la complexité à neuf paramètres optimisables, sont dérivées de modèles existants. Elles seront jugées uniquement sur leur capacité à reproduire les débits observés au pas de temps journalier, suivant une procédure de calage-contrôle sur périodes indépendantes

Dans les trois prochains chapitres, nous exposons les étapes de construction du cadre comparatif de notre étude, avec le regroupement d'un vaste échantillon de données, le choix et l'évaluation d'une procédure d'optimisation des paramètres et la sélection d'un jeu de critères d'appréciation des performances. Nous présenterons ensuite au chapitre 5 les résultats des tests réalisés.

Chapitre 2

Chapitre 2

Description de l'échantillon de données utilisé pour tester les modèles pluie-débit

2.1. Introduction

Le lien entre les constructions intellectuelles que sont les modèles et le monde réel repose sur les données qui caractérisent ce dernier. Notre approche de comparaison de structures de modèles pluie-débit et par la suite notre tentative d'amélioration d'une de ces structures repose sur un large échantillon de données constitué au début de notre étude et qui est ensuite resté inchangé pendant toute la durée de la recherche. Des données sur 429 bassins versants ont été rassemblées. Dans ce chapitre, nous nous proposons de donner un rapide aperçu de la localisation des bassins étudiés, de quelques unes de leurs caractéristiques ainsi que de l'origine et du type de données collectées. Etant donné la taille de notre échantillon et le but de notre recherche (évaluation de la capacité des modèles à reproduire les débits observés), nous n'avons pas cherché ici à donner une description très détaillée de ces bassins, avec des descripteurs physiques notamment. Nous avons plutôt essayé de voir comment cet échantillon de bassins peut être appréhendé dans notre étude et comment les caractéristiques de ces données peuvent favoriser l'évaluation des modèles.

2.2. Observations, instruments de connaissance du bassin et de construction de modèles

L'observation des phénomènes naturels est un outil essentiel à la connaissance et la compréhension des processus physiques. En hydrologie comme dans beaucoup de sciences de la Terre et de l'environnement, les données sont aussi des éléments de référence indispensables pour évaluer la qualité des modèles. Les observations représentent donc un élément fondamental des démarches de modélisation en hydrologie. Nascimento (1995) expose que la démarche de modélisation d'un système repose fondamentalement sur trois points: les entrées $I(t)$, l'opérateur de transformation OT (le système) et les sorties $S(t)$. Ici, le système représenté est le bassin versant. Connaissant $I(t)$ et $S(t)$, le modélisateur est alors confronté au problème de l'identification de OT .

Dans les démarches empiriques, c'est typiquement le problème qu'essaie de résoudre le modélisateur, en validant la représentation OT du système sur un grand nombre de données. Les approches conceptuelles, elles, au contraire n'utilisent pas les données pour l'élaboration de la structure du modèle, puisque la démarche consiste à proposer une représentation, une perception du monde réel en fonction notamment de connaissances acquises dans le domaine

de la physique. Les données servent seulement à fixer les paramètres laissés libres dans la structure pour permettre une adaptation du modèle à chaque bassin. Les exercices de test de la structure du modèle, au cas par cas, en confrontant les sorties du modèle aux données, permettent alors de valider ou d'infirmer les hypothèses émises sur les concepts utilisés.

Les observations (qui sont des données pour le modèle) constituent donc, quelle que soit l'approche adoptée, le lien au réel par lequel on peut évaluer l'adéquation du modèle dans sa représentation du système naturel. L'approche comparative que nous voulons réaliser a pour but de confronter les structures des modèles à la réalité des bassins dans un très grand nombre de cas. Il s'agit en fait d'évaluer leur capacité à représenter les systèmes réels, et donc déterminer lesquels traduisent le mieux le comportement global du bassin.

2.3. Description de l'échantillon

Nous allons fournir ici un descriptif de l'origine et du type de données utilisées, en distinguant successivement les bassins en France, aux Etats-Unis, en Australie, en Côte-d'Ivoire et au Brésil (par ordre décroissant du nombre de bassins). Pour tous les bassins, les données nécessaires pour l'utilisation des modèles se composent de chroniques concomitantes de pluie et débit au même pas de temps, et de données d'évapotranspiration potentielle (ETP). Etant donné le caractère global des modèles auxquels nous nous intéressons (pas de prise en compte explicite dans le modèle de l'hétérogénéité spatiale du bassin), les données d'entrée du modèle (pluie et ETP) sont des moyennes réalisées la plupart du temps à partir de plusieurs stations de mesures. Pluies et débits ont toujours été collectés au pas de temps journalier auquel nous nous intéressons ici.

2.3.1. Données en France

Nous disposons, dans notre échantillon, de données sur 307 bassins répartis sur le territoire métropolitain français. Ces données ont, pour la plupart, déjà été utilisées dans diverses études, thèses ou tests de modèles réalisés au *Cemagref*. Parmi ces bassins, nous retrouvons ceux utilisés par Edijatno (1991), Makhlouf (1994), Nascimento (1995) et Edijatno *et al.* (1999) dans leurs travaux successifs de recherche sur le modèle GR. Des bassins issus des travaux de Baudez (1997) et Loumagne *et al.* (1999) sont également inclus. Enfin, nous avons utilisé l'échantillon de bassins du Massif Central constitué par Andréassian (thèse en cours, 2001) pour ses recherches sur l'impact de l'évolution de la couverture forestière sur l'hydrologie du bassin versant.

Les données hydrométriques journalières utilisées sont issues de la banque de données HYDRO du Ministère de l'Environnement, qui assure le stockage de données d'environ 3300 stations en France. Cette base de données distingue de grandes unités hydro-géographiques, qui divisent le territoire français en six zones correspondant aux six Agences de bassin. Nous disposons de données sur chacun de ces grands bassins, à savoir:

- 31 stations sur le bassin Rhin-Meuse,
- 3 stations sur le bassin Artois-Picardie,
- 56 stations sur le bassin Seine-Normandie,
- 65 stations sur le bassin Loire-Bretagne,
- 61 stations sur le bassin Adour-Garonne,
- 91 stations sur le bassin Rhône-Méditerranée-Corse.

La répartition des stations hydrométriques est représentée à la Figure 2.1 La liste des stations est fournie en Annexe 2.

Pour chacune de ces stations, des données de pluies ont été rassemblées. Elles proviennent de la banque de données PLUVIO de Météo-France qui regroupe des données sur environ 15000 stations pluviométriques. Pour les 307 bassins versants, nous avons des données au pas de temps journalier de 740 postes pluviométriques, choisis en fonction de leur localisation par rapport aux bassins concernés. Certaines de ces stations sont utilisées pour plusieurs bassins (bassins emboîtés ou adjacents). Lorsque plusieurs postes pluviométriques étaient présents par bassin, la pluie de bassin au jour i a été calculée comme la moyenne des hauteurs de pluie tombée à chaque poste pluviométrique. Une moyenne arithmétique simple a été appliquée ici systématiquement. En cas de lacune sur un des postes pluviométriques au jour i , la station n'était alors pas prise en compte ce jour-là.

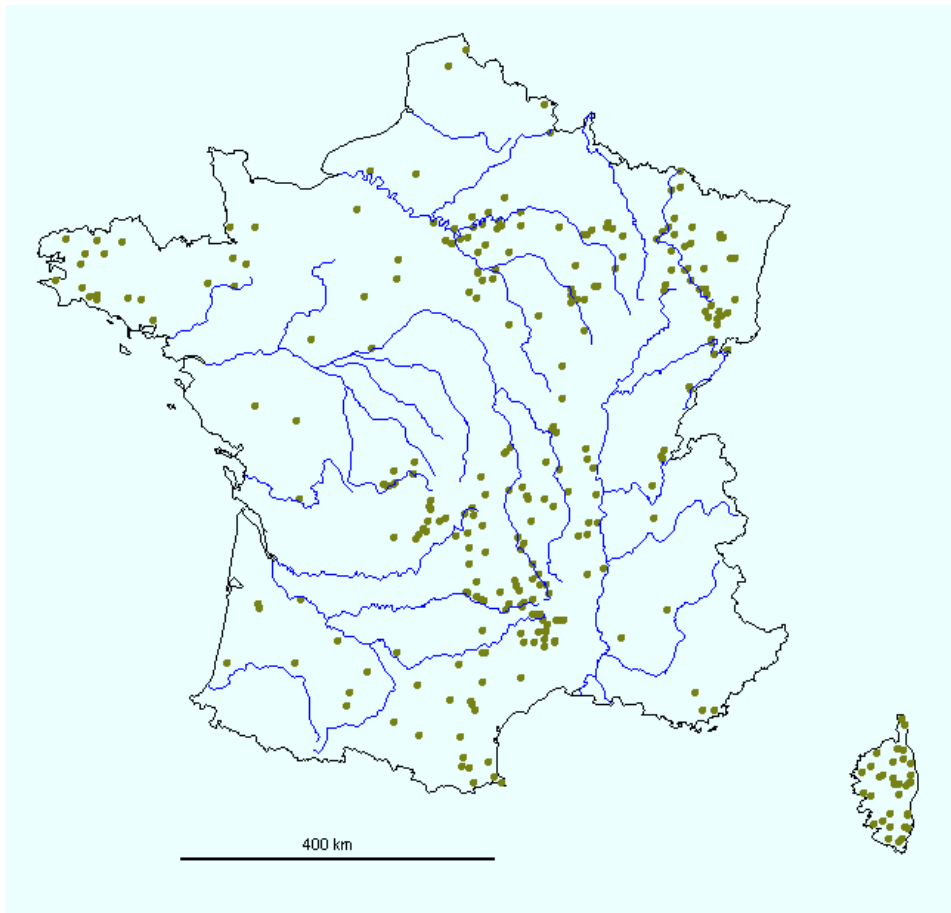


Figure 2.1: Localisation des 307 stations hydrométriques françaises

Les données fournies par Météo-France font apparaître un code neige permettant d'identifier la nature des précipitations. En cas de précipitations neigeuses, la hauteur d'eau équivalente est donnée, avec un code correspondant. Cette information a été utilisée ici dans un module neige simple, proposé par Makhoulf (1994), dont les paramètres (éventuellement optimisables) ont été fixés à des valeurs moyennes. Nous avons choisi d'appliquer ici ce module, quelle que soit la structure du modèle testé, car il permet d'améliorer sensiblement les résultats (gain de deux à trois points). Ne disposant d'information sur les précipitations neigeuses que dans les données de Météo-France, ce module de fonte de neige n'a été appliqué que sur les bassins français.

Les données d'ETP sont fournies également par Météo-France. Ce sont des données calculées à partir de la formule de Penman (1948). Les séries utilisées ici sont en fait constituées de moyennes décennales interannuelles. Pour toutes les années, les mêmes 36 valeurs décennales sont donc utilisées. La raison de l'utilisation de telles moyennes provient de l'assez faible variabilité interannuelle de cette variable et de la moindre sensibilité du modèle à cette entrée par rapport aux pluies (voir par exemple à ce sujet l'étude de Paturel *et al.*, 1995). Des tests effectués par Andréassian (thèse en cours, 2001) sur des bassins du Massif Central indiquent qu'une meilleure description de la variabilité temporelle de l'ETP ne permet pas d'améliorer de façon significative les performances des modèles testés.

Pour revenir à des valeurs journalières, une formule de lissage polynomiale a été utilisée (voir Edijatno, 1991), qui permet d'assurer des variations continues de l'ETP et d'éviter les brusques changements lorsque l'on passe d'une décennie à la suivante. Nous disposons de données sur 65 stations météorologiques. Pour chaque bassin versant, une ou plusieurs stations ont été utilisées, suivant la taille du bassin. Dans le cas de plusieurs stations, une moyenne (éventuellement pondérée) a été réalisée.

2.3.2. Données aux Etats-Unis

Les données des 82 bassins américains ont deux origines, la base de données de l'Agricultural Research Service (ARS) et la base de données de la Model Parameter Estimation Experiment (MOPEX). Une liste de ces bassins est fournie en Annexe 2.

◆ Données ARS

Notre échantillon compte 45 bassins versants provenant de la base de données de l'ARS (Thurman et Roberts, 1995). Cette dernière est gérée par le Water Data Center (WDC) qui est un groupe de l'ARS au sein de l'United States Department of Agriculture (USDA). Le WDC est chargé du stockage et de la dissémination des données collectées par les gestionnaires de bassins.

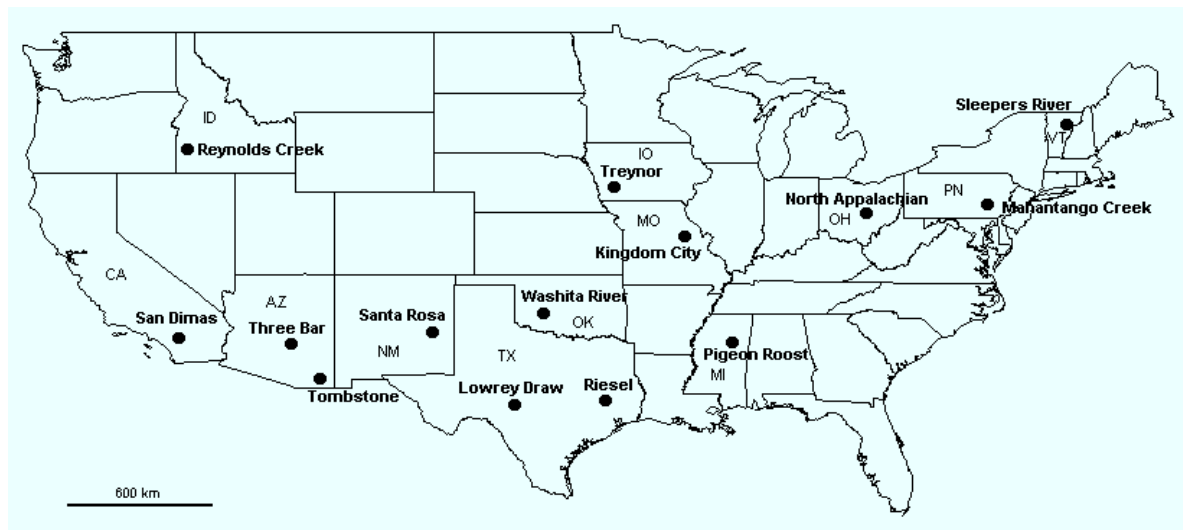


Figure 2.2: Localisation des sites des bassins versants américains expérimentaux issus de la base de données ARS

Ces bassins sont principalement de petits bassins versants expérimentaux ruraux. Ils sont situés dans les Etats de l'Arizona, de la Géorgie, de l'Idaho, de l'Iowa, du Nouveau Mexique, du Mississippi, du Missouri, de l'Ohio, de l'Oklahoma, de la Pennsylvanie, du Texas et du Vermont. A ces bassins s'ajoutent deux autres bassins versants expérimentaux (en Californie

et en Arizona) non géré par l'USDA. La localisation des sites est illustrée à la Figure 2.2, avec dans la plupart des cas, plusieurs stations par site.

Les données ARS recueillies étaient à pas de temps variable. Elles ont été converties à un pas de temps journalier, en adaptant les programmes de conversion mis au point par Faucher *et al.* (1996). Les valeurs de pluies de bassin ont été calculées par moyenne pondérée des stations pluviométriques disponibles. Généralement, nous n'avons eu que peu de pluviomètres pour calculer la pluie de bassin.

Nous ne disposons que de données de température minimale et maximale journalière pour tenir compte de la demande évaporative. Nous avons donc calculé des valeurs d'ETP à partir de la formule de Hargreaves et Samani (1982) basée sur la température et la durée théorique d'ensoleillement. Des valeurs journalières moyennes interannuelles ont été calculées.

◆ Données MOPEX

La base de données MOPEX a été créée dans le but de réunir des données sur un large échantillon de bassins versants à travers le monde et de mettre ces données à la disposition de la communauté scientifique pour permettre le test de modèles hydrologiques et atmosphériques, et l'estimation de leurs paramètres, à partir de caractéristiques physiques notamment.

Les 37 bassins versants issus de cette base de données sont situés dans les Etats de l'Arkansas, du Kansas, du Missouri, du Nouveau-Mexique, de l'Oklahoma et du Texas. Leur localisation est illustrée à la Figure 2.3. Les données d'ETP, moyennes journalières interannuelles, dérivent de l'atlas de Farnsworth *et al.* (1982).

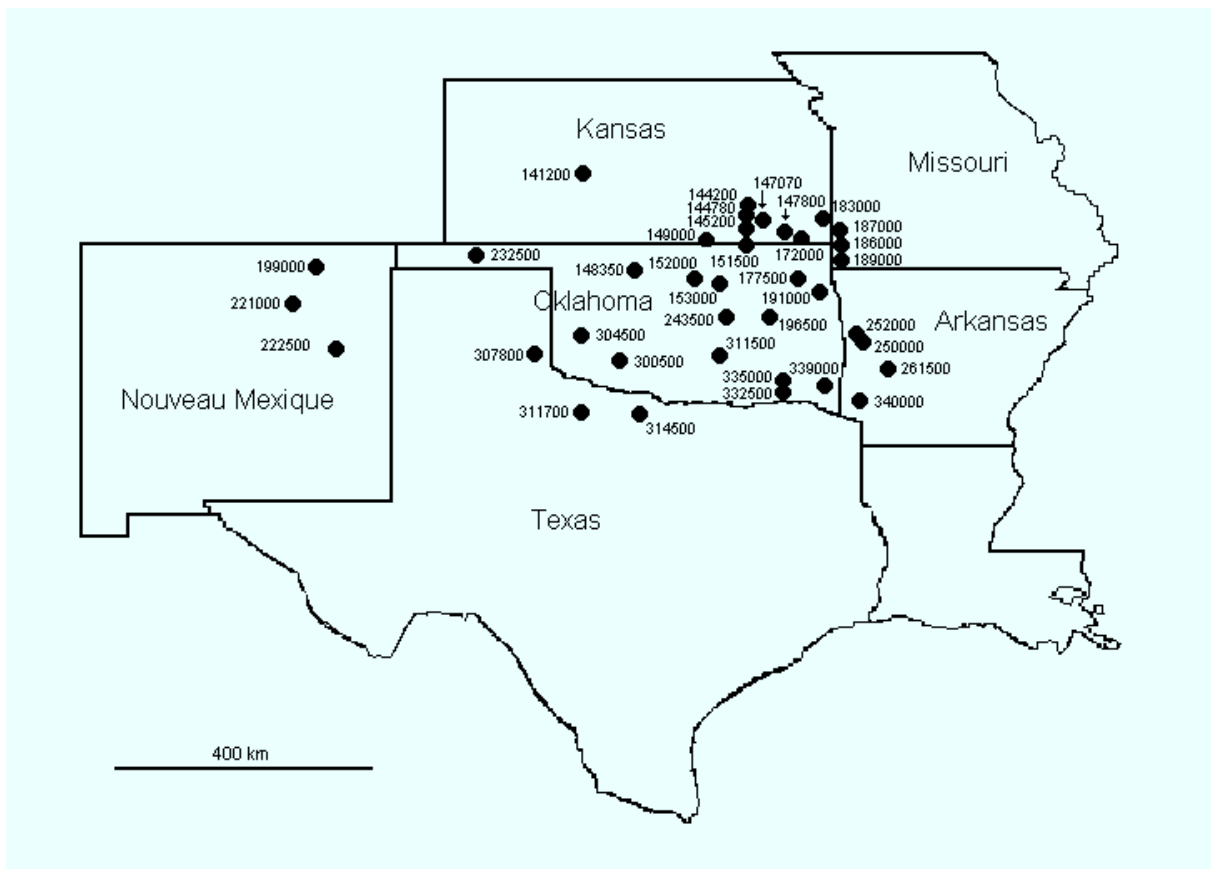


Figure 2.3: Localisation des stations hydrométriques des bassins versants issus de la base MOPEX

2.3.3. Données en Australie

Les chroniques pluie-débit-évaporation des 26 bassins australiens utilisés ici ont été préparées par Chiew et McMahon (1994) pour tester le modèle MODHYDROLOG. Ces bassins ont été sélectionnés à partir de la base de l'Australian Bureau of Meteorology, avec l'assistance des agences de l'eau de l'Etat et des Territoires, dans le cadre du projet 'Monitoring Climate Change and its Impacts on Australia's Water Resources' mis en place par l'Australian Water Resources Council.

Tout ou partie de cet échantillon a également été utilisé par Chiew *et al.* (1993, 1995, 1996), Summer *et al.* (1997), Ye *et al.* (1997) et Thyer *et al.* (1999) pour tester divers modèles (MODHYDROLOG, mSFB, GSFB, BEST, IHACRES, LASCAM). Bon nombre de ces bassins ont également été utilisés par Chiew et McMahon (1993) dans une étude de détection de tendance dans des séries hydrologiques. Chiew et McMahon (1994) mentionnent que cet échantillon regroupe une large gamme de conditions physiques et climatiques rencontrées en Australie, avec des bassins semi-arides dans le sud-ouest (notamment utilisés par Ye *et al.*, 1997) ou des conditions tropicales humides dans le Queensland. La localisation des bassins est illustrée à la Figure 2.4 et la liste des stations est fournie en Annexe 2.

Les données de pluies fournies sont des moyennes de bassin. Les données d'ETP sont des valeurs journalières datées, estimées à partir de la formule de Morton (1983). L'étude de Chiew et McMahon (1991) montre que cette formule est fiable dans le contexte australien.

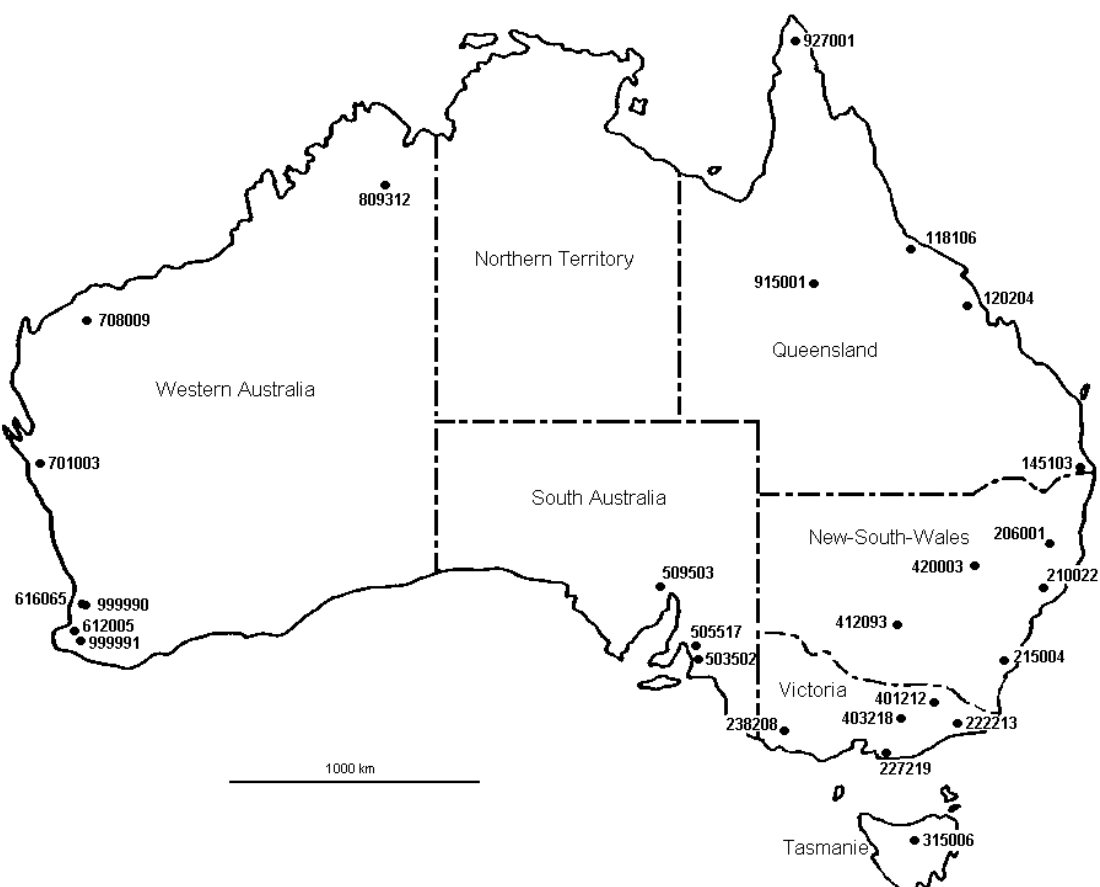


Figure 2.4: Localisation des 26 stations hydrométriques australiennes (adapté d'après Chiew et McMahon, 1994)

2.3.4. Données en Côte d'Ivoire

Les jeux de données des 10 bassins de Côte-d'Ivoire ont été fournis par l'équipe hydrologie de l'antenne de l'Institut de Recherche pour le Développement (IRD, ex-ORSTOM) à Abidjan. L'étude de ces bassins par l'IRD s'inscrit dans le cadre des programmes de recherche ERREAU (Evaluation Régionale des Ressources en EAU) et ICCARE (Identification et Conséquences d'une variabilité du Climat en AfRIque de l'ouest non sahélienne) dans le cadre de FRIEND-AOC (Flow Regimes from International Experimental and Network Data - Afrique de l'Ouest et Centrale non sahélienne) coordonné par l'Unesco. Ce programme a pour objet l'identification, dans toute la sous-région AOC, d'une éventuelle fluctuation climatique et l'étude de ses conséquences sur les ressources en eau (voir par exemple Servat *et al.*, 1997).

La liste des bassins est fournie en Annexe 2 et leur localisation apparaît à la Figure 2.5. Les données d'ETP sont calculées à partir de la formule de Penman (1948) et sont au pas de temps décadaire.

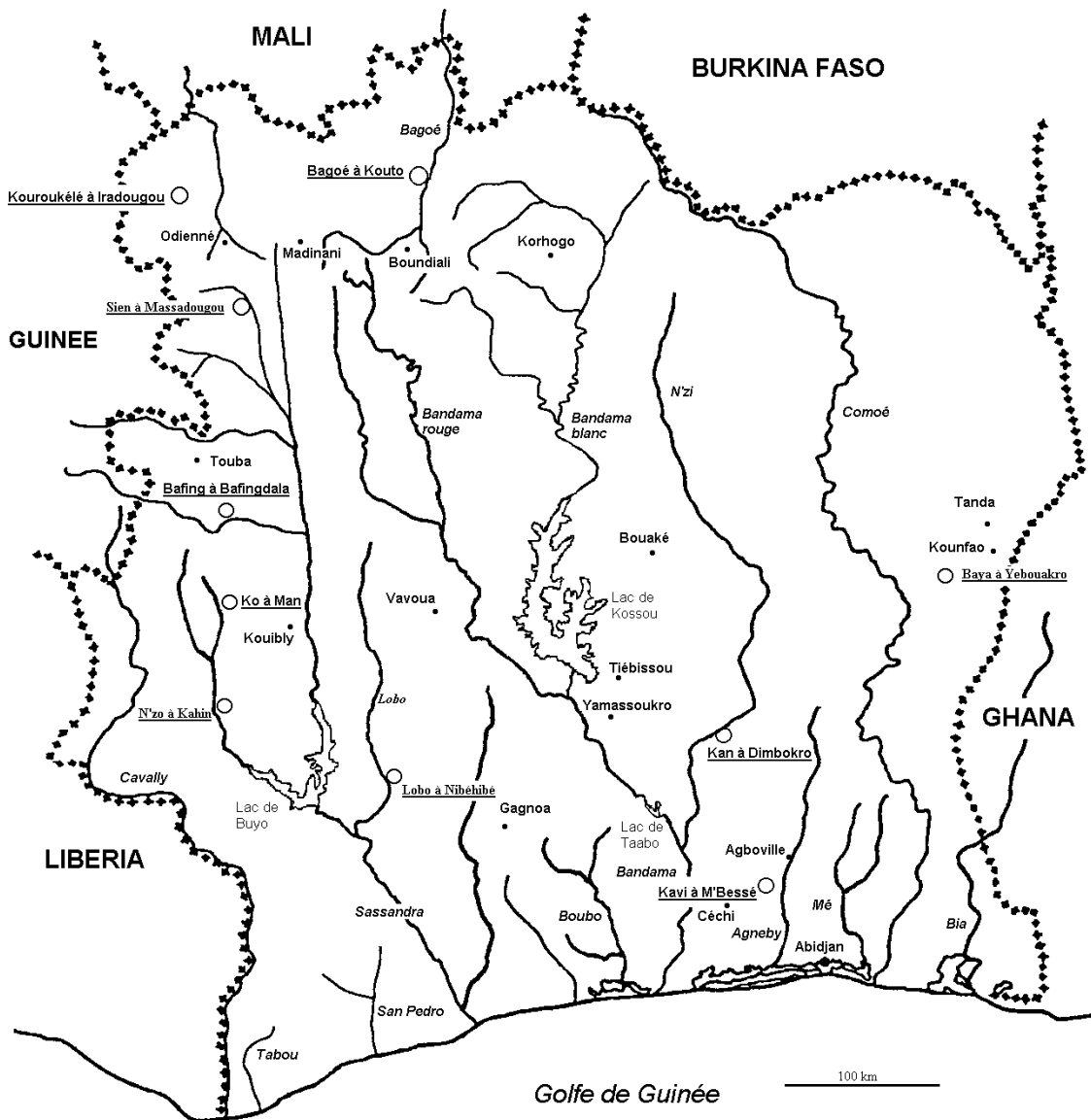


Figure 2.5: Localisation des bassins de Côte-d'Ivoire (adapté d'après Dezetter, 1991 et Ouédraogo, 1996)

Des modèles tels que GR2M, GR3, CREC, MODGLO ou VUB ont déjà été testés sur certains de ces bassins par Servat et Dezetter (1991,1992), Paturel *et al.* (1995), Ouédraogo (1996) et Ouédraogo *et al.* (1998).

2.3.5. Données au Brésil

Les données des quatre bassins versants brésiliens proviennent de l'université de Minas Gerais, Belo Horizonte, au Brésil. La liste des bassins est donnée en Annexe 2 et une carte de localisation est fournie à la Figure 2.6. Ces bassins ont été utilisés par Melo et Nascimento (1999) pour des tests comparatifs de modèles.

Les données d'ETP sont des données moyennes interannuelles, issues de mesures par bac Colorado.

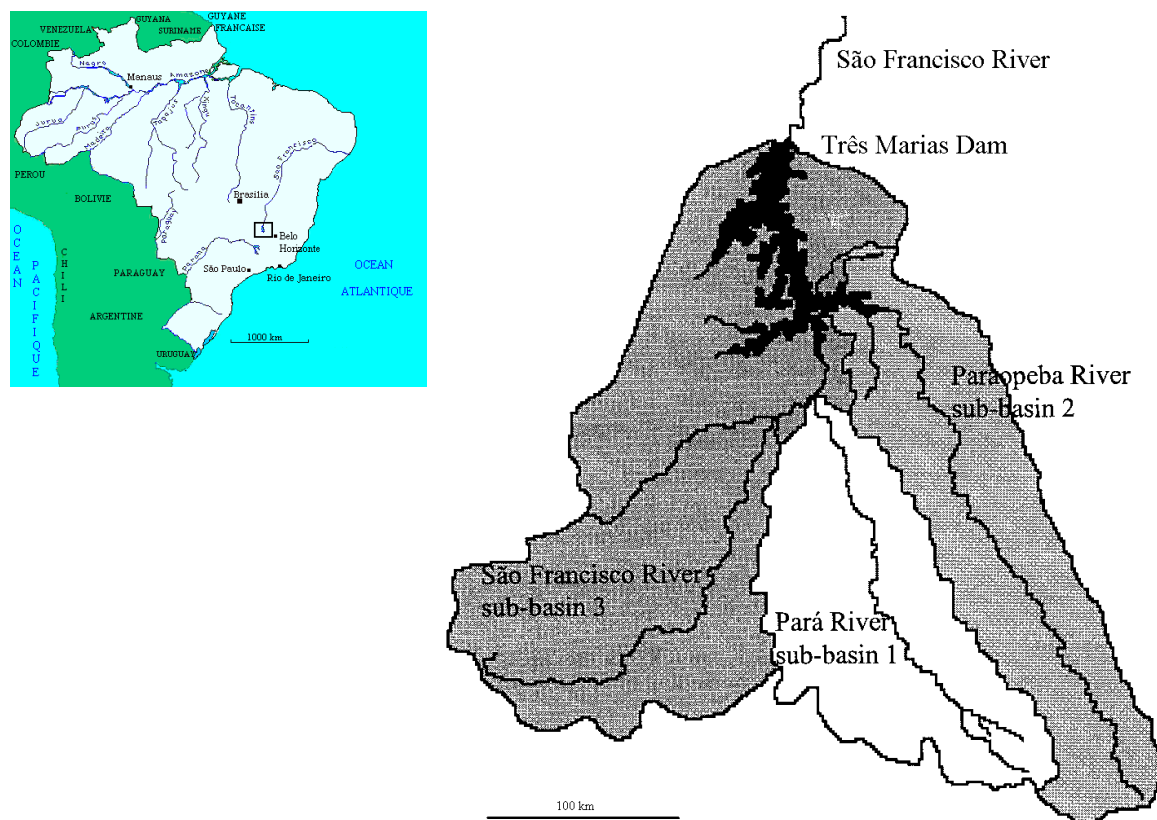


Figure 2.6: Localisation des bassins versants brésiliens (adapté d'après Melo et Nascimento, 1999)

2.4. Quelques caractéristiques des bassins

L'échantillon des 429 bassins que nous venons de présenter est très vaste. Au total, ce sont plus de deux millions de triplets pluie-débit-évapotranspiration qui vont nous permettre de tester les qualités des structures de modèle. Nous n'avons pas voulu entrer ici dans le détail des caractéristiques de chaque bassin, mais plutôt avoir une vue d'ensemble de notre échantillon.

La taille des bassins versants varie sur une large gamme, de 0,1 à plus de 10000 km², avec une superficie médiane de 100 km². 10 % des bassins ont une superficie inférieure à 10 km² et 10 % une superficie supérieure à environ 2500 km². Ceci est illustré à la Figure 2.7. Les deux

bassins les plus grands sont les bassins de la Seine à Paris (43800 km²) et le bassin du São Francisco au barrage de Três Marias (50600 km²). De telles tailles de bassins sont généralement considérées comme incompatibles avec l'utilisation d'un modèle global. Or dans les tests que nous avons effectués, il n'est pas apparu de limitations des performances des modèles sur ces grands bassins. Les bassins les plus petits sont essentiellement des bassins versants expérimentaux américains.

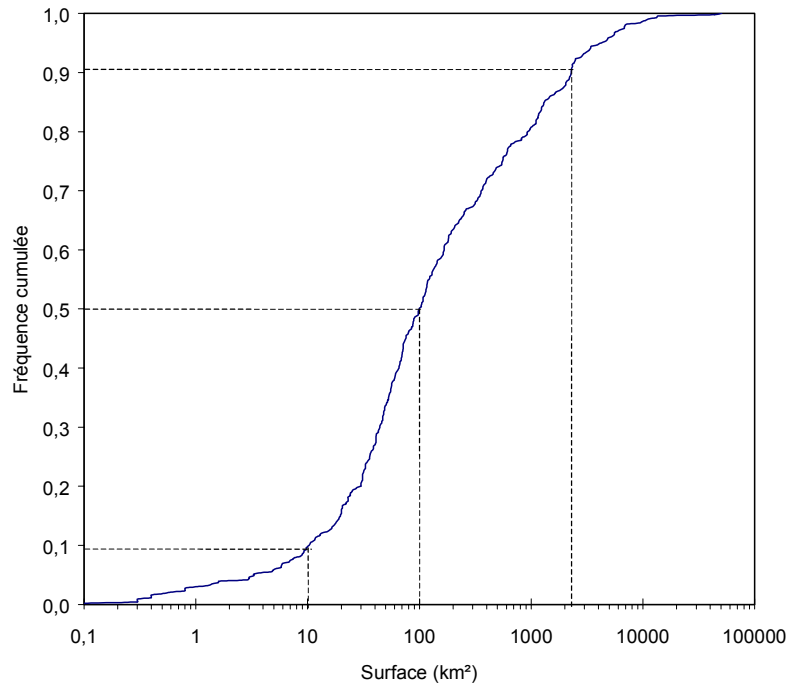


Figure 2.7: Distribution des superficies des 429 bassins versants

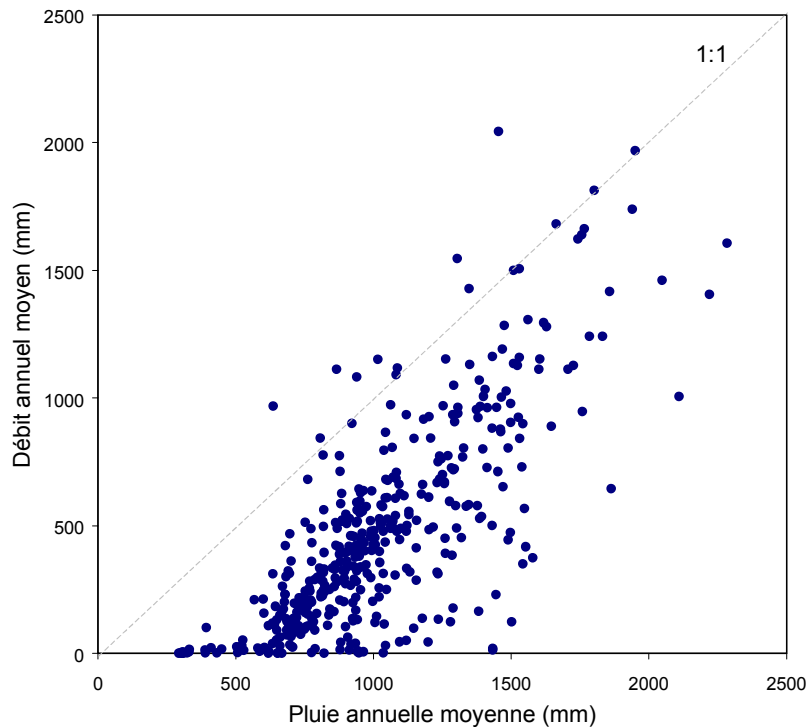


Figure 2.8: Pluie et débit annuels moyens des 429 bassins versants

Une grande variété de conditions hydro-climatiques est présente dans notre échantillon. On trouve ainsi des conditions semi-arides en Australie ou dans le sud des Etats-Unis, avec des cours d'eau connaissant des débits seulement quelques jours dans l'année, ou des conditions tropicales humides dans le sud de la Côte-d'Ivoire ou le nord de l'Australie. Des conditions tempérées sont présentes par exemple en France. Il y a cependant en France une assez grande diversité climatique, avec des influences méditerranéennes au sud ou continentales à l'est.

L'ETP annuelle moyenne varie sur les différents bassins entre 630 et 2040 mm, la pluie annuelle moyenne entre 300 et 2300 mm et le débit annuel moyen entre 0,2 et 2040 mm. La Figure 2.8, illustrant les pluies et débits annuels moyens des 429 bassins, montre que pour un petit nombre de bassins versants, les débits excèdent les pluies. Cette situation problématique pour l'application de modèles pluie-débit peut avoir plusieurs origines, parmi lesquelles une mauvaise définition des limites du bassin, une sous-estimation de la pluie de bassin, une surestimation des débits ou encore des apports de nappe provenant de l'extérieur des limites topographiques du bassin (non correspondance des bassins topographique et géologique).

Il existe par ailleurs au sein de l'échantillon une grande variété de comportements entre les saisons, avec des saisons pluvieuses et sèches très marquées et des saisons de hautes eaux et basses eaux très contrastées, ou inversement des régimes assez uniformes, tant au niveau de la pluie que des débits tout au long de l'année.

A partir des informations sur les caractéristiques physiques dont nous disposons, notamment sur des bassins versants américains ou sur les bassins australiens, il apparaît que l'échantillon regroupe une grande variété de couverts végétaux ou de caractéristiques géologiques, pédologiques ou topographiques.

2.5. Quelle critique des données dans le contexte de notre étude ?

La connaissance du contexte hydro-climatique ou de l'évolution du bassin, et la critique préalable des données sont probablement des aspects importants de toute étude hydrologique ponctuelle. En effet, une telle analyse préparatoire peut révéler l'existence de conditions limitant l'application du modèle hydrologique utilisé ou des problèmes dans la qualité des données (typiquement les chroniques pluie-débit). Ainsi, ces aspects peuvent être très utiles dans la critique des résultats du modèle et l'évaluation des incertitudes sur ses sorties.

2.5.1. Les facteurs limitant l'application des modèles

Les paramètres des modèles que nous nous proposons de tester sont calés d'après des séries de données, censées fournir au modèle l'information suffisante pour permettre d'identifier les caractéristiques du comportement du bassin. La validité de l'utilisation du jeu de paramètres obtenus, en dehors de la période de calage, repose sur une hypothèse de stationnarité du comportement du bassin. Or cette dernière est conditionnée par une évolution non-significative des conditions climatiques (nous entendons ici l'absence de changement climatique à long terme et non des variations interannuelles) et du contexte physique du bassin.

Le test des modèles en mode calage/contrôle que nous mettons en œuvre doit normalement être mené dans ce cadre de stationnarité. L'absence de ces conditions place le modèle dans une situation où l'utilisateur lui demande de donner des réponses qu'il n'est pas censé pouvoir fournir. Idéalement, il faudrait donc que chacun des 429 bassins de notre échantillon vérifie ces hypothèses. Cette situation n'est évidemment pas vérifiée ici.

La stationnarité des conditions climatiques est souvent difficile à évaluer car elle n'est observable que sur le long terme (typiquement plusieurs décennies), et il semble probable que certains des bassins utilisés sont soumis à une évolution régionale du climat. A titre d'exemple, nous ne citerons que les bassins ivoiriens, situés dans une zone soumise depuis 25 ans à une diminution progressive des précipitations (Servat *et al.*, 1996). Ces évolutions placent les modèles en situation défavorable dans la mesure où elles peuvent modifier le comportement du bassin. Ainsi, l'élévation de la température, en provoquant une sécheresse, peut induire une modification du couvert végétal ou des caractéristiques des sols, et donc modifier le comportement hydrologique.

Ces modifications des conditions physiques du bassin, induites ou non par des changements climatiques, sont cependant généralement lentes (à l'exception de 'catastrophes' telles que des incendies par exemple) en comparaison aux modifications induites par les activités humaines. Ainsi, l'implantation d'un réservoir sur une rivière, d'ouvrages de dérivation d'une partie du débit pour des besoins en irrigation, des pompages pour subvenir aux besoins d'adduction en eau potable, des changements d'occupation des sols ou de pratiques agricoles (remembrement, changement de couvert végétal, drainage, imperméabilisation de surfaces), peuvent très rapidement avoir une influence significative voire majeure sur le comportement hydrologique du bassin. Les bassins non soumis à de telles influences sont de plus en plus rares. Notre échantillon compte des bassins soumis à des modifications de ce genre. A titre d'exemple, nous pouvons citer le bassin de la Mountain Fork à Eagletown, Oklahoma, aux Etats-Unis (bassin US339000, MOPEX), dont les débits sont régularisés depuis 1966. La Figure 2.9 montre de façon nette l'influence de cet ouvrage sur la réponse du bassin, avec écrêtement des crues. Il est évident qu'il sera très difficile voire impossible pour le modèle de rendre compte du comportement artificialisé du bassin observé après cette date.

Les procédures de renaturalisation des débits peuvent permettre dans certains cas de réévaluer les débits qui existeraient en l'absence de l'ouvrage ou des pompages influençant le régime, mais cela suppose que les données nécessaires (par exemple volumes prélevés dans le cas de pompages) soient disponibles, ce qui n'était pas le cas ici.

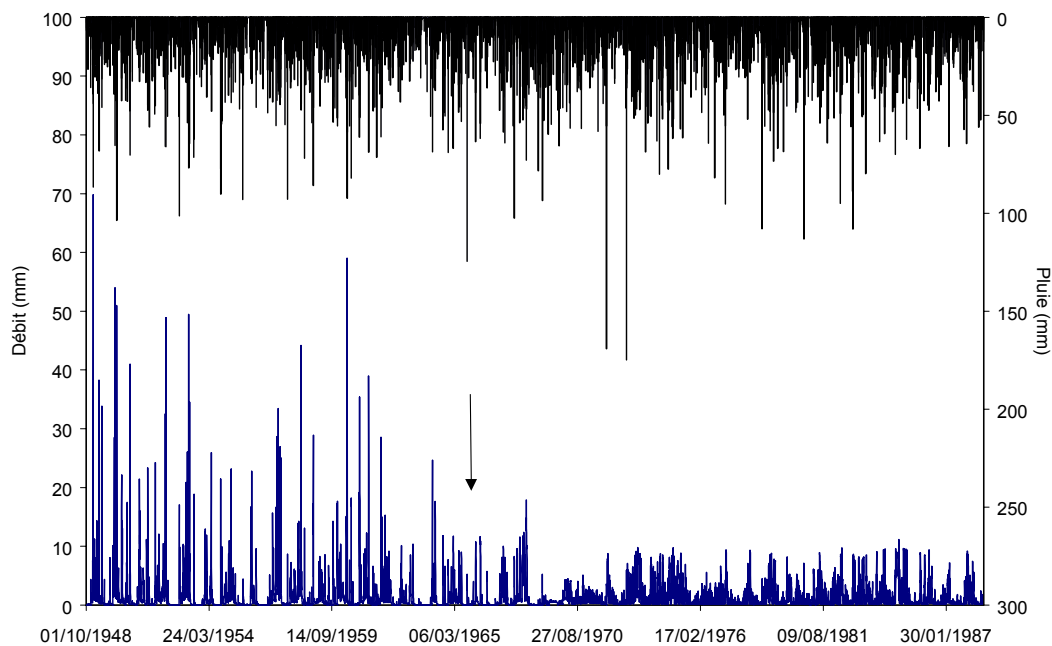


Figure 2.9: Chroniques pluie-débit pour le bassin de la Mountain Fork à Eagletown, Oklahoma, aux Etats-Unis (bassin US339000, MOPEX)

2.5.2. Qualité des données

Les données représentent un élément majeur du processus de modélisation. Leur qualité peut influencer sensiblement les valeurs calées des paramètres et donc les sorties du modèle. Elle dépend généralement de plusieurs facteurs d'ordre métrologique ou méthodologique. Au niveau des mesures, des erreurs d'estimation dans les variables utilisées en entrée du modèle ou en sortie (pour le calage) sont toujours présentes. Ainsi, l'estimation des débits est sujette à de nombreuses incertitudes, parmi lesquelles on trouve des erreurs de précision des instruments de mesure, des erreurs de la courbe de tarage, des problèmes de précision ou de robustesse dans la mesure des débits extrêmes (débits d'étiage, débits de crue), etc.. L'estimation de la pluie est elle aussi difficile avec les problèmes de captation des pluviomètres par exemple. L'estimation de l'ETP, variable non mesurée, mais calculée par un modèle, est également source d'erreurs. Enfin, le traitement de ces données, par les moyennes spatiales et/ou temporelles effectuées, introduit également des erreurs sur l'estimation des valeurs 'réelles'.

Dans notre échantillon, ces problèmes sur les données sont évidemment présents. Nous avons pu observer par exemple des débits constants en étiage sur de longues périodes, des écrêtements systématiques des crues à une valeur précise pouvant indiquer l'existence d'une plaine d'inondation, des débits évoluant parfaitement linéairement entre deux dates semblant indiquer un comblement de lacune par interpolation ou encore des bassins où les débits annuels sont plus grands que les pluies annuelles (voir Figure 2.8).

2.5.3. Quelle attitude adopter dans notre étude ?

Apporter un éclaircissement sur tous les points précédents relatifs à la validité de l'hypothèse de stationnarité et la qualité des données représente un travail colossal dans le cadre de notre échantillon de 429 bassins versants. Cela supposerait, en effet, de mener une démarche systématique d'investigation pour connaître l'historique du bassin sur la période de disponibilité des données et les conditions de mesure (avec les éventuels problèmes rencontrés). Supposons désormais que toutes ces informations aient pu être collectées. Deux attitudes peuvent être adoptées. La première consiste en une critique basée sur le jugement du modélisateur. Cette démarche est cependant fortement conditionnée par la subjectivité de la personne qui juge les données. Par ailleurs, elle conduit inévitablement à éliminer les bassins ou données présentant des problèmes flagrants et à laisser de côté les problèmes plus insidieux. Une seconde attitude est de mettre en place une méthode objective (*i.e.* indépendante de la démarche de modélisation) de test des bassins et des données pour estimer leur 'validité', ainsi que des seuils d'acceptabilité des incertitudes. Or, à notre connaissance, il n'existe pas à l'heure actuelle de tels tests qui puissent être appliqués de façon systématique. D'ailleurs, la variété des sources d'erreurs sur les données est telle qu'elles sont souvent considérées comme des phénomènes aléatoires, qui sont neutres vis-à-vis de la modélisation pluie-débit.

La mise en œuvre d'une critique des données paraît donc délicate dans notre étude, et à la limite peu souhaitable compte tenu de l'absence de protocole rigoureux et d'informations suffisantes. Par ailleurs, il faut se rappeler que nous réalisons une étude comparative. Il est probable que l'existence de problème dans les données ou de non-stationnarité des caractéristiques du bassin affecteront de la même manière les différents modèles. Éliminer les bassins où il semble y avoir des problèmes n'aurait pour effet que de rehausser le niveau moyen des performances, sans changer vraisemblablement les conclusions de la comparaison, tout en risquant d'introduire un biais du fait que ce tri serait inspiré par un raisonnement

pluie-débit. Nous avons donc décidé de conserver dans notre échantillon l'ensemble des bassins versants que nous avons pu collecter, sans effectuer la moindre sélection a priori.

Néanmoins, pour vérifier l'hypothèse précédente de la non-influence de l'absence de critique de données sur les résultats de la comparaison, nous avons essayé d'identifier les bassins pour lesquels il semble y avoir des problèmes dans les données. D'après les remarques précédentes, notre approche n'a pu être que subjective. Elle s'est appuyée sur l'observation des graphiques des chroniques de données, comme celui de la Figure 2.9, et dans le cas des données américaines de la base MOPEX sur les informations disponibles. L'observation des séries de pluie ne permet pas de critique pertinente, du fait de la nature chaotique des données, avec très peu d'autocorrélation dans le temps contrairement aux débits. C'est donc uniquement sur ces derniers que nous avons porté un jugement. Ce sont finalement 42 bassins¹, soit environ 10 % de l'échantillon, qui ont été jugés litigieux et peu propices à une modélisation pluie-débit de qualité. Cette sélection est grossière et n'a permis d'identifier que les problèmes paraissant les plus flagrants. Nous verrons au chapitre 5 l'effet de la non-prise en compte de ces données litigieuses sur les résultats.

2.6. Identification de périodes pour tester les modèles

2.6.1. Choix des périodes

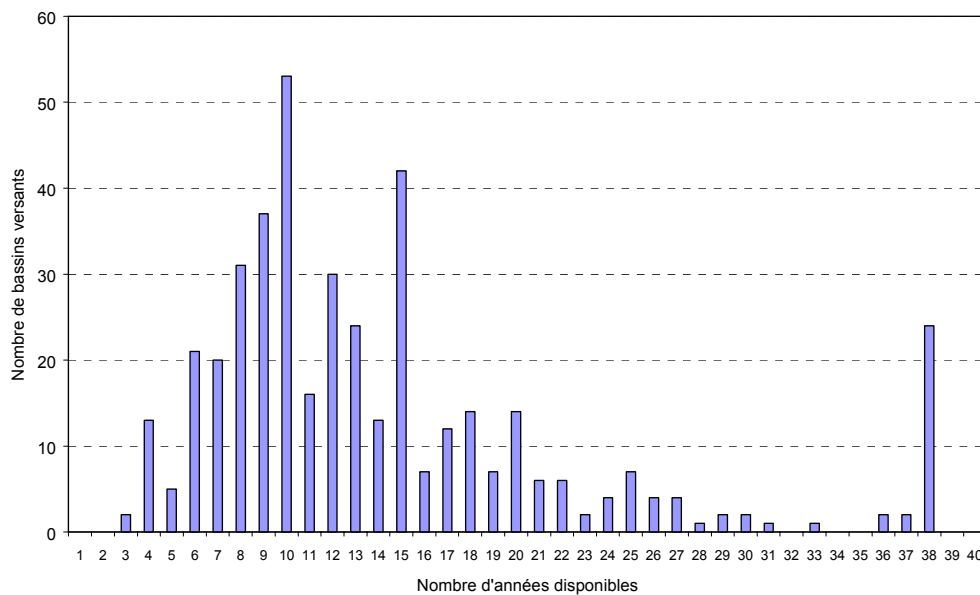


Figure 2.10: Répartition des bassins en fonction de la longueur des séries disponibles

Nous avons vu au chapitre précédent que nous adoptons pour le test des modèles une procédure en calage-contrôle ('split sample test') sur des périodes indépendantes. Cela

¹ Les bassins présentant des séries de débits jugées problématiques en vue de l'application des modèles pluie-débit sont les suivants: A6621210, A9013050, E4035710, H1932010, H3613010, H3613020, H4232040, H4243010, H5213310, H9923010, K2714010, K3206010, L0615810, O7692510, U1004010, U1025010, Y1355410, Y8624010, Y9215010, Y9605220, Y9605230, US141200, US183000, US340000, US339000, US314500, US232500, US222500, US221000, ARS63004, ARS68013, ARS69012, ARS69013, ARS69016, ARS71002, ARS71003, AU412093, AU509503, AU915001, CI150400, CI160120, CI250130.

suppose donc que nous découpons la série de données disponibles en sous-périodes sur chaque bassin. Sur l'échantillon des 429 bassins versants que nous avons rassemblés, les longueurs des séries de données sont très variables, comme l'illustre la Figure 2.10, avec des séries allant de 3 à 38 années.

Le découpage en seulement deux sous-périodes aurait été peu intéressant pour les bassins où nous disposons de longues séries de données. Nous avons préféré exploiter ces plus longues séries en y distinguant jusqu'à six sous-périodes indépendantes. Ce sont donc entre deux et six sous-périodes qui ont permis de tester les structures sur chaque bassin.

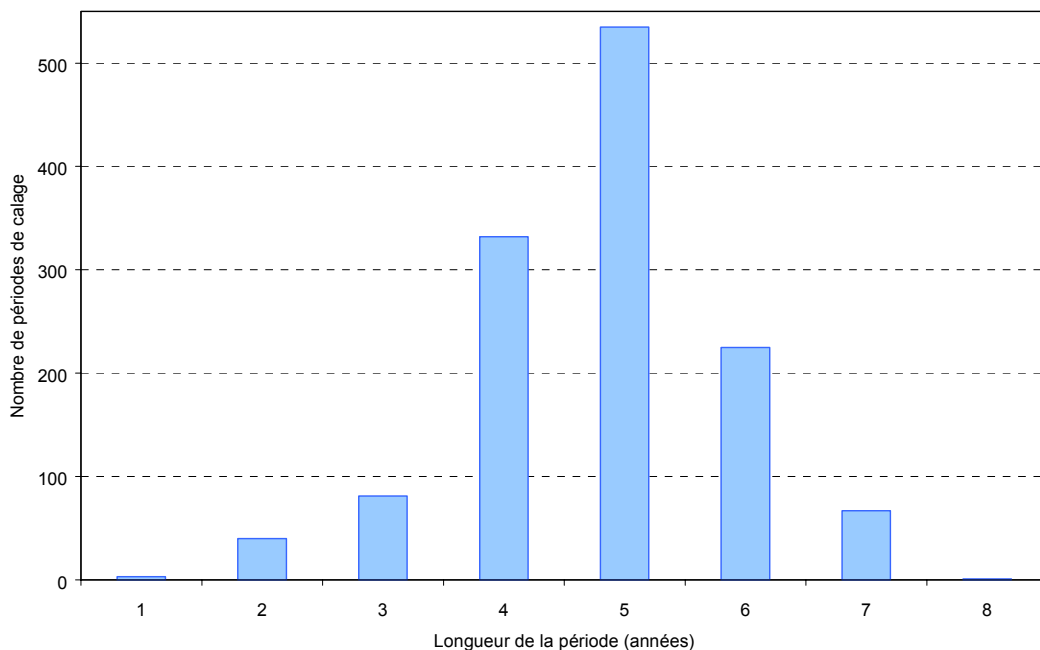


Figure 2.11: Répartition des différentes périodes en fonction de leur longueur

Ces sous-périodes sont constituées d'années calendaires. Leur longueur est comprise entre une et huit années, avec une majorité de sous-périodes de quatre, cinq et six années (voir Figure 2.11). Si les périodes les plus longues incluent généralement une bonne variabilité d'événements hydrologiques et sont assez représentatives de la gamme de conditions pouvant se produire sur le bassin, les périodes les plus courtes en revanche peuvent être très dépendantes de conditions sèches ou humides affectant la période considérée. Ces caractéristiques sont susceptibles d'avoir un impact non négligeable lors du calage des paramètres. Si les périodes longues permettent l'obtention de paramètres généralement plus robustes, le fait de tester le modèle dans des conditions contrastées, peut-être même évolutives, un peu à la manière du 'differential split-sample test' préconisé par Klemeš (1986), permet de mieux mettre en valeur les qualités de robustesse du modèle.

Pays	Australie	Brésil	Côte-d'Ivoire	Etats-Unis	France	Total
Nombre de bassins	26 (6.1)	4 (0.9)	10 (2.3)	82 (19.1)	307 (71.6)	429
Nombre de périodes de calage	63 (4.9)	12 (0.9)	41 (3.2)	312 (24.3)	856 (66.7)	1284
Nombre de tests en contrôle	98 (3.1)	24 (0.7)	156 (4.9)	1146 (35.8)	1780 (55.5)	3204

Tableau 2.1: Répartition des bassins, périodes de calage et tests en contrôle en fonction des différents pays d'origine (entre parenthèses, proportion du total en %)

Au total, ce sont finalement 1284 périodes de calage qui ont été identifiées, soit en moyenne trois périodes par bassin. Pour chaque bassin, on cale successivement le modèle sur chaque période et on effectue les tests en contrôle sur toutes les périodes restantes. Cela correspond à un total de 3204 tests en contrôle. La répartition des bassins, périodes et tests en contrôle est résumée dans le Tableau 2.1 en fonction des pays d'origine. On constate la prédominance des bassins français, avec deux tiers des périodes de calage et plus de la moitié des tests en contrôle. Les résultats des tests seront donc assez fortement influencés par le contexte climatique français. Les bassins étrangers fournissent cependant un complément très appréciable, avec 44 % des contrôles et des conditions climatiques non rencontrées en France.

2.6.2. Initialisation des modèles

La qualité des simulations des modèles dépend en partie de la détermination des inconnues du système. Si les paramètres sont les inconnues les plus importantes (déterminées lors de la phase de calage), les états initiaux du système sont également des inconnues qui peuvent jouer un rôle important sur les simulations. On comprend aisément que les modules de production et de routage du modèle ne réussissent pas à donner des simulations de débit satisfaisantes si les niveaux des réservoirs dont ils dépendent prennent des valeurs complètement erronées. Si cette influence de mauvaises conditions initiales peut avoir lieu en phase de simulation, elle peut également poser problème en phase d'optimisation: à ce stade, la procédure de calage peut en effet essayer de compenser les erreurs des états initiaux en jouant sur la valeur des paramètres. Ces derniers n'ont alors plus de réelle consistance.

Le choix d'une bonne initialisation peut donc être déterminant dans la démarche de simulation des débits. En fait, la persistance des problèmes sur les pas de temps qui suivent l'initialisation sera fonction de la mémoire du système liée à la valeur des paramètres et des caractéristiques climatiques de la période. L'avantage des modèles à réservoirs est que les états du système sont généralement bornés, et tendent à rejoindre une position nulle en l'absence d'excitations (c'est-à-dire sans entrées d'eau). Les erreurs sur les états initiaux du système tendent donc à s'atténuer avec le temps. Ainsi, une mauvaise initialisation peut être totalement gommée en l'espace de quelques dizaines de pas-de-temps.

Au cas par cas, ce problème est généralement facilement décelable à partir de la comparaison graphique des hydrogrammes simulés ou observés, comme nous avons pu l'observer dans un travail antérieur de test de modèles (Perrin, 1997). L'optimisation des états initiaux lors de la phase de calage n'est pas souhaitable car elle complique l'optimisation par ajout de paramètres. Nous avons donc préféré réaliser l'initialisation des systèmes en choisissant des états a priori, dépendants de la saison et/ou de conditions moyennes sur le bassin (le module des débits peut être utilisé pour initialiser des réservoirs de routage). Cette initialisation, utilisée seule, est bien trop grossière pour pouvoir être fiable. Nous avons donc utilisé au début de chaque période une année complète de mise en route, qui permet d'atténuer ou d'effacer totalement les effets de cette première initialisation grossière. Elle permet au modèle de rejoindre un fonctionnement normal. Au cours de cette année, les résultats du modèle ne sont pas pris en compte pour le calcul de la fonction objectif et des critères de qualité.

Ce type d'initialisation a déjà été utilisé par Chiew et McMahon (1994) et Edijatno *et al.* (1999). Elle est apparue satisfaisante dans la majorité des cas. La question de l'influence de l'initialisation sur les résultats est brièvement abordée dans l'Annexe 5 qui se rapporte aux travaux du chapitre 6. La Figure 2.12 illustre la façon dont les séries de données ont été divisées, avec la période de mise en route, dans le cas d'une série de 37 ans divisée en quatre sous-périodes.

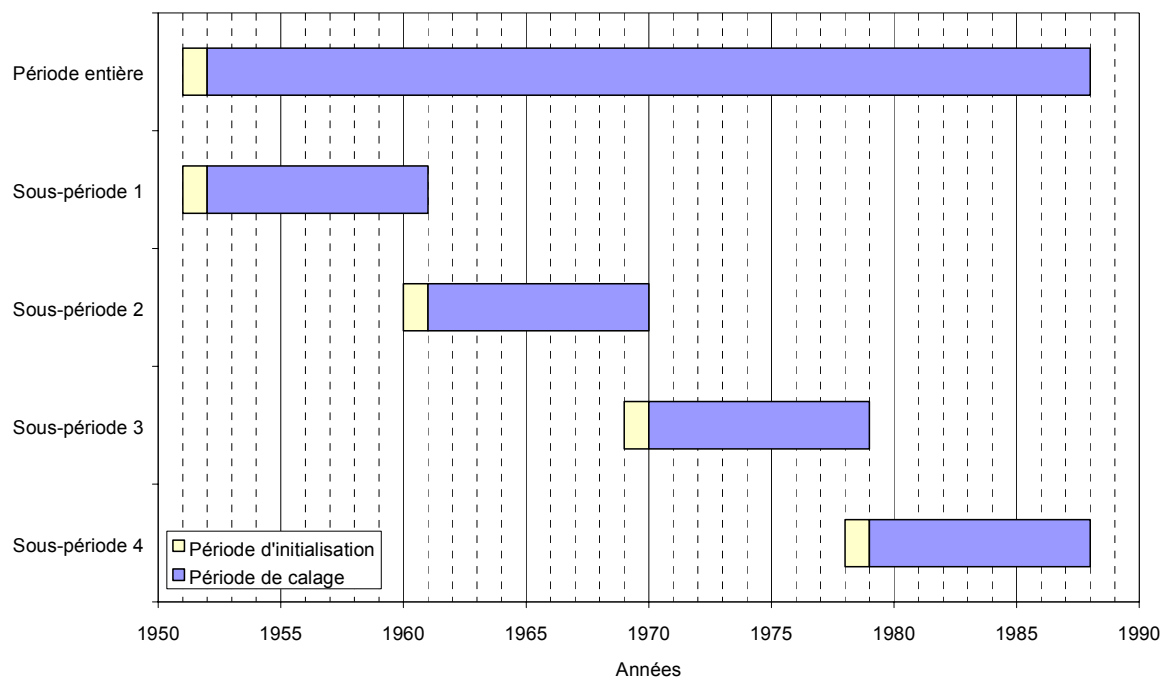


Figure 2.12: Exemple de division de la période d'étude totale en quatre sous-périodes de 9 ans, avec une année de mise en route

2.7. Classification des bassins en fonction de descripteurs hydro-climatiques

2.7.1. Introduction

Nous n'avons eu, jusqu'à présent, qu'un rapide aperçu des caractéristiques des bassins. Le but de cette partie est d'essayer d'identifier des classes de bassins versants regroupés en fonction des similarités de leurs caractéristiques hydro-climatiques. Cette démarche a pour double intérêt de voir s'il existe des ressemblances de comportement entre les bassins de notre échantillon, et s'il est possible d'envisager de faire une correspondance entre types de comportement et structures de modèle, en fonction des performances de ces dernières. Ce deuxième point sera examiné au chapitre 5.

La caractérisation et la typologie des régimes d'écoulement ont été largement étudiées par de nombreux auteurs. Pardé (1933) fut l'un des précurseurs de la classification de régimes très variés à travers le monde. Plus récemment, cette question a fait l'objet d'importants travaux, notamment dans le cadre du programme de recherche FRIEND (Arnell *et al.*, 1997).

Diverses approches ont été proposées pour l'identification de régions homogènes ou le regroupement de bassins en fonction de similarités de caractéristiques ou de comportement. Juncker (1971) considère, dans une approche très globale, l'existence ou non de sept composants du cycle hydrologique (précipitation, végétation, sol, évaporation potentielle, pente, perméabilité et ruissellement de surface), dans une procédure en tout ou rien, pour distinguer 48 comportements possibles qu'il a cartographiés à une échelle planétaire. Cette approche reste cependant par nature très géographique, en définissant des régions homogènes.

D'autres auteurs ont adopté une approche à des échelles plus fines, en explorant les similarités entre bassins. Nathan et McMahon (1990) proposent, sur un échantillon de 184 bassins en Australie, plusieurs techniques d'agrégation des bassins à partir de leurs caractéristiques physiques et de divers critères de similarité. Burn (1997) et Burn et Goel (2000), quant à eux, constituent des regroupements de bassins en appliquant des tests d'homogénéité sur des groupes de bassins à partir de variables telles que les quantiles de crue ou des mesures de saisonnalité des régimes. Que ce soit à partir de variables physiques ou hydrologiques, la formation de groupes de bassins reste cependant très sensible aux variables descriptives choisies et à la technique de rapprochement utilisée (Nathan et McMahon, 1990).

Notre démarche s'inscrit, en quelque sorte, dans la lignée de ces derniers travaux. Nous avons voulu effectuer des regroupements de bassins en utilisant des similarités de leurs caractéristiques. Nous avons néanmoins appliqué une technique beaucoup plus simple que celles utilisées par les auteurs cités précédemment. Nous nous sommes contentés d'opérer une classification hiérarchique ascendante sur notre échantillon de bassins. Cette technique permet d'agréger progressivement les bassins entre eux, en groupant les bassins les plus proches selon une distance euclidienne calculée dans l'espace des variables explicatives. Si l'on considère que chaque bassin a un poids, cette technique consiste en fait à répartir l'inertie totale de l'échantillon de manière à minimiser l'inertie au sein des groupes (individus proches) et à maximiser l'inertie entre les groupes (groupes bien dissociés et dont les centres sont éloignés). Des explications de la méthode sont données par exemple par Saporta (1990). Cette opération a été réalisée à l'aide du logiciel de statistiques STATlab.

2.7.2. Choix de variables explicatives

Pour caractériser le fonctionnement des bassins, nous avons choisi ici des caractéristiques climatiques et hydrologiques simples, calculées à partir des séries de données utilisées par les modèles. C'est en fait le seul type de descripteurs dont nous avons disposé au cours de notre étude. Nous n'avons en effet pas pu réunir de descripteurs physiographiques sur l'ensemble de notre échantillon de bassins. De tels descripteurs, qui caractérisent des facteurs intervenant dans la relation pluie-débit, auraient été complémentaires de ceux utilisés ici.

Neuf variables quantitatives ont ainsi été choisies, notamment à partir des travaux de Tangara (1991), auxquelles a été ajoutée la superficie du bassin versant, seule variable descriptive physique du bassin utilisée ici.

Les dix variables retenues sont définies par:

- ETPAN: l'ETP annuelle moyenne (mm);
- PMOYAN: la pluie annuelle moyenne (mm);
- COEFP: le coefficient d'irrégularité des pluies (%), défini par:

$$I = 100 \cdot \frac{P_{mx} - P_{mn}}{\overline{P_m}} \quad \text{Eq. (2.1)}$$

où P_{mx} est la pluie mensuelle moyenne du mois le plus pluvieux, P_{mn} la pluie mensuelle moyenne du mois le moins pluvieux et $\overline{P_m}$ la pluie moyenne mensuelle;

- QMOYAN: la lame d'eau annuelle moyenne écoulée (mm);
- COEFQ: le coefficient d'irrégularité des débits (%), calculé en appliquant l'Eq. (2.1) aux débits;

- REND: le rendement du bassin (%), défini par le rapport entre lame d'eau annuelle écoulée et pluie annuelle;
- BFI: le Base Flow Index, ou indice d'écoulement de base (%), défini comme le rapport entre le volume d'écoulement de base et le volume d'écoulement total;
- NCRU: le nombre annuel moyen de pics de crues, une crue étant définie comme un événement au cours duquel le débit excède quatre fois le module des débits;
- NETI : le nombre annuel moyen de jour d'étiage, l'étiage étant défini comme un débit inférieur au quart du module;
- SURF: la superficie du bassin versant (km²).

En utilisant ces groupes de variables, nous faisons ainsi intervenir à la fois les conditions climatiques et quelques caractéristiques de la réponse du bassin. Ces dernières sont intégratrices des caractéristiques physiques des bassins. On obtient ainsi une représentation relativement bonne du contexte hydrologique de chaque bassin, ce qui est intéressant en vue d'établir une correspondance entre bassins et modèles.

Ces variables ne sont pas toutes indépendantes. Il existe des corrélations significatives entre certains de ces descripteurs, par exemple entre le débit moyen annuel et le rendement du bassin. Cette non-indépendance n'est cependant pas gênante pour notre procédure de classification. Par ailleurs, les superficies des bassins versants variant sur plusieurs ordres de grandeur (de 0,1 à 50000 km² environ), la racine carrée de la surface a été utilisée ici pour limiter l'influence de cette variable lors de la classification et éviter l'isolement artificiel des plus grands bassins.

2.7.3. Résultats de la classification

Bien que Arnell *et al.* (1997) fassent référence à neuf classes de régimes sur des bassins européens, nous nous sommes restreints ici à cinq classes de fonctionnement de bassins. Leur composition est donnée dans le Tableau 2.2 et les caractéristiques moyennes des bassins de chacune des classes sont données par le Tableau 2.3.

Les bassins de la classe 1 (essentiellement des bassins français d'une grande moitié nord de la France) sont caractérisés par:

- un indice de débit de base élevé,
- une faible variabilité intra-annuelle des régimes d'écoulement et de pluviométrie,
- un faible nombre de crues et un faible nombre de jours d'étiage.

Les bassins de la classe 2 sont proches de ceux de la classe 1, notamment sur les variables ETPAN et PMOYAN, avec cependant des écoulements plus faibles et plus irréguliers, ce qui se traduit par davantage de crues, de jours d'étiages et des rendements et BFI plus faibles. Ils sont par ailleurs de taille moins élevée que les bassins de la classe 1.

Les bassins de la classe 3 sont caractérisés par:

- une forte pluviométrie,
- des rendements élevés et des forts débits.

Les bassins de la classe 4 sont caractérisés par:

- une ETP assez élevée
- une pluviométrie assez faible, présentant en revanche une assez forte variabilité,

- des débits faibles et assez irréguliers, avec des rendements faibles et de faibles débits de base,
- des crues nombreuses et un nombre élevé de jours d'étiage.

Les bassins de la classe 5 sont caractérisés par:

- une très forte ETP et des pluies très faibles et très irrégulières,
- un très faible nombre de jours d'écoulement,
- une faible superficie de bassin,
- de faibles rendements et la quasi-absence de débit de base.

Cette description reste sommaire et ne permet pas d'établir de correspondances avec des types de milieux physiques, étant donné l'absence de descripteurs physiographiques.

Classe	Nombre de bassins	Part de l'échantillon total (%)	Dont				
			Bassins américains	Bassins australiens	Bassins brésiliens	Bassins français	Bassins ivoiriens
1	144	34	9	1	4	129	1
2	83	19	16	6	0	60	1
3	100	23	1	3	0	96	0
4	90	21	46	14	0	22	8
5	12	3	10	2	0	0	0

Tableau 2.2: Composition des classes de bassins versants

Variables		Classes de bassins				
		1	2	3	4	5
ETPAN (mm)	Moyenne	788	819	834	1292	1625
	Intervalle	633-1504	636-1323	636-1808	689-2046	1579-1738
PMOYAN (mm)	Moyenne	952	930	1404	826	379
	Intervalle	634-1579	620-1553	636-2284	300-1543	294-621
COEFP (%)	Moyenne	197	191	219	272	398
	Intervalle	144-329	146-271	150-393	164-400	299-498
QMOYAN (mm)	Moyenne	409	283	1036	156	12
	Intervalle	0,6-824	6-686	488-2043	2-557	0,1-101
COEFQ (%)	Moyenne	226	319	262	467	956
	Intervalle	39-356	225-435	50-470	245-774	717-1444
REND (%)	Moyenne	40,8	29,2	74,7	18,0	2,6
	Intervalle	0,1-70,9	0,7-68,8	33,6-1,53	0,01-66,0	0,01-22,2
BFI (%)	Moyenne	67,6	45,8	51,7	32,4	0,1
	Intervalle	34,6-98,5	19,4-70,5	12,5-95,3	1,1-74,9	0-0,9
NCRU	Moyenne	2,7	6,5	4,7	6,6	4,1
	Intervalle	0-6,9	3,3-10,8	0-8,2	1,1-16	0,8-6,7
NETI (j)	Moyenne	51,2	122,3	88,2	202,8	338,7
	Intervalle	0-143,6	27,3-215,7	0-168,1	75,3-332,4	193,4-361,3
SURF (km ²)	Médiane	250	104	60	118	25
	Intervalle	0,4-50600	1,3-6895	1,4-3420	0,1-9889	2,9-225

Tableau 2.3: Caractéristiques des cinq classes de bassins versants (en gras, valeurs maximales pour les 5 classes; en italique, valeurs minimales)

La variance inter-classes, part de l'inertie de l'échantillon existant entre les classes de la partition, n'est que de 54 %, ce qui signifie que les classes ne sont pas très nettement dissociées. Un exemple de répartition des classes est donné à la Figure 2.13 dans le cas des variables NCRU et BFI. On constate sur cette figure la faible séparation des classes de bassins.

Ce sont respectivement les variables BFI, NCRU, QMOYAN, ETP et COEFQ qui contribuent principalement à l'excentricité des classes 1 à 5. La classe 5, avec seulement 12 bassins, regroupe des bassins en conditions semi-arides, pour lesquels il n'existe un débit que quelques jours dans l'année. Ces bassins, en majorité de faible surface, ont des caractéristiques assez marginales par rapport aux autres bassins.

La faible distinction observée entre classes peut être attribuée au fait que notre échantillon regroupe un continuum de bassins: pour chacune des caractéristiques utilisées, la répartition des bassins ne fait pas apparaître de zones d'exclusion, où aucun bassin ne serait présent. Il devient alors difficile de donner à certains bassins une appartenance nette à telle ou telle classe.

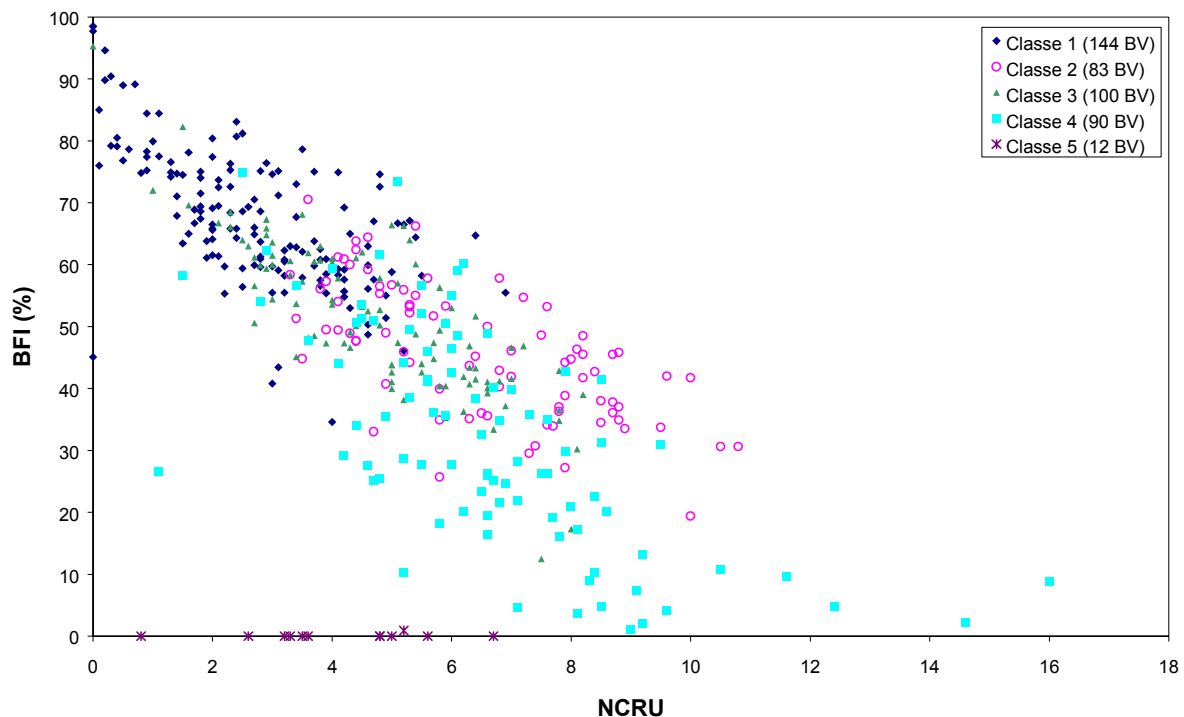


Figure 2.13: Répartition des bassins pour deux variables NCRU et BFI

La répartition proposée est par ailleurs très dépendante des variables explicatives utilisées. Ainsi, nous nous sommes aperçus qu'en enlevant certaines variables parmi les dix utilisées, les classes correspondantes avaient des compositions différentes. Nous rejoignons ici les conclusions de Nathan et McMahon (1990) qui avaient identifié le même problème dans leur étude sur des bassins du Sud-Est de l'Australie.

2.8. Conclusion

Ce chapitre nous a permis, au travers de la présentation de l'échantillon des 429 bassins versants, de souligner la diversité des conditions climatiques et hydrologiques dont nous avons disposé pour tester les structures de modèles pluie-débit. Cet échantillon est constitué de bassins situés majoritairement en France mais aussi aux Etats-Unis, en Australie, en Côte-d'Ivoire et au Brésil. Le regroupement d'un tel nombre de bassins permet de s'affranchir de façon relativement satisfaisante de la dépendance des conclusions de notre étude vis-à-vis des caractéristiques particulières d'un petit nombre de bassins, étendant de ce fait la validité de ce

travail comparatif. Par ailleurs, lors de la démarche de recherche d'amélioration des structures des modèles, cet échantillon permettra d'asseoir avec plus de certitude les modifications avancées.

Tous les bassins collectés ont été ici indistinctement conservés pour tester les modèles. En effet, les procédures d'élimination de données défectueuses ou de bassins influencés par des activités humaines sont généralement subjectives voire biaisées et difficilement applicables de façon systématique. Nous ne disposons pas non plus, et de loin, de suffisamment d'informations pour porter un regard critique complètement pertinent sur ces données. Par ailleurs, le contexte comparatif adopté dans toute notre étude limite l'intérêt de mener une critique des données, dont le résultat ne serait probablement que de rehausser le niveau global des performances des modèles, sans changer les conclusions générales de l'étude. Nous vérifierons cette hypothèse dans la suite grâce à une distinction opérée sur un petit nombre de bassins présentant des caractéristiques jugées peu propices à l'application des modèles pluie-débit.

En se basant sur des variables hydro-climatiques des bassins calculées à partir des séries de données, nous avons pu opérer une dissociation de notre échantillon en 5 classes, présentant des caractéristiques relativement contrastées. Cependant ces classes dépendent assez fortement des variables utilisées et ne sont pas identifiées très distinctement. Nous nous fonderons néanmoins sur cette classification pour essayer d'établir des correspondances entre bassins et modèles dans le chapitre 5 de présentation des résultats.

Chapitre 3

Chapitre 3

Choix et évaluation d'une méthode pour le calage des paramètres des modèles pluie-débit: la méthode 'pas-à-pas'

3.1. Introduction

Les modèles sélectionnés dans notre étude de comparaison sont des modèles conceptuels ou empiriques dont les structures dépendent de paramètres. Ces derniers permettent à l'utilisateur d'adapter le modèle aux caractéristiques particulières du bassin considéré, de la même façon que les dimensions d'un patron sont adaptées par le tailleur aux mensurations de la personne à habiller. Une partie des concepteurs de ces modèles avaient l'intention de pouvoir appliquer directement leurs modèles sur des bassins dénués de données hydrométriques (bassins non jaugés), en attribuant aux paramètres des significations physiques, les rendant ainsi quantifiables par des mesures de terrain (par exemple conductivité hydraulique des sols, épaisseur de la couche de sol, etc.). D'autres ont proposé des structures avec des paramètres non directement quantifiables, s'appuyant sur des algorithmes mathématiques pour déterminer les valeurs des paramètres, laissant de côté dans un premier temps l'application des modèles sur des bassins non jaugés.

Quelle que soit l'attitude initiale de développement adoptée, l'interprétation physique des paramètres de tels modèles reste très délicate, comme nous le verrons au chapitre 8, et par conséquent, la majorité, pour ne pas dire la totalité, de ces modèles recourent aujourd'hui à des techniques mathématiques d'estimation des paramètres. Ces dernières permettent, dans une phase d'optimisation (aussi appelée calage), de déterminer les valeurs des paramètres les plus adaptées au bassin d'étude.

L'optimisation des paramètres de systèmes non-linéaires que sont les modèles pluie-débit est une étape délicate et certains auteurs s'accordent pour dire que la qualité des paramètres d'un modèle dépend notamment de la puissance et de la robustesse de l'algorithme utilisé (Duan *et al.*, 1992). Cette phase d'optimisation n'est cependant pas un problème hydrologique proprement dit, mais plutôt un corollaire mathématique de la modélisation lié à la nature et à la complexité intrinsèque du modèle (structure et nombre de paramètres). Néanmoins, la complexité de certains modèles pluie-débit a conduit au développement de méthodes d'optimisation de plus en plus sophistiquées, destinées à pallier les défaillances de méthodes traditionnelles d'optimisation.

Dans notre étude, nous avons choisi de mettre toutes les structures des modèles testés sur un pied d'égalité, notamment dans l'effort consenti pour évaluer les valeurs des paramètres sur chaque bassin. Nous désirons ainsi appliquer une même méthode pour estimer les jeux de paramètres de tous les modèles. Le mode de calage automatique a été préféré aux techniques

manuelles ou semi-automatiques, et nous centrerons la suite de la discussion essentiellement sur ce premier type d'optimisation, plus rapide et moins subjectif.

Nous justifions dans ce chapitre le choix de la méthode d'optimisation sélectionnée en deux temps. Nous menons tout d'abord une analyse bibliographique, essentielle pour décrire quelques aspects de la problématique d'optimisation, avoir un aperçu des algorithmes couramment utilisés, et faire une synthèse des études comparatives entre diverses méthodes. On comprend mieux ainsi leurs avantages et inconvénients respectifs dans le contexte de la modélisation hydrologique. Cette réflexion orientera notre choix vers un type de méthode.

Ensuite, sans nous engager dans une évaluation exhaustive des algorithmes de calage existants, nous renforçons les raisons de notre choix en effectuant des tests d'efficacité de l'algorithme sélectionné, à partir de quatre structures de modèles de divers degrés de complexité (nombre de paramètres optimisables).

3.2. Optimisation en hydrologie: une problématique complexe

L'optimisation des paramètres d'un modèle pluie-débit a pour but de trouver le jeu de paramètres qui rapproche le plus possible le comportement du modèle de celui du bassin modélisé, la similitude des comportements étant quantifiée par un critère (fonction objectif) servant à l'optimisation des paramètres et mesurant ce degré de similitude. Le calage demande donc le choix à la fois d'un critère de qualité, d'une méthode pour identifier les paramètres, ainsi que de séries de données destinées à fournir l'information nécessaire au calage (Sorooshian et Gupta, 1985).

Cette phase de calage contribue également à compenser les erreurs internes ou imprécisions du modèle et les erreurs généralement contenues dans les données d'entrée. C'est une phase délicate, souvent d'autant plus délicate que le modèle a un nombre élevé de paramètres. Avec l'amélioration des moyens de calcul, on opte de plus en plus pour des méthodes automatiques de calage, par opposition à des techniques manuelles utilisant généralement des critères graphiques. Si ces procédures automatisées permettent généralement de gommer la subjectivité inhérente aux approches manuelles et sont souvent beaucoup plus rapides que ces dernières, elles sont en revanche confrontées à toute une série de problèmes numériques qui compliquent la phase d'optimisation.

Johnston et Pilgrim (1976) et Duan *et al.* (1992) ont répertorié certains de ces problèmes. Il s'agit notamment de:

- l'interdépendance entre paramètres: le changement de la valeur d'un paramètre peut être compensé par la modification d'autres paramètres, ceci entraînant des solutions équivalentes en terme de valeur du critère d'optimisation,
- la faible sensibilité de la fonction objectif à la modification de certains paramètres: d'importantes modifications de la valeur d'un paramètre n'entraînent que de faibles modifications de la fonction objectif, ce qui correspond à des régions plates de la surface de réponse,
- la présence de points de l'espace des paramètres où la fonction critère n'est pas différentiable (du fait par exemple de la présence de seuils dans le fonctionnement de certains réservoirs), et la structure non convexe de la surface de réponse,
- la présence d'optima locaux ou également l'existence de plusieurs zones de convergence.

Ibbitt et O'Donnell (1971a) donnent des illustrations très claires de ces problèmes au niveau des caractéristiques de la surface de réponse.

Cet ensemble de facteurs conduit souvent à une situation où l'optimum global est difficilement observable, ou même à la non-unicité d'un tel optimum (Sorooshian et Gupta, 1983). Sorooshian et Gupta (1985) définissent ainsi la notion d'identifiabilité des jeux de paramètres: la structure d'un modèle M paramétrée par θ est globalement identifiable si et seulement si différentes valeurs de paramètres donnent lieu à différentes sorties du modèle. Compte tenu de la multitude des problèmes énoncés précédemment, cette équivalence est rarement vérifiable en modélisation hydrologique. Il devient alors vain d'essayer de trouver un quelconque optimum global, quelles que soient les performances de la méthode d'optimisation employée.

Par ailleurs, le jeu de paramètres qui caractérise le comportement du bassin devrait être une spécification intrinsèque du bassin. Le bassin réel donne en effet une réponse unique (chronique de débits) à une série de pluie indépendamment de la façon dont un observateur le perçoit. On devrait donc s'attendre à ce que les jeux de paramètres caractérisant le bassin ne varient guère si l'on change de fonction objectif, de technique d'optimisation ou de série de calage. Or, essentiellement du fait du caractère beaucoup trop approximatif du modèle et des problèmes numériques d'identifiabilité, cette indépendance n'est vérifiée pour aucun de ces points. Les paramètres obtenus restent dépendants à la fois du choix de la fonction objectif (Sorooshian, 1981; Sefe et Boughton, 1982; Ibbitt et Hutchinson, 1984), de la méthode d'optimisation (Duan *et al.*, 1992; Gan et Biftu, 1996) ou des séries de données utilisées (Gupta et Sorooshian, 1983; Sorooshian *et al.*, 1983; Allred et Haan, 1991; Yapo *et al.*, 1996).

De nombreux auteurs s'accordent donc pour dire que l'estimation des paramètres de modèles hydrologiques doit passer par une réflexion profonde sur les problèmes internes de ces modèles qui limitent l'efficacité des méthodologies de calage (Pickup, 1977; Sorooshian, 1985; Hendrickson *et al.*, 1988; Sorooshian *et al.*, 1993; Gan et Biftu, 1996). Pickup (1977) et Gupta et Sorooshian (1983) montrent par exemple que l'étude fine des caractéristiques structurelles du modèle et la reparamétrisation judicieuse de certaines de ses composantes peuvent améliorer l'observabilité de l'optimum et la facilité de calage. Une attention particulière doit également être donnée à la qualité des séries employées pour caler les modèles. Celles-ci doivent être considérées davantage en terme de représentativité et de qualité d'information qu'en termes quantitatifs de longueur de séries de calage.

Les difficultés d'optimiser des paramètres de modèles hydrologiques cachent donc des problèmes liés à la définition, l'existence et l'identifiabilité d'un jeu optimal de paramètres pour le modèle, ceci provenant principalement de caractéristiques internes au modèle et secondairement des questions d'erreurs de données. Les qualités qui peuvent alors être exigées d'une méthode de calage semblent donc devoir se limiter à la capacité à fournir des paramètres 'conceptuellement acceptables' et stables (plutôt que d'hypothétiques optima mathématiques), le degré de 'réalisme' des paramètres ainsi identifiés pouvant être évalué par le niveau de performance du modèle au stade de la validation. Cependant, la première réponse apportée aujourd'hui aux problèmes d'optimisation a été une sophistication des méthodes de calage, bien avant une réflexion approfondie sur les remèdes à apporter aux problèmes structurels des modèles. Dans la suite, nous présentons quelques-unes de ces techniques.

3.3. Quelques méthodes de calage

Pour tenter de remédier à certains des problèmes mentionnés précédemment, beaucoup de recherches se sont concentrées durant les trente dernières années sur l'amélioration des algorithmes d'optimisation. Le développement progressif des moyens de calcul a permis l'émergence de nouvelles méthodes automatiques. D'importants travaux ont été notamment

réalisés aux Etats-Unis sur le modèle Sacramento du National Weather Service River Forecast System (voir par exemple Brazil et Hudlow, 1980; Brazil et Krajewski, 1987; Sorooshian *et al.*, 1993).

En se basant sur la synthèse présentée par Sorooshian et Gupta (1995), nous exposons dans les paragraphes suivants une classification et un aperçu des méthodes d'optimisation applicables aujourd'hui en hydrologie, pour mieux comprendre les différences de stratégies entre ces méthodes. La discussion porte ici sur les méthodes permettant de fournir un seul jeu de paramètres optimisés. Elle ne considère pas celles construites dans un but d'une quantification des incertitudes, telles que les méthodes de type Monte-Carlo qui fournissent des distributions de probabilités sur les paramètres (par exemple la méthode GLUE de Beven et Binley, 1992, ou de façon plus sophistiquée, l'association méthodes Monte-Carlo – chaîne de Markov telle que l'algorithme Métropolis proposé par Kuczera et Parent, 1998).

Les différentes stratégies spatiales de recherche d'un optimum dans l'espace des paramètres permettent de répartir les algorithmes en deux grandes catégories, avec d'un côté les méthodes locales et de l'autre les méthodes globales.

3.3.1. Les méthodes locales

Les méthodes locales adoptent une stratégie itérative dans laquelle, partant d'un point de l'espace des paramètres, on se déplace dans une direction qui améliore continûment la valeur de la fonction critère, jusqu'à ne plus pouvoir générer d'amélioration. Le jeu de paramètres trouvé correspond alors à l'optimum de la fonction. Au sein des méthodes locales, on peut distinguer deux sous-groupes, les méthodes directes et les méthodes de gradient.

Les méthodes directes utilisent comme information au cours de l'optimisation la seule valeur de la fonction critère aux différents points testés de l'espace des paramètres. A partir d'un jeu initial de paramètres, la méthode choisit un pas de recherche et une direction pour effectuer des déplacements dans l'espace des paramètres, et calculer la valeur de la fonction au nouveau point. S'il y a amélioration, l'opération est renouvelée à partir de ces nouveaux paramètres. Sinon, on choisit une nouvelle direction et/ou un nouveau pas de recherche à partir de ce même point.

L'originalité d'une méthode locale directe réside dans la stratégie du choix de la direction et du pas de recherche. De nombreuses méthodes existent. Nous citerons à titre d'exemple la méthode de Rosenbrock (1960), la méthode du Pattern Search ou PS (Hooke et Jeeves, 1961) ou encore la méthode du Simplex (Nelder et Mead, 1965) qui est l'une des plus efficaces et des plus utilisées. La méthode 'pas-à-pas' (Michel, 1989; Nascimento, 1995), utilisée dans cette étude, fait également partie de cette catégorie.

Les méthodes de gradient utilisent comme information à la fois la valeur de la fonction critère et celle du gradient de la fonction pour décider de la stratégie d'évolution dans l'espace des paramètres. Dans le cas où ces méthodes utilisent le Hessien de la fonction, c'est-à-dire la matrice des dérivées partielles du second ordre de la fonction par rapport aux paramètres, ce sont alors des méthodes dites de Newton. La difficulté d'évaluer le Hessien incite parfois à en utiliser des approximations à partir de dérivées du premier ordre. Les méthodes sont alors dites de quasi-Newton. On peut citer notamment l'algorithme de Davidon-Fletcher-Powell (Fletcher et Powell, 1963) ou encore la méthode de Levenberg-Marquardt (Levenberg, 1944; Marquardt, 1963), version modifiée de celle de Gauss-Newton.

Théoriquement, dans un problème idéal d'optimisation où la surface de réponse est convexe (fonction uni-modale présentant une seule zone de convergence), l'optimum est atteint avec une méthode locale quel que soit le point de départ choisi pour les paramètres. Les méthodes

locales ont été initialement conçues pour traiter de tels cas d'optimisation. Or le problème est généralement non convexe dans le cas des modèles hydrologiques, avec des fonctions multi-modales.

3.3.2. Les méthodes globales

Les méthodes globales sont conçues pour pouvoir traiter efficacement des problèmes où la fonction à optimiser est multi-modale. Les méthodes globales, par opposition aux locales, explorent une partie beaucoup plus grande de l'espace des paramètres. Elles adoptent diverses stratégies d'exploration, que l'on peut classer en trois catégories: déterministes, stochastiques ou combinatoires.

Les méthodes déterministes pratiquent une discrétisation systématique de l'espace des paramètres en le découpant en mailles. Pour chaque nœud du maillage, on calcule la valeur de la fonction objectif, et il est possible, avec une discrétisation suffisamment fine, de localiser l'optimum global de la fonction. Ces méthodes permettent également d'apporter des renseignements sur la structure de la surface de réponse et de localiser les optima secondaires. Elles n'avancent pas d'hypothèses a priori sur la forme de l'hypersurface ou la nature des résidus du modèle. Cependant, pour qu'elles soient efficaces, ces méthodes nécessitent un maillage fin de l'espace des paramètres, ce qui entraîne des temps de calcul à l'heure actuelle prohibitifs. Elles sont donc rarement utilisées, si ce n'est dans des études détaillées sur la structure des surfaces de réponse. On peut citer à titre d'exemple la méthode Exhaustive Gridding ou EG (Duan *et al.*, 1992; Nascimento, 1995).

Les méthodes stochastiques échantillonnent l'espace des paramètres de façon aléatoire. Les jeux de paramètres sont générés en utilisant des distributions de probabilité. Si on utilise, de façon assez primitive, une loi de probabilité uniforme, on donne à chaque zone de l'espace des paramètres la même probabilité de contenir l'optimum de la fonction. C'est par exemple le cas de la méthode Uniform Random Sampling ou URS (Duan *et al.*, 1992; Nascimento, 1995). Si une telle méthode n'avance pas d'hypothèses a priori sur la forme de l'hypersurface, sa faiblesse vient du fait que la recherche n'est pas guidée et que l'information obtenue à chaque calcul de la fonction objectif n'est pas utilisée pour le reste de l'optimisation. Cela oblige donc à opérer un grand nombre d'essais et la procédure devient coûteuse en calcul. D'autres méthodes, en revanche, utilisent les valeurs de la fonction obtenues au cours de l'optimisation pour adapter la distribution de probabilité utilisée pour échantillonner les paramètres. On peut accroître ainsi la probabilité de chercher l'optimum dans une zone correspondant aux meilleures valeurs de la fonction objectif de l'itération précédente. Il s'agit par exemple de la méthode Adaptative Random Sampling ou ARS (Masri *et al.*, 1976; Pronzato *et al.*, 1984; Brazil et Krajewski, 1987).

Dans la lignée des méthodes stochastiques, de nouvelles méthodes ont été proposées au cours de ces dix dernières années, faisant appel à des stratégies fondées sur des concepts d'évolution de populations de points dans des systèmes naturels. Celle du Simulated Annealing ('recuit simulé') ou SA (Kirkpatrick *et al.*, 1983; Press *et al.*, 1992) est basée sur l'analogie entre la façon dont des métaux fondus se refroidissent, et l'optimisation d'une fonction avec de nombreux degrés de libertés, que l'on veut minimiser. Dans cette méthode, on part d'une population de points de l'espace des paramètres, représentant les atomes du métal. Chaque point est caractérisé par la valeur de la fonction objectif représentant l'énergie de l'atome. D'une itération à l'autre, on fait progressivement baisser la température du système grâce à un paramètre de contrôle. A chaque itération, on fait varier aléatoirement la valeur des paramètres. Si le nouveau point obtient une fonction objectif inférieure (déplacement de l'atome vers un état d'énergie inférieure), il est conservé dans l'échantillon. Sinon, sa

conservation est fonction d'une loi de probabilité (algorithme de Metropolis *et al.*, 1953). On évolue ainsi vers un minimum d'énergie, minimum de la fonction objectif. La particularité de la méthode est de pouvoir accepter des points qui ne minimisent pas l'énergie du système, ce qui permet d'éviter d'être piégé trop facilement dans des minima locaux.

Le concept d'évolution naturelle d'une population a également été utilisé, avec un principe de sélection favorisant les individus les mieux portants (Holland, 1975). Des méthodes d'algorithmes génétiques (Wang, 1991, 1997; Franchini, 1996) ont ainsi été développées. Comme précédemment, on part d'une population de points choisis dans l'espace des paramètres. On fait évoluer cette population à la manière de l'évolution génétique d'une population biologique. La sélection des individus est faite en fonction de leurs qualités génétiques (valeur de la fonction critère), les parents les mieux portants ayant le plus de chance de donner une descendance. D'une itération à l'autre, la descendance est générée par mixage des chromosomes (qui représentent la position des points dans l'espace). On tend ainsi vers la meilleure descendance, optimum de la fonction objectif.

Partant de réflexions sur les méthodes multi-départ, Duan *et al.* (1992, 1993) ont créé, dans le même esprit que les algorithmes génétiques, la méthode Shuffled Complex Evolution - University of Arizona ou SCE-UA. Cette algorithme fait évoluer parallèlement des sous-groupes d'une population qui se mélangent périodiquement (échange d'information). On retrouve des concepts d'évolution similaires à ceux de l'algorithme génétique, avec des principes différents de génération de la descendance s'appuyant sur un schéma proche de celui du Simplex. Comme dans la méthode SA, les méthodes d'évolution génétique permettent de retenir d'une itération à la suivante des individus qui n'améliorent pas la qualité de la population totale, ce qui limite la convergence sur des optima locaux.

Les méthodes combinatoires ont été suggérées par de nombreux auteurs, pour améliorer les performances des méthodes locales (par exemple Ibbitt et O'Donnell, 1971b, et Johnston et Pilgrim, 1976). Elles utilisent généralement une méthode stochastique d'échantillonnage pour générer des jeux de paramètres qui serviront ensuite de points de départ à des méthodes locales d'optimisation. Ce sont les méthodes multidépart, qui diminuent considérablement la probabilité d'échec d'une méthode locale. On peut notamment citer l'association de la méthode locale du Simplex avec une méthode stochastique classique pour donner la méthode du Multistart Simplex ou MSX (Duan *et al.*, 1992; Sorooshian *et al.*, 1993; Nascimento, 1995; Gan et Biftu, 1996) ou encore l'association du Simplex et de l'ARS précédemment citée (Duan *et al.*, 1992).

D'autres méthodes combinatoires ont également été proposées pour affiner la recherche de méthodes globales par des méthodes locales. Ainsi Isabel et Villeneuve (1986) associent une méthode de recherche au hasard de type Monte-Carlo (Karnopp, 1963) avec la méthode du Simplex et la méthode du gradient conjugué de Fletcher-Powell. On peut également citer la combinaison du SA et du Simplex, appliquée par Summer *et al.* (1997) et Thyer *et al.* (1999) sur le modèle mSFB. Les algorithmes génétiques ont également été couplés à des méthodes locales, notamment par Franchini *et al.* (1998) qui y ont associé une méthode de programmation quadratique séquentielle (Schittowski, 1985), méthode s'apparentant à une méthode de Newton pour affiner la recherche de l'optimum.

3.4. Avantages et inconvénients des méthodes de calage: des tests comparatifs

Le paragraphe précédent a présenté les grands types de méthodes utilisées pour l'optimisation des paramètres de modèles hydrologiques. Nous avons pu voir les différences de stratégie entre elles. Ces méthodes présentent chacune avantages et inconvénients. Nous allons essayer maintenant, au travers de nombreuses études existant sur la comparaison de ces méthodes, de distinguer les forces et faiblesses des méthodes locales et globales, et d'analyser ensuite leurs efficacités relatives.

La plupart des études comparatives ont été menées sur des données synthétiques où les séries de débit sont calculées à partir de séries de pluie et d'ETP et d'un jeu connu de paramètres. On connaît donc dans un tel cas le jeu optimum des paramètres, auquel on compare le jeu de paramètres obtenus par calage. Ces études fournissent donc des résultats davantage théoriques que pratiques puisqu'elles se placent dans les cas idéaux d'absence d'erreurs de données et de modélisation. Les études comparatives menées sur données réelles sont, quant à elles, plus représentatives de cas concrets. Ne connaissant pas généralement l'optimum réel, on compare alors les résultats des différentes méthodes sur la valeur finale de la fonction objectif, et, s'il y a plusieurs optimisations dans différentes conditions, sur la stabilité des paramètres obtenus.

3.4.1. Quelques avantages et inconvénients des méthodes locales

Les méthodes de gradients sont souvent handicapées par les irrégularités de l'hypersurface (Ibbitt et O'Donnell, 1971; Johnston et Pilgrim, 1976; Pickup, 1977). La fonction critère, si elle est généralement continue, présente des dérivées discontinues ou seulement continues par morceaux. Il est alors difficile d'évaluer le Hessien dans les méthodes de type Newton. S'il est possible d'approximer le Hessien (méthode du quasi-Newton) à partir du calcul analytique des dérivées partielles dans le cas de modèles simples (Gupta et Sorooshian, 1985), ce calcul analytique reste souvent très complexe voire impossible pour la majorité des modèles. La principale raison de ces discontinuités, suggérée déjà par Ibbitt et O'Donnell (1971) et identifiée par Hendrickson *et al.* (1988), est la présence de seuils de fonctionnement dans la structure du modèle, conjuguée à son fonctionnement non-continu (pas de temps discret). Pour ces raisons, les méthodes de gradient sont généralement moins robustes et plus sensibles aux irrégularités de l'hypersurface de la fonction (ou de ses dérivées) que des méthodes directes, que ce soit sur des données réelles ou des données synthétiques (Johnston et Pilgrim, 1976; Pickup, 1977; Hendrickson *et al.*, 1988). L'introduction de courbes en 'S' pour lisser les fonctions de seuil, suggérée par Kitanidis et Bras (1980), génère des perturbations dans les dérivées et ne solutionne pas entièrement le problème.

Un problème commun à toutes les méthodes locales est le risque d'être piégé sur un optimum local. Par nature, les méthodes locales n'ont en effet pas été conçues explicitement pour pouvoir prendre en compte des optima multiples. Il semble donc que les résultats d'une optimisation par une méthode locale soient dépendants du point de départ de l'optimisation et de la stratégie d'optimisation. Ibbitt et O'Donnell (1971b) montrent pour huit méthodes locales, en utilisant des points de départ différents, que la convergence s'effectue généralement vers des points différents. La convergence globale n'est que rarement atteinte. Sur le modèle de Boughton (12 paramètres) et utilisant des données synthétiques, Pickup (1977) teste quatre méthodes locales. Aucune ne réussit à identifier le jeu de paramètres utilisé pour calculer la série synthétique des débits et chacune converge vers des paramètres différents. Dans sa comparaison des méthodes du Simplex et de Rosenbrock, (Sorooshian, 1980) n'apporte pas de conclusion très nette quant à la supériorité de l'une ou l'autre des

méthodes. Utilisant également des données synthétiques, Gupta et Sorooshian (1985) montrent, avec deux méthodes locales, la difficulté de converger vers l'optimum lorsque l'on choisit des points de départ trop éloignés de la zone de convergence globale. Ce problème de convergence est aussi observé par Tanakamaru (1995) dans le cas du modèle Tank.

Cependant, il apparaît que, si les optima globaux ne sont pas atteints, les méthodes locales fournissent généralement des résultats corrects en terme de valeur de la fonction objectif (Johnston et Pilgrim, 1976). Par ailleurs, les problèmes de convergence peuvent être attribués, dans le cas de modèles complexes, davantage à des problèmes d'identifiabilité d'un seul optimum du fait notamment de l'interdépendance entre paramètres qu'à de réelles faiblesses des méthodes locales (Pickup, 1977).

Au niveau de l'efficacité en terme de temps de calcul, les méthodes locales présentent le gros avantage d'être des méthodes rapides. Les méthodes de gradient apparaissent plus rapides que les méthodes directes (Gupta et Sorooshian, 1985; Hendrickson *et al.*, 1988).

3.4.2. Quelques avantages et inconvénients des méthodes globales

De nombreuses comparaisons incluant des méthodes globales ont été menées durant les dix dernières années, certainement du fait de l'intérêt grandissant pour ces techniques d'optimisation. Ces méthodes ne présentent pas théoriquement, dans les cas où l'optimum est identifiable, les problèmes liés aux optima secondaires mentionnés précédemment pour les méthodes locales dans l'optimisation de fonctions multi-modales. Nous pouvons, à travers les études comparatives existantes, essayer d'apprécier l'efficacité des méthodes globales.

Sur données synthétiques, il apparaît que la procédure globale la plus rapide et la plus apte à localiser l'optimum soit la méthode SCE-UA, comme l'ont montré Duan *et al.* (1992) sur le modèle SIXPAR, Sorooshian *et al.* (1993) sur le modèle Sacramento, Tanakamaru (1995) et Cooper *et al.* (1997) sur le modèle Tank ou encore Franchini *et al.* (1998) sur le modèle ADM. D'autres méthodes testées dans ces études telles que la méthode MSX (Duan *et al.*, 1992; Sorooshian *et al.*, 1993) ou les algorithmes génétiques couplés à des méthodes locales (Tanakamaru, 1995; Franchini *et al.*, 1998) donnent également de bons résultats. Les méthodes URS et ARS ne fournissent pas de bons résultats et sont prohibitives en temps de calcul (Duan *et al.*, 1992). Les algorithmes génétiques seuls (Tanakamaru, 1995; Cooper *et al.*, 1997; Franchini et Galeati, 1997) ne fournissent pas non plus de résultats satisfaisants.

Sur données réelles, la supériorité de la méthode SCE-UA par rapport à d'autres méthodes globales est moins nette. Si Thyer *et al.* (1999) montrent la supériorité de la méthode SCE-UA sur une méthode combinant les procédures du Simulated Annealing et du Simplex dans le cas du modèle de Boughton, Sorooshian *et al.* (1993) trouvent des résultats sensiblement équivalents pour la méthode SCE-UA et la méthode MSX dans le cas du modèle Sacramento. Gan et Biftu (1996) formulent des observations similaires dans le cas des modèles Sacramento, Smar, Nam et Xinanjiang. Il est intéressant de voir que dans ces études, les performances obtenues en contrôle sont également équivalentes. De la même façon, les résultats de Ndiritu et Daniell (1999) montrent que, si le perfectionnement d'une procédure d'optimisation basée sur un algorithme génétique permet d'améliorer la qualité de l'optimum, en revanche au niveau du contrôle il n'y a pas d'amélioration consécutive des performances. Franchini *et al.* (1998), quant à eux, rencontrent de gros problèmes dans l'estimation des paramètres pour le modèle ADM avec la méthode SCE-UA et deux autres méthodes combinatoires.

Dans les discussions de leurs résultats, tous ces auteurs font néanmoins mention des problèmes (cités précédemment) d'observabilité d'un optimum global pour ces modèles en

raison de leur structure complexe, du choix de la fonction objectif et des erreurs dans les données. Les problèmes d'observabilité peuvent également entraîner une instabilité des paramètres. Explorant une large part de l'espace des paramètres et cherchant un optimum numérique, on peut trouver par des méthodes globales des jeux de paramètres très différents d'une méthode à une autre et des jeux de paramètres sans réelle parenté conceptuelle, comme le rapportent Gan et Biftu (1996). Dans ces cas, le problème d'optimisation est donc couplé à une complexité inappropriée du modèle, c'est-à-dire une complexité dictée par des concepts et non pas par la recherche de l'efficacité réelle du modèle. C'est alors très probablement ce dernier qui doit être remis en cause.

Les temps de calage et la puissance de calcul nécessaires aux méthodes globales ont été les raisons majeures limitant leur développement et leur utilisation. Il a fallu attendre de disposer d'ordinateurs puissants, dans les années 80, pour les rendre utilisables. Les temps de calcul nécessaires à la plupart d'entre elles sont encore aujourd'hui des facteurs limitant leur implémentation dans un contexte opérationnel.

Enfin, il est intéressant de retenir de l'étude de Cooper *et al.* (1997) que, si les méthodes globales ont été conçues notamment pour éviter la dépendance des résultats vis-à-vis du point de départ constatée pour les algorithmes locaux, cette dépendance n'est pas inexistante pour les méthodes globales testées (SCE-UA, algorithme génétique et Simulated Annealing). La plupart des procédures globales nécessitent en effet la détermination d'une population initiale de points, que la méthode fait ensuite évoluer. Les résultats de cette étude, obtenus sur données synthétiques pour le modèle Tank (8 paramètres), tendent à montrer la dépendance des résultats vis-à-vis de cette population initiale, c'est-à-dire en fait l'incapacité des méthodes, même globales, à identifier un réel optimum.

3.4.3. Méthodes locales contre méthodes globales

Il existe peu de comparaisons incluant à la fois des méthodes locales et globales. On constate à travers ces quelques études que l'avantage théorique que devraient avoir les méthodes globales sur les locales dans l'identification d'un optimum n'est généralement pas vérifié dans la pratique.

Dans les cas théoriques sans erreurs de données (données synthétiques), Tanakamaru (1995) trouve les méthodes globales plus efficaces que les locales. En revanche, d'après les études existantes, nous n'avons pas constaté de supériorité des méthodes globales dans le cas de données réelles. Dans leur étude très complète impliquant quatre modèles pluie-débit (9 à 15 paramètres) testés sur des données réelles de huit bassins versants, Gan et Biftu (1996) rapportent que les valeurs de la fonction critère obtenues par les méthodes SCE-UA et MSX ne sont meilleures que de façon marginale par rapport à celles obtenues par la méthode du Simplex. Ils montrent par ailleurs que les performances des modèles calculées a posteriori en contrôle sont similaires et parfois meilleures pour les jeux de paramètres obtenus par la méthode du Simplex. De la même manière, Franchini et Galeati (1997) montrent que dans le cas du modèle ADM, la méthode locale du Pattern Search donne des résultats aussi satisfaisants voire meilleurs que différents algorithmes génétiques.

Au niveau des temps de calcul, les méthodes locales sont plus rapides que les procédures globales. Par exemple, la méthode du Simplex est environ trente fois plus rapide que la méthode SCE-UA, elle-même environ trois fois plus rapide que la méthode MSX (Duan *et al.*, 1992; Gan et Biftu, 1996).

3.5. Choix d'une méthode d'optimisation

3.5.1. Vers une méthode locale

Plusieurs points ressortent de l'analyse précédente. Il existe aujourd'hui des méthodes fiables pour l'estimation de paramètres. Dans les cas théoriques (données synthétiques), les paramètres obtenus par une méthode globale permettent dans une majorité de cas un meilleur ajustement des sorties du modèle aux données de calage que par des méthodes locales. Cependant, cet avantage devient faible voire inexistant lorsque des données réelles sont utilisées.

L'avantage pouvant exister au calage pour les méthodes globales ne garantit cependant pas aux paramètres identifiés une meilleure efficacité dans la pratique, c'est-à-dire en phase de contrôle du modèle lorsque l'on sort de la période de calage. En effet, elles tendent à favoriser les performances au calage, avec des gains finalement marginaux sur les méthodes locales, au détriment de la stabilité des paramètres. Ceux-ci peuvent en effet être identifiés dans des régions très différentes de l'espace des paramètres lorsque l'on change de période. Les méthodes locales, au contraire, en partant d'un point fixe, favorisent cette stabilité, qui sera une nécessité lorsque l'on passera à l'étape de régionalisation ou d'explication des paramètres.

En fait, l'efficacité de puissantes méthodes globales telles que SCE-UA, théoriquement supérieures, est limitée par la complexité structurelle de la plupart des modèles, rendant l'optimum global généralement non-unique et peu identifiable. Les raisons profondes des problèmes d'optimisation des paramètres semblent donc résider bien davantage dans les maladresses ou les excès de paramétrisation, et surtout dans l'inadéquation des formulations mathématiques des structures de modèles.

Pour ces raisons, nous avons choisi d'utiliser dans notre travail une méthode locale d'optimisation. Ce choix est renforcé par des considérations pratiques, la taille de notre échantillon de données (1284 périodes de calage) et le nombre de modèles testés (38 structures) multipliant de façon importante le temps de calcul avec des méthodes globales. Dans la continuité des travaux de Nascimento (1995) et Edijatno *et al.* (1999), la méthode 'pas-à-pas' a été adoptée. Nous présentons cette méthode dans le paragraphe suivant.

3.5.2. La méthode 'pas-à-pas'

La méthode 'pas-à-pas' (Michel, 1989; Nascimento, 1995) a été développée à la Division Hydrologie du Cemagref d'Antony. La Figure 3.1 illustre les étapes adoptées par l'algorithme. C'est une méthode locale qui opère une optimisation (maximisation ou minimisation) d'une fonction objectif choisie par l'utilisateur (indépendamment de la méthode). Nous adopterons dans ce chapitre une maximisation du critère de Nash-Sutcliffe (1970) calculé sur les racines carrées des débits, noté ici R^2 et présenté au chapitre 4.

Supposons que l'on ait à caler un vecteur x de paramètres ayant n composantes (x_1, x_2, \dots, x_n) . On désire trouver le vecteur x^* qui maximise le critère R^2 dont la valeur dépend, par l'intermédiaire des débits calculés, des valeurs des paramètres (x_1, x_2, \dots, x_n) .

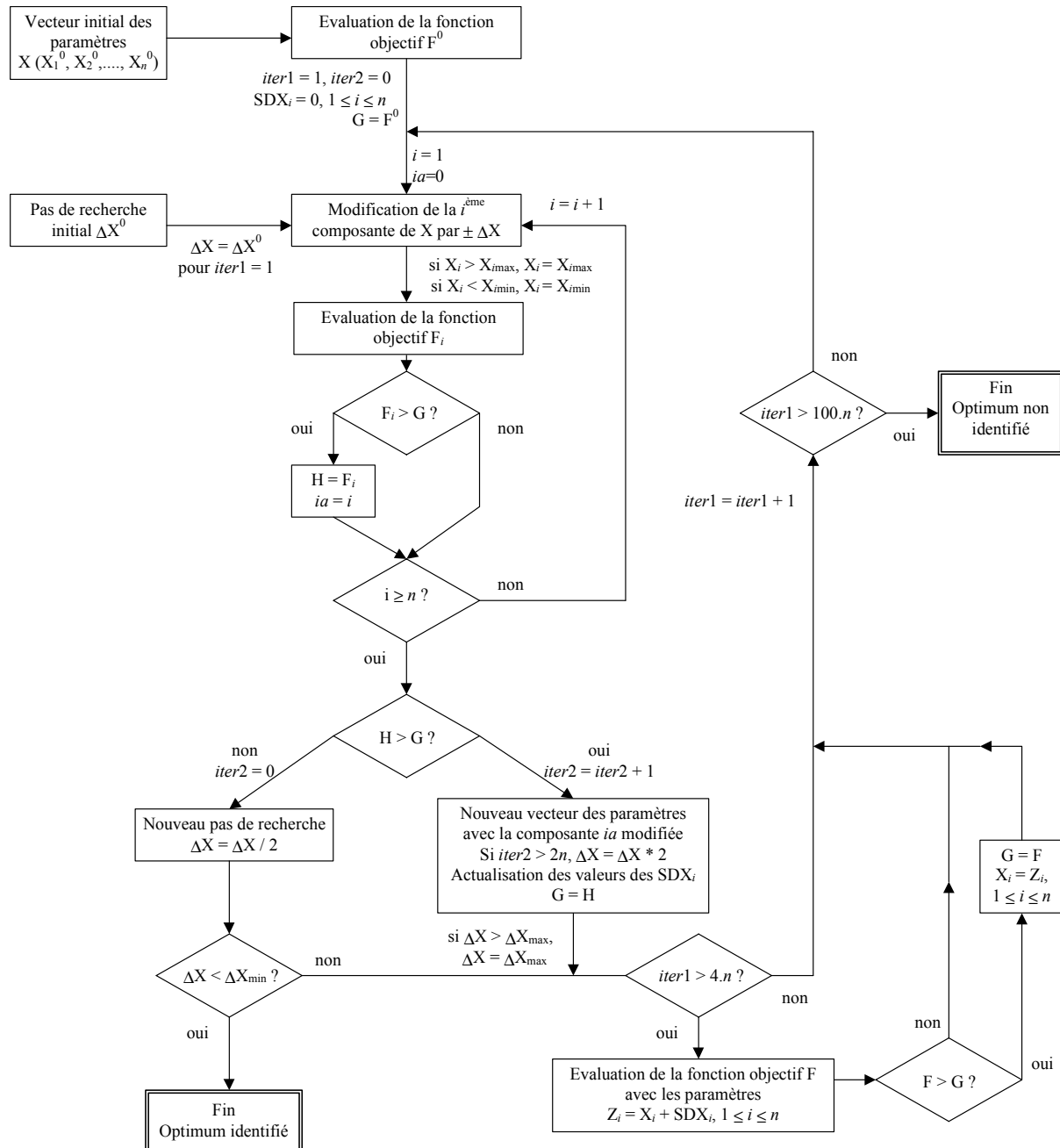


Figure 3.1: Diagramme schématique de la méthode de calage 'pas-à-pas'

Le processus d'optimisation est itératif. La méthode adopte une stratégie de déplacement, la plupart du temps, le long des axes de l'espace des paramètres, avec un pas de recherche pouvant varier d'une itération à l'autre. L'amplitude du pas de recherche étant ici la même pour tous les paramètres, des transformations mathématiques préalables (transformations logarithmiques ou puissance par exemple) peuvent être appliquées pour garantir des sensibilités grossièrement équivalentes à ce pas de recherche pour tous les paramètres. Ces transformations sur les paramètres sont choisies en tenant compte de la façon dont les paramètres interviennent dans le modèle, et sont propres à celui-ci.

La recherche démarre à partir d'un vecteur initial des paramètres x^0 , de composantes $(x_1^0, x_2^0, \dots, x_n^0)$. On calcule alors la valeur correspondante de la fonction objectif. On fait ensuite varier successivement chacune des valeurs des paramètres d'une déviation initiale $\pm \Delta x_{ini}$ (ici on adopte $\Delta x_{ini} = 0,64$, ce qui revient à multiplier ou diviser par deux la valeur du

paramètre lorsque la transformation est logarithmique). Lorsque la valeur du $i^{\text{ème}}$ paramètre est modifiée, les deux vecteurs de paramètres testés sont donc $(x_1^0, x_2^0, \dots, x_i^0 + \Delta x_{inib}, \dots, x_n^0)$ et $(x_1^0, x_2^0, \dots, x_i^0 - \Delta x_{inib}, \dots, x_n^0)$. A chaque fois la valeur de la fonction objectif correspondante est calculée.

Lorsque tous les paramètres ont été modifiés un à un, deux cas de figure peuvent se présenter:

1. il y a une amélioration de la valeur de la fonction objectif pour un ou plusieurs des nouveaux vecteurs des paramètres. On retient dans ce cas le vecteur x^1 qui correspond à la meilleure amélioration de la fonction. Ce vecteur, dont les composantes correspondent à celles du vecteur x^0 sauf la composante i qui correspond à $x_i^0 + \Delta x$ ou $x_i^0 - \Delta x$, devient le nouveau vecteur 'initial' de recherche. S'il y a $2n$ améliorations successives de la fonction dans une même direction, le pas de recherche est alors multiplié par deux pour accélérer la recherche (dans la limite de Δx_{\max} , que l'on prendra ici égal à 2). La procédure peut également être affinée et/ou accélérée en utilisant, au delà de $4n$ itérations de calcul, un pas de recherche correspondant à un lissage exponentiel sur les déplacements effectués aux itérations précédentes. Ceci a pour but d'accélérer et d'améliorer la recherche, en particulier dans le cas où il existe sur la surface de réponse une vallée qui ne soit pas dans la direction de l'axe de variation d'un des paramètres. Ce sont les seuls cas où la méthode adopte un déplacement non parallèle à l'un des axes de l'espace des paramètres.
2. il n'y a aucune amélioration de la valeur de la fonction critère pour aucune des modifications des paramètres. On affine alors la recherche en divisant le pas de recherche par deux, et on réitère les modifications des paramètres sur le même vecteur initial x^0 avec ce nouveau pas de recherche.

A chaque itération, on vérifie que les valeurs des paramètres appartiennent au domaine des paramètres transformés possibles préalablement spécifié, ce domaine correspondant à une zone de sensibilité algébrique des paramètres définie a priori.

La procédure s'arrête lorsque la déviation minimale Δx_{\min} spécifiée sur les paramètres (ici, on adopte $\Delta x_{\min} = 0,01$) n'apporte plus d'amélioration de la fonction critère. On a alors localisé l'optimum x^* (ou un optimum) avec une précision relative à la déviation minimale. Par sécurité, pour ne pas avoir des temps de calcul trop longs, la procédure se termine si le nombre d'itérations est supérieur à $100n$ (cas rare que nous n'avons jamais rencontré). La méthode n'a alors pu localiser un optimum sur la surface de réponse.

Nascimento (1995) a réalisé une évaluation de cette méthode en l'appliquant au modèle simple GR4J à quatre paramètres. Pour cela, il l'a comparée à des méthodes globales suivant la méthodologie de Duan *et al.* (1992) sur des données du bassin versant de l'Orgeval. Les résultats montrent que malgré l'existence de maxima locaux et de régions de l'hypersurface faiblement sensibles à la variation des paramètres, la méthode donne de bons résultats. Le faible nombre de paramètres du modèle GR4J et le bon degré d'indépendance entre eux permettent à la méthode 'pas-à-pas' d'avoir ce bon comportement. Même si les résultats ne concernent qu'un bassin, la méthode paraît fiable à plus de 90 %.

Si l'utilisation d'une méthode locale d'optimisation telle que la méthode 'pas-à-pas' ne semble pas a priori être un facteur limitant la validité de notre travail de comparaison de structures de modèles, il ne faut pas pour autant nier certaines limitations intrinsèques d'une telle méthode, notamment liées au choix du point de départ de l'algorithme. Pour les cas d'optimisation problématiques, Nascimento (1995) conseille d'adopter une procédure multidépart. Dans les paragraphes suivants, nous avons essayé d'évaluer l'influence du choix du point de départ de l'optimisation sur les performances de la méthode et des modèles.

3.6. Une évaluation de la méthode 'pas-à-pas': influence du point de départ

3.6.1. Introduction

La méthode 'pas-à-pas', méthode locale d'optimisation, nécessite le choix d'un jeu initial de paramètres comme point de départ du calage. La méthode n'explorant pas tout l'espace des paramètres mais seulement un voisinage de ce point, on comprend l'influence que peut avoir le choix de ces paramètres initiaux sur le résultat final de l'optimisation. Quelques études ont été menées sur l'influence du point de départ pour des méthodes locales (Ibbitt et O'Donnell, 1971b; Gupta et Sorooshian, 1985; Hendrickson *et al.*, 1988). Toutes ont été réalisées sur données synthétiques. Il apparaît que plus le paramètre initial est éloigné du paramètre optimum, moins la méthode réussit à converger vers cet optimum. Néanmoins, pour un paramètre initial calculé à partir du paramètre optimum ayant subi jusqu'à 30% de perturbation, les paramètres optimisés sont généralement égaux ou proches des paramètres optimaux. Au delà, les résultats ne sont plus fiables.

Pour minimiser cette dépendance des résultats de l'optimisation vis-à-vis du point de départ, nous avons vu précédemment que des procédures multidépart sont proposées, dans lesquelles on utilise la méthode locale à partir de chaque élément d'un échantillon de paramètres initiaux (déterminé généralement de façon stochastique), en conservant au final le jeu de paramètres correspondant à la meilleure valeur de la fonction objectif. Si ces méthodes ont le mérite d'explorer une grande partie de l'espace des paramètres, elles sont néanmoins très gourmandes en temps de calcul.

Plutôt que de tester une méthode multidépart avec un grand nombre de points de départ, il est apparu plus intéressant d'adopter une démarche progressive, en complexifiant la méthode de calage par l'introduction d'un deuxième point de départ. Nous pouvons ainsi observer l'effet du double départ sur les performances de l'algorithme d'optimisation. Nous nous plaçons ici dans un contexte de données réelles. Les performances de la méthode seront donc quantifiées par les résultats des modèles.

Après une description de la méthodologie d'évaluation employée, nous essaierons de montrer l'influence de l'utilisation de ce deuxième jeu initial de paramètres sur les performances en calage et a posteriori en contrôle, ainsi que les conséquences sur les valeurs des paramètres optimisés.

3.6.2. Méthodologie d'évaluation

Quatre structures de modèles pluie-débit (GR4J, IHAC, TOPM et HBV0) ont été choisies pour étudier l'influence du point de départ dans l'optimisation. Ce sont des structures présentant des paramétrisations variées et des degrés de complexité différents, avec respectivement quatre, six, six, et neuf paramètres à caler. Les versions de IHAC et TOPM utilisées ici ont un paramètre de moins que les versions testées dans la comparaison. Les différences sont explicitées en Annexe 1. Au cours des premiers tests effectués sur ces quatre modèles, nous avons pu nous rendre compte que par facilité de calage décroissante (correspondant au nombre moyen d'itérations nécessaires pour le calage), on trouve GR4J, IHAC, TOPM et HBV0, ces deux derniers modèles présentant des temps moyens de calage voisins. Nous nous sommes servi ici seulement du sous-échantillon de données françaises, regroupant 307 bassins versants qui représentent 856 périodes de calage et 1780 tests en contrôle.

Alors qu'une méthode multidépart sélectionne au hasard des points de départ, ces derniers pouvant être très excentrés dans l'espace des paramètres par rapport aux valeurs habituelles de convergence, nous avons choisi ici de prendre des points de départ tout à fait plausibles. Une première étape a été de considérer un seul point de départ choisi pour donner des résultats satisfaisants sur l'ensemble de l'échantillon. Ensuite, comme première étape vers une approche multi-départ, nous avons voulu tester l'influence de l'introduction d'un deuxième point de départ. Pour déterminer les deux nouveaux points de départ de l'optimisation, nous avons effectué sur chacun des échantillons de 856 jeux de paramètres obtenus lors de la première optimisation avec un point de départ une classification hiérarchique ascendante (avec une métrique euclidienne) à l'aide du logiciel STATlab. Nous avons ainsi pu à chaque fois grouper les jeux de paramètres en deux classes, ce qui revient à distinguer deux groupes de bassins-périodes pour chacun des modèles. Les deux nouveaux vecteurs initiaux de paramètres ont été calculés comme la moyenne arithmétique des jeux de paramètres appartenant à chacune des classes. Par cette démarche, nous obtenons donc deux points de départ, chacun étant centre de gravité d'une classe. Chaque individu (vecteur de paramètres) issu de la première phase d'optimisation se trouve alors en moyenne plus proche de l'un des deux nouveaux points de départ qu'il ne l'était de l'unique point de départ initial. Si les optima trouvés précédemment sont stables, les deux points de départ ainsi localisés doivent rendre la recherche plus efficace et également plus complète en donnant à l'algorithme la possibilité d'explorer une zone plus étendue de l'espace.

Il était intéressant de voir par ailleurs si l'utilisation de deux points de départ caractérisant deux groupes de paramètres permettait ou non d'accentuer la dissociation de ces groupes, l'augmentation de cette séparation étant indicatrice de la dépendance des paramètres optimisés au point de départ choisi.

Une nouvelle série de calages a alors été effectuée en utilisant la méthode 'pas-à-pas' successivement à partir de ces deux jeux de paramètres initiaux. Pour chaque bassin-période, on retient des deux jeux de paramètres calés celui qui permet d'obtenir la meilleure valeur de la fonction objectif, et on effectue les contrôles correspondants avec le jeu de paramètres retenu.

Enfin, à partir des nouveaux échantillons de paramètres obtenus par optimisation, nous avons réitéré la procédure de classification et de calage à deux points de départ pour vérifier si le réajustement de ces deux points de départ induisait une stabilisation des résultats. Nous avons finalement réalisé une troisième classification pour analyser les résultats de cette troisième série de calages. Dans la suite, nous désignerons par:

- Phase 1: première série de calages, à partir d'un point de départ, et première classification;
- Phase 2: deuxième série de calages, à partir de deux points de départ, et deuxième classification;
- Phase 3: troisième série de calages, à partir de deux points de départ réajustés, et troisième classification.

3.6.3. Classification et points de départ des optimisations

L'utilisation de la classification ascendante hiérarchique a permis de scinder chaque échantillon de jeux de paramètres en deux classes, desquelles nous avons pu extraire des paramètres moyens (centre d'inertie de chaque classe) servant de points de départ à l'optimisation en deux temps avec la méthode 'pas-à-pas'. Le Tableau 3.1 montre pour chacun des modèles les résultats des classifications.

		Phase 1	Phase 2	Phase 3
GR4J	Taux de variance intra-classe	0.80	0.79	0.79
	Taux de variance inter-classe	0.20	0.21	0.21
	Cardinal classe 1	499	402	401
	Cardinal classe 2	357	454	455
IHAC	Taux de variance intra-classe	0.84	0.82	0.84
	Taux de variance inter-classe	0.16	0.18	0.16
	Cardinal classe 1	363	387	289
	Cardinal classe 2	493	469	567
TOPM	Taux de variance intra-classe	0.87	0.86	0.85
	Taux de variance inter-classe	0.13	0.14	0.15
	Cardinal classe 1	591	535	375
	Cardinal classe 2	265	321	481
HBV0	Taux de variance intra-classe	0.85	0.89	0.84
	Taux de variance inter-classe	0.16	0.11	0.16
	Cardinal classe 1	407	297	378
	Cardinal classe 2	449	559	478

Tableau 3.1: Caractéristiques des trois classifications successives effectuées sur le nuage des paramètres optimisés

Le taux de variance intra-classes, qui représente la part relative de l'inertie du nuage des paramètres contenue dans les classes, toujours très supérieur au taux de variance inter-classes, indique tout d'abord que les classes obtenues ne sont que faiblement dissociées et faiblement homogènes, c'est-à-dire que les classifications obtenues ne sont pas de très bonne qualité en terme de dissociation de groupes d'individus. Ce résultat s'explique par le fait que l'échantillon de bassins-périodes présente des conditions hydrologiques très variables, comme nous l'avons vu au chapitre 2, générant par conséquent un nuage de paramètres relativement continu.

Cette continuité du nuage des paramètres est illustrée à la Figure 3.2 dans le cas le plus favorable où la variance inter-classes est la plus élevée (modèle GR4J, classification 2). Nous représentons les projections du nuage de paramètres et celles des centres des deux classes, sur les deux plans formés respectivement par les paramètres 1 et 2, et les paramètres 3 et 4.

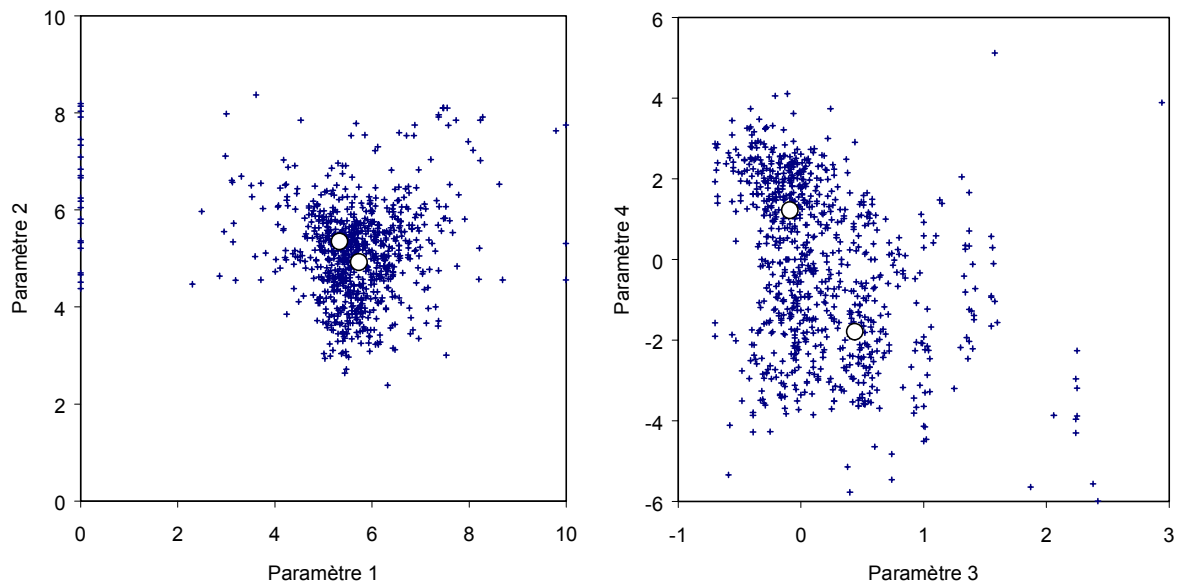


Figure 3.2: Projections du nuage de points sur des plans de l'espace des paramètres transformés du modèle GR4J, avec les deux centres de gravité des classes

Par ailleurs, le gain pour le taux de variance inter-classes, lors du passage des phases 1 à 2, ou des phases 1 à 3, n'est pas systématique et reste relativement faible: il est au maximum de

1,8 % en valeur absolue (TOPM). Dans un cas, il se traduit au contraire par une perte de 4,4 % pour HBV0. Ceci tend à montrer que l'ajout d'un deuxième point de départ lors de l'optimisation avec la méthode 'pas-à-pas' n'a pas un réel impact sur la distribution globale des paramètres optimisés dans l'espace des paramètres. Autrement dit, l'exploration d'une plus grande partie de l'espace ne semble pas favoriser réellement la dispersion des paramètres calés.

Grâce au tableau des contributions fourni par le logiciel, il est possible d'analyser l'influence relative des différents paramètres dans la constitution des classes. Cette contribution représente la part de l'inertie (variance) du nuage de points expliquée par la variable explicative considérée. Il est possible en particulier d'analyser si les paramètres jouant un rôle prépondérant voient ce rôle renforcé lorsque l'on ajoute un deuxième point de calage. Pour tous les modèles, il existe un, deux ou au maximum trois paramètres qui apportent l'essentiel des contributions. Assez logiquement, ces paramètres sont ceux pour lesquels l'écart entre les deux centres de gravités des classes de paramètres est le plus grand.

Pour le modèle GR4J, la contribution la plus forte (50 %) à la dissociation des classes est apportée par le paramètre 4 (ici le paramètre d'échanges souterrains), différenciant les bassins exportant de l'eau et ceux recevant de l'eau. Sa contribution augmente dans les phases 2 (61 %) et 3 (58 %). Le paramètre 3, constante de temps des hydrogrammes unitaires, apporte une contribution relativement importante et assez stable (environ 35 %).

Pour IHAC, deux paramètres apportent des contributions assez similaires: il s'agit des paramètres 3 et 4 (les deux constantes de vidange des réservoirs de routage rapide et lent), avec respectivement 32 et 34 % pour la première classification. Ces contributions fluctuent moyennement pour les deux autres classifications, avec 37 et 30 % pour le paramètre du réservoir rapide et 24 et 28 % pour le réservoir lent. Dans une moindre mesure mais avec une contribution croissante, le paramètre 2 (paramètre de dissociation des écoulements lents et rapides) voit sa contribution augmenter de 15 à 24 %.

Pour TOPM, le paramètre essentiel de séparation des classes est le paramètre 1, paramètre qui intervient dans la vidange du réservoir exponentiel, avec une contribution de 82 % pour la première classification. Cette contribution tombe cependant à 70 % lors de la deuxième classification et chute à 26 % pour la troisième. Pour cette dernière, la principale contribution (49 %) est alors apportée par le paramètre 6, qui intervient dans la courbe de distribution de l'indice topographique.

Enfin, pour HBV0, c'est la contribution du paramètre 4, constante de vidange du réservoir profond qui semble jouer le rôle le plus important dans la séparation des deux classes, avec trois contributions successives de 48, 41 et 42 %. On peut également noter les contributions similaires et relativement constantes (environ 15 %) du paramètre d'infiltration maximale vers le réservoir profond (paramètre 5) et de l'exposant intervenant dans la vidange du réservoir sol (paramètre 7).

Le rôle discriminant fort de certains paramètres des modèles n'est généralement pas renforcé – à part pour le modèle GR4J – par la prise en compte de deux points initiaux de calage qui normalement mettent en valeur le rôle séparateur de ces paramètres. Les contributions, d'une classification à l'autre, ne fluctuent pas uniformément, les paramètres les plus discriminants conservant généralement leur prépondérance. Le cas de TOPM est à cet égard singulier puisque le paramètre essentiel de dissociation des classes lors de la première classification perd son rôle prépondérant au profit d'un autre paramètre, ceci avec une amélioration constante au cours des trois optimisations de la variance inter-classes. Ceci pourrait être un indice d'interdépendance entre paramètres.

Concernant la stabilité des jeux de paramètres initiaux utilisés pour les optimisations 2 et 3 (calculés dans les phases 1 et 2), si certains des paramètres varient peu ou restent inchangés, d'autres, en revanche, voient leur valeur très largement modifiée. Le Tableau 3.2 indique les modifications relatives moyennes (en %) des composantes des points de départ d'une classification à l'autre. Ces modifications sont généralement assez importantes, mais souvent plus petites entre les deux dernières classifications qu'entre les deux premières, ce qui semble signifier que les changements apportés par le réajustement des paramètres initiaux sont moindres par rapport à ceux entraînés par l'ajout d'un deuxième point de départ. Une fois encore, le cas de TOPM est singulier, ce qui peut être rapproché des observations faites précédemment.

		GR4J	IHAC	TOPM	HBV0
Phase 1 -> Phase 2	Classe1	13,2	16,7	14,1	23,6
	Classe2	13,6	28,3	35,5	28,8
Phase 2 -> Phase 3	Classe1	2,4	20,1	53,1	10,9
	Classe2	8,0	22,4	80,4	5,9

Tableau 3.2: Modifications relatives moyennes (en %) des coordonnées des centres de gravités des classes

Cependant, si les centres de gravité des classes varient, et si la contribution des différentes composantes des points initiaux à la dissociation des classes peut également être variable, il est difficile de tirer des conclusions quant à la stabilité générale du nuage des jeux de paramètres obtenus en calage. La forte variabilité des cardinaux des deux classes, mise en évidence dans le Tableau 3.1, n'est pas forcément significative de grands changements de résultats, étant donné la faible séparation des nuages des deux classes.

Pour étudier cette stabilité, nous avons regardé dans un premier temps l'évolution du nombre de bassins appartenant à l'une ou l'autre des classes. Nous considérons qu'un bassin s'apparente entièrement à l'une des classes si les jeux de paramètres issus des calages sur toutes les périodes appartiennent à la même classe. La Figure 3.3 montre sur l'échantillon de 307 bassins, le nombre total de bassins appartenant entièrement à l'une des deux classes, et le nombre de bassins appartenant seulement à la classe 1.

Le nombre de bassins appartenant aux deux classes évolue peu d'une optimisation à l'autre, montrant que l'introduction d'un deuxième point de départ n'améliore pas (ni ne détériore) la capacité de la méthode à identifier des jeux de paramètres similaires d'une période à l'autre sur un même bassin. D'autre part, derrière cette stabilité globale, il apparaît que les cardinaux des classes 1 et 2 sont beaucoup plus variables: les bassins ne sont pas toujours apparentés à la même classe et ce ne sont pas forcément les mêmes bassins qui peuvent être considérés comme appartenant entièrement à l'une des classes. Les plus fortes variations du cardinal de la classe 1 sont observées pour TOPM et HBV0. Ces résultats restent cependant dépendants de la méthode de classification employée.

Il ressort de tout ce qui précède que l'utilisation de deux points de départ au lieu d'un seul lors de l'optimisation n'influence guère la stabilité d'ensemble de l'échantillon des paramètres calés. En revanche, une certaine variabilité semble apparaître au sein de cet échantillon, que nous allons essayer de mieux quantifier dans la suite.

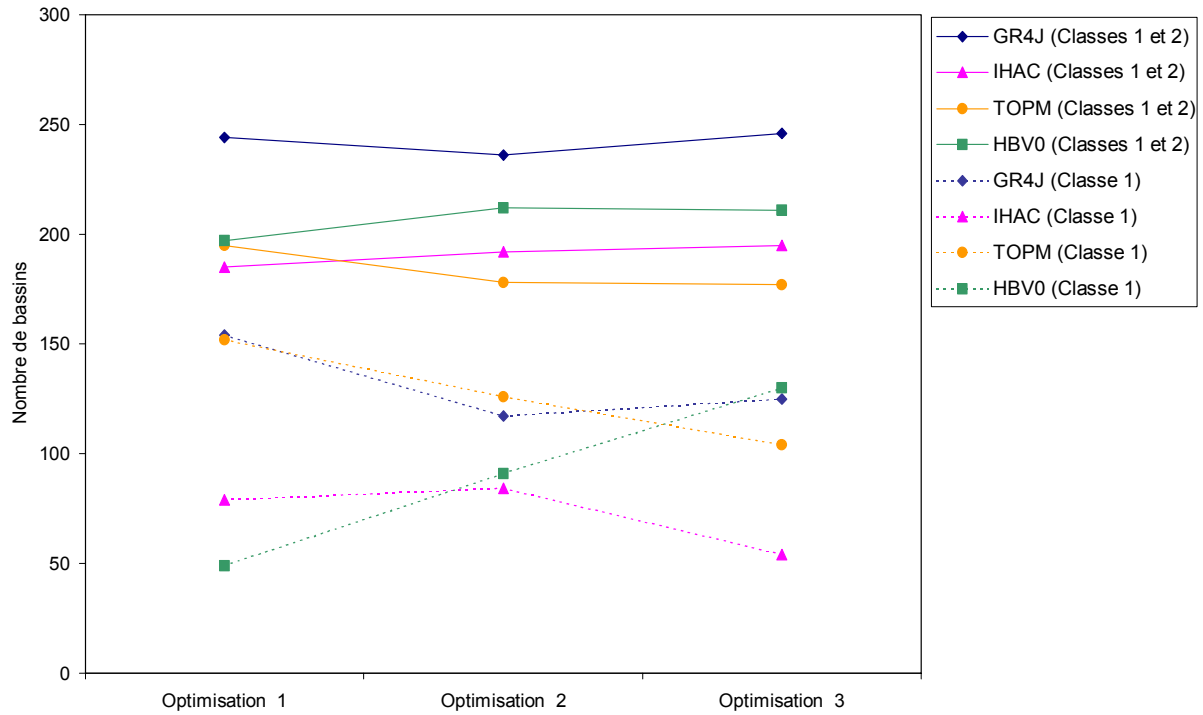


Figure 3.3: Bassins versants appartenant à l'une des classes de la classification

3.6.4. Analyse des paramètres calés

Pour expliquer plus en détail les fluctuations relatives constatées sur les couples de points de départ en passant d'une phase à une autre (Tableau 3.2), le Tableau 3.3 fournit pour chaque échantillon de 856 jeux de paramètres les variations absolues moyennes des paramètres en passant d'une optimisation à la suivante. Pour rendre ces variations absolues moyennes plus significatives, nous avons indiqué entre parenthèses le taux de variation (en %) par rapport à la moyenne initiale des paramètres.

	Paramètres	Phase1->Phase2	Phase2->Phase3		Paramètres	Phase1->Phase2	Phase2->Phase3
GR4J	1	0,21 (4)	0,06 (1)	IHAC	1	0,13 (2)	0,02 (1)
	2	0,09 (2)	0,03 (1)		2	0,6 (128)	0,21 (91)
	3	0,02 (13)	0,01 (6)		3	0,40 (93)	0,12 (28)
	4	0,15 (68)	0,05 (26)		4	0,54 (20)	0,18 (6)
HBV0	1	0,26 (4)	0,17 (3)		5	0,12 (44)	0,04 (17)
	2	0,33 (97)	0,27 (79)		6	0,05 (3)	0,01 (1)
	3	0,4 (32)	0,27 (25)	TOPM	1	1,61 (150)	1,24 (126)
	4	0,8 (14)	0,26 (5)		2	0,79 (16)	0,75 (15)
	5	0,79 (188)	0,26 (236)		3	0,08 (2)	0,06 (1)
	6	0,06 (38)	0,03 (16)		4	0,05 (17)	0,04 (14)
	7	1,12 (700)	0,85 (354)		5	0,35 (8)	0,22 (5)
	8	1,09 (45)	0,51 (19)		6	1,62 (91)	1,4 (94)
	9	1,46 (258)	0,81 (1350)				

Tableau 3.3: Différences absolues moyennes et relatives (en %) des paramètres d'une optimisation à l'autre

Il semble y avoir une relative stabilisation des résultats sur les paramètres moyens lorsque l'on ne fait que réajuster les points de départ (optimisations 2 à 3). Ceci est confirmé par les valeurs des différences absolues moyennes toujours plus petites dans le passage des phases 2 à 3 que dans le passage des phases 1 à 2.

La variabilité des paramètres permet en partie d'expliquer les fluctuations de leurs contributions dans les différentes classifications. Ainsi, par exemple, les variations importantes des contributions à la différenciation des classes des paramètres 1 et 6 de TOPM sont à relier aux variations importantes observées sur ces paramètres aux cours des trois optimisations.

Par ailleurs, l'importante variabilité constatée pour certains paramètres du modèle HBV0 peut être synonyme de non sensibilité ou d'existence d'interdépendance et donc de compensations entre paramètres. Il existe en effet, comme nous le verrons par la suite, de nombreux cas où des jeux différents de paramètres conduisent à des performances similaires du modèle.

Le Tableau 3.4 présente le nombre de calages où la méthode 'pas-à-pas' a convergé vers le même jeu de paramètres entre les optimisations 1 et 2, 1 et 3, et 2 et 3.

	GR4J	IHAC	TOPM	HBV0
Opt1->Opt2	394	107	2	8
Opt1->Opt3	374	102	4	5
Opt2->Opt3	540	150	6	6

Tableau 3.4: Nombre de convergences vers le même optimum (sur un total de 856 optimisations)

La méthode de recherche ne converge pratiquement jamais vers le même optimum pour les structures HBV0 et TOPM, même lorsque l'on ne fait que réajuster les points de départ de l'optimisation. Pour IHAC, la méthode prouve une plus grande capacité à converger vers le même optimum, même si la proportion (environ un huitième des bassins-périodes) reste assez faible. Enfin, pour GR4J, la méthode réussit à converger dans pratiquement la moitié des cas vers le même optimum lorsque l'on passe d'un à deux points de départ, et dans plus de 60 % des cas lorsque l'on ne fait que réajuster les points initiaux.

Ces résultats confirment tout d'abord la dépendance des paramètres optimisés vis-à-vis du point de départ de la méthode. D'autre part, si la différence entre un modèle à 4 paramètres (GR4J) et un modèle à 9 paramètres (HBV0) semble indiquer que la méthode est plus efficace pour de faibles niveaux de complexité du modèle, la comparaison des résultats de IHAC et TOPM, tous deux à 6 paramètres, montre que la structure et/ou la paramétrisation du modèle jouent un rôle essentiel dans la capacité de la méthode à converger vers le même optimum pour des points de départ différents. Il semble donc qu'il y ait des problèmes d'identifiabilité qui limitent la capacité de la méthode à converger vers un même optimum.

Cette remarque est renforcée par le Tableau 3.5, qui donne des informations similaires au Tableau 3.4, mais nous nous sommes intéressés ici aux points de convergence similaires et non plus identiques. Deux jeux de paramètres sont considérés similaires si la variation moyenne des paramètres n'excède pas la valeur du critère d'arrêt de l'optimisation (0,01). Cette condition moins stricte que l'égalité de tous les paramètres permet d'évaluer la capacité de la méthode à converger vers des optima situés dans des zones de l'espace très voisines.

	GR4J	IHAC	TOPM	HBV0
Opt1->Opt2	723	440	54	96
Opt1->Opt3	729	417	40	83
Opt2->Opt3	769	545	69	153

Tableau 3.5: Nombre de convergence vers des optima similaires (sur un total de 856 optimisations)

La différence entre IHAC et TOPM est ici renforcée. La convergence vers des optima similaires devient meilleure pour HBV0 que pour TOPM, même si les optima de ces deux

structures restent faiblement identifiables. Pour GR4J, on constate que la méthode converge vers un optimum similaire dans plus de 80 % des cas.

La méthode 'pas-à-pas' étant très efficace pour le modèle le plus simple, il semble que ce sont des problèmes d'identifiabilité d'un optimum qui viennent handicaper la méthode de calage pour les structures plus complexes. Si l'identifiabilité dépend en partie du nombre de paramètres (plus il y a de paramètres, plus il risque d'y avoir des interdépendances entre eux), la paramétrisation semble jouer un rôle essentiel dans l'observabilité d'un optimum. Nous rejoignons en cela les remarques formulées par Gan et Biftu (1996).

L'introduction d'un deuxième point de départ lors de l'optimisation peut améliorer la stabilité des valeurs des paramètres (la variation des paramètres est moins forte lors du réajustement des points de départ). Cependant, elle ne permet pas d'atténuer les problèmes d'optimisation liés à l'interdépendance entre paramètres.

Ces remarques sur les paramètres doivent être complétées par une analyse des performances du modèle dans les diverses situations d'optimisation. En effet, nous venons de comparer les optima trouvés par la méthode de calage au niveau de la valeur des paramètres obtenus. Il est important de voir si l'introduction d'un deuxième point de départ permet d'améliorer la qualité des optima trouvés. Travaillant sur données réelles, nous ne pouvons faire ici que des comparaisons relatives entre les résultats, puisqu'on ne connaît pas (s'ils existent) les vrais optima.

3.6.5. Performances en calage

Considérons tout d'abord l'échantillon d'un point de vue global. La Figure 3.4 présente l'évolution des distributions cumulées des critères d'optimisation obtenus au cours des trois séries d'optimisations pour les structures GR4J et TOPM. Ces distributions varient très peu d'une situation d'optimisation à une autre. Ces courbes sont pratiquement confondues dans le cas de la structure GR4J. Le Tableau 3.6 complète ces commentaires. Il indique les performances globales sur l'échantillon lors des trois séries de calage au niveau du critère d'optimisation.

Les gains en terme de performance moyenne au calage sont faibles: de 0,2 à 0,4 % en valeur absolue en passant des optimisations 1 à 2, c'est-à-dire en ajoutant un second point de départ. Le plus faible gain (0,2 points) est obtenu pour le modèle GR4J. On trouve ensuite IHAC avec 0,3 % puis TOPM et HBV0 avec 0,4 %. Il apparaît que l'introduction d'un deuxième point de départ permet en général à la méthode de recherche de localiser un optimum à peine meilleur qu'avec un seul point de départ. Cette situation ressort en dépit du fait que l'on augmente les chances de trouver un meilleur optimum, étant donné que l'on couvre une région de l'espace plus étendue pendant le processus d'optimisation.

Par ailleurs, après un réajustement des deux paramètres initiaux, les performances obtenues à l'optimisation 3 restent pratiquement inchangées par rapport à l'optimisation 2, excepté dans le cas de TOPM où une légère amélioration est constatée.

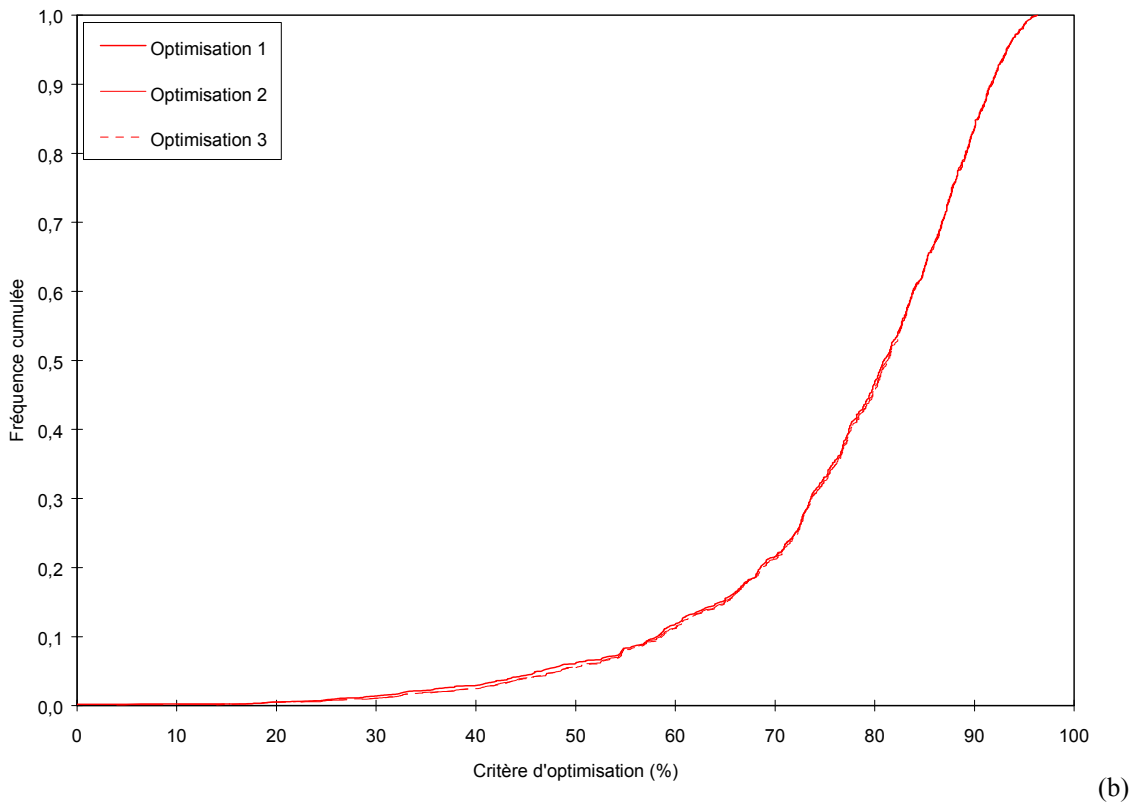
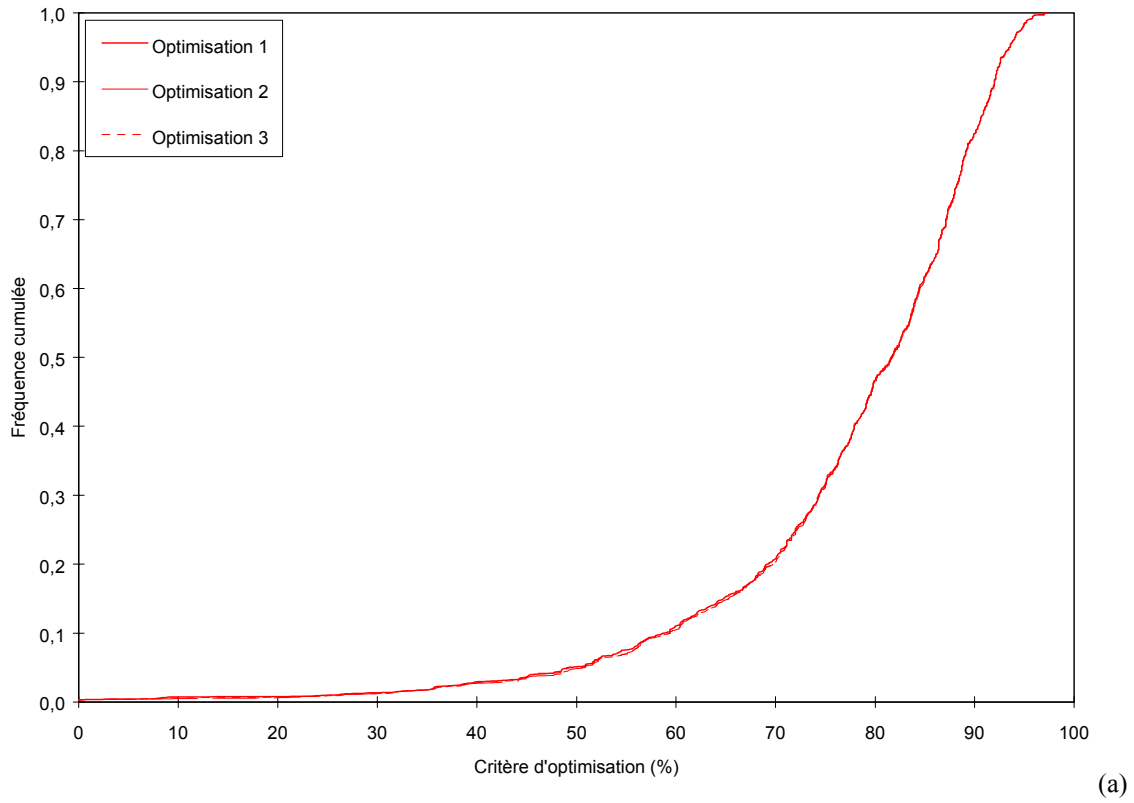


Figure 3.4: Distributions des critères d'optimisation pour les structures (a) GR4J et (b) TOPM

		Optimisation 1	Optimisation 2	Optimisation 3
GR4J	Maximum	97.2	97.2	97.2
	Moyenne	77.8	78.0	78.0
	Minimum	-84.1	-1.9	-31.9
	F(0)	0.003	0.003	0.003
IHAC	Maximum	94.6	94.7	94.6
	Moyenne	75.1	75.4	75.4
	Minimum	-0.5	0.5	0.8
	F(0)	0.001	0.000	0.000
TOPM	Maximum	96.3	96.3	96.3
	Moyenne	77.3	77.7	77.8
	Minimum	-117.0	4.0	14.0
	F(0)	0.001	0.000	0.000
HBV0	Maximum	96.5	96.5	96.4
	Moyenne	78.8	79.2	79.2
	Minimum	-63.1	-63.1	-63.1
	F(0)	0.001	0.001	0.001

Tableau 3.6: Performances globales sur l'échantillon pour les trois séries de calage (maximum, moyenne, minimum des performances et pourcentage F(0) de critères négatifs)

Considérons maintenant l'évolution des performances bassin-période par bassin-période. Dans le Tableau 3.7, nous donnons les écarts minimaux, maximaux et moyens obtenus d'une série d'optimisation à l'autre. N1 représente le nombre de bassins-périodes où l'on a eu une diminution de la fonction objectif, N2 celui où la fonction objectif est restée identique et N3 celui où la fonction objectif a augmenté. Entre parenthèses sont indiqués pour N1 le nombre de calages avec une diminution de plus de 10 % en valeur absolue de la fonction objectif, pour N3 le nombre de calages avec une augmentation de plus de 10 % et pour N2 le nombre de calages où la fonction objectif est restée inchangée avec changement de la valeur des paramètres. Dans ce tableau, deux valeurs de la fonction objectif sont considérées identiques (équivalentes) si elles diffèrent de moins de 0,1 %, ce qui correspond à un seuil en deçà duquel les différences sont considérées comme non significatives.

	GR4J		IHAC		TOPM		HBV0	
	Opt1->Opt2	Opt2->Opt3	Opt1->Opt2	Opt2->Opt3	Opt1->Opt2	Opt2->Opt3	Opt1->Opt2	Opt2->Opt3
Minimum	-1,3	-14,8	-2,2	-2,3	-12,9	-10,4	-17,7	-13,1
Maximum	52,2	30,0	55,3	1,4	121,0	37,2	18,0	12,5
Moyenne	0,2	0,0	0,3	0,0	0,4	0,1	0,4	0,0
N1 ↘	10 (0)	14 (2)	16 (0)	24 (0)	60 (2)	90 (1)	110 (2)	129 (1)
N2 =	795 (401)	826 (286)	670 (562)	794 (644)	598 (596)	649 (642)	393 (385)	560 (554)
N3 ↗	51 (4)	16 (1)	170 (1)	38 (0)	198 (5)	117 (2)	353 (4)	167 (1)

Tableau 3.7: Différences dans les valeurs de la fonction objectif entre les trois optimisations (différences maximales, minimales et moyennes; N1: nombre de bassins-périodes avec diminution de la fonction objectif, N2 nombre de bassins-périodes sans changement de la fonction objectif; N3: nombre de bassins-périodes avec augmentation de la fonction objectif)

Dans un certain nombre de cas, il y a une diminution de la fonction objectif. Cette situation peut paraître surprenante, mais elle s'explique par le fait que d'une série d'optimisation à l'autre, on ne reprend jamais les points de départ initiaux précédents. Si tel avait été le cas, il n'y aurait pu y avoir de situation avec diminution de la fonction objectif. Le nombre de cas de diminutions reste cependant assez faible, et ce d'autant plus que le modèle est simple. Par ailleurs, il n'y a que très peu de cas où la procédure d'optimisation conduit à une importante diminution ou augmentation (+ ou - 10 points) de la fonction objectif.

Lors de l'ajout d'un deuxième point de départ (passage des optimisations 1 à 2), pour la majorité des calages, la valeur de la fonction objectif reste inchangée, sauf dans le cas du modèle HBV0. En rapprochant ces résultats des commentaires formulés pour le Tableau 3.4, on constate par ailleurs qu'il arrive souvent que la méthode localise des optima similaires de la fonction objectif avec des jeux de paramètres différents, renforçant l'idée qu'il existe sur la surface de réponse des solutions équivalentes.

Par ailleurs, l'amélioration de la fonction objectif en passant de un à deux points de départ se produit dans 6 % des cas pour GR4J, 20 % pour IHAC, 23 % pour TOPM et 41 % pour HBV0. Dans le cas de GR4J, ces résultats sont à rapprocher des ordres de grandeur de 90 % de fiabilité de la méthode 'pas-à-pas' énoncés par Nascimento (1995) concernant les performances en calage.

Les résultats obtenus consécutivement au réajustement des points de départ sont généralement plus stables, avec une proportion accrue de fonctions objectif inchangées et une diminution des cas de forte amélioration ou de forte dégradation de la fonction. Cette meilleure stabilité est à rapprocher de la meilleure stabilité des paramètres constatée précédemment.

Il ressort de l'analyse précédente les points importants suivants:

- l'utilisation de deux points de départ au lieu d'un seul dans le processus d'optimisation, qui double la durée de la procédure, ne permet qu'une très faible amélioration des performances moyennes en calage.
- le nombre de cas où la valeur de la fonction critère reste inchangée, en ajoutant un point de départ, est généralement élevé. Cette constance de la fonction critère s'accompagne parfois ou souvent (selon les modèles) de la modification des paramètres du modèle. Il existe donc des cas où l'on a des solutions équivalentes sur l'hypersurface de la fonction critère, c'est-à-dire qu'il peut exister plusieurs jeux de paramètres acceptables qui donnent des solutions de qualité semblable.
- le nombre de calages où l'on constate une amélioration sensible de la fonction objectif reste toujours minoritaire, généralement inférieur à un quart du nombre des bassins-périodes (sauf dans un cas où la proportion atteint 41 %).
- le réajustement du couple de points de départ modifie de façon négligeable les résultats moyens sur l'échantillon, indiquant une relative stabilité des résultats, à mettre en lien avec la relative stabilité correspondante des paramètres optimisés.

Ces points tendent à indiquer que l'amélioration des performances en calage consécutive au perfectionnement de la technique d'optimisation est limitée. Nos résultats montrent qu'il existe des situations d'équivalence d'optima sur l'hypersurface de la fonction critère. Il devient alors impossible à la méthode de calage de converger toujours vers un même optimum lorsque l'on modifie le point de départ (cette incapacité serait d'ailleurs la même pour une autre méthode d'optimisation). Cependant, la faible amélioration des performances que nous avons constatée semble indiquer que la procédure de recherche, avec un seul point de départ, est capable d'identifier des optima de qualité presque aussi bonne que celle de l'optimum général.

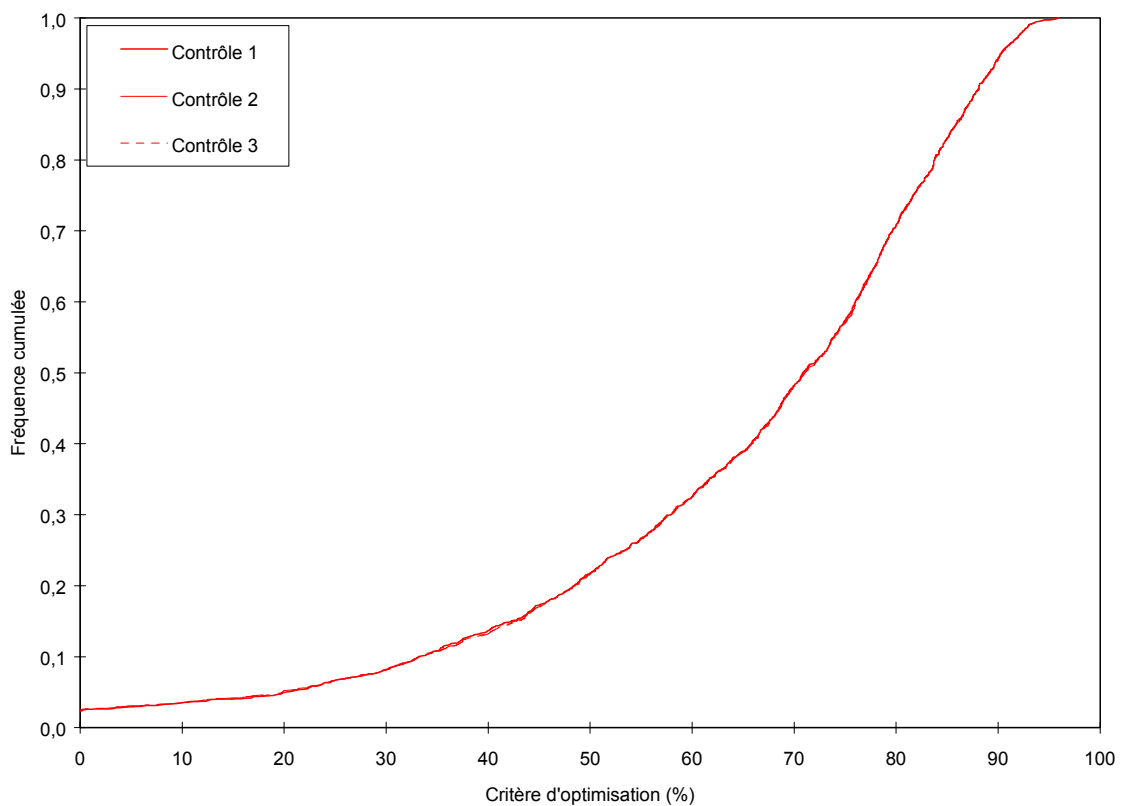
Dans le paragraphe suivant, nous étudions l'influence du calage à deux départs sur les performances en contrôle, sur lesquels se fondent l'essentiel de notre appréciation des structures de modèle dans notre étude de comparaison.

3.6.6. Performances en contrôle

Les performances en contrôle sont ici jugées en fonction du même critère que celui utilisé en calage. Le Tableau 3.8 regroupe les résultats obtenus par les quatre structures de modèles sur les 1780 tests en contrôles.

		Optimisation 1	Optimisation 2	Optimisation 3
GR4J	Maximum	97.1	97.1	97.1
	Moyenne	70.0	70.2	70.0
	Minimum	-537.1	-328.7	-507.3
	F (0)	2.4	2.4	2.3
IHAC	Maximum	93.6	93.5	93.5
	Moyenne	63.2	63.2	63.2
	Minimum	-672.9	-673.8	-676.2
	F (0)	2.3	2.4	2.3
TOPM	Maximum	95.7	95.7	95.6
	Moyenne	67.7	68.6	68.7
	Minimum	-514.4	-488.6	-485.1
	F (0)	2.3	1.8	1.8
HBV0	Maximum	95.2	95.3	95.3
	Moyenne	67.6	67.9	67.9
	Minimum	-598.5	-613.5	-575.5
	F (0)	2.7	2.5	2.3

Tableau 3.8: Performances des quatre structures de modèle en contrôle (performances maximales, minimales et moyennes et pourcentage F(0) des critères négatifs)



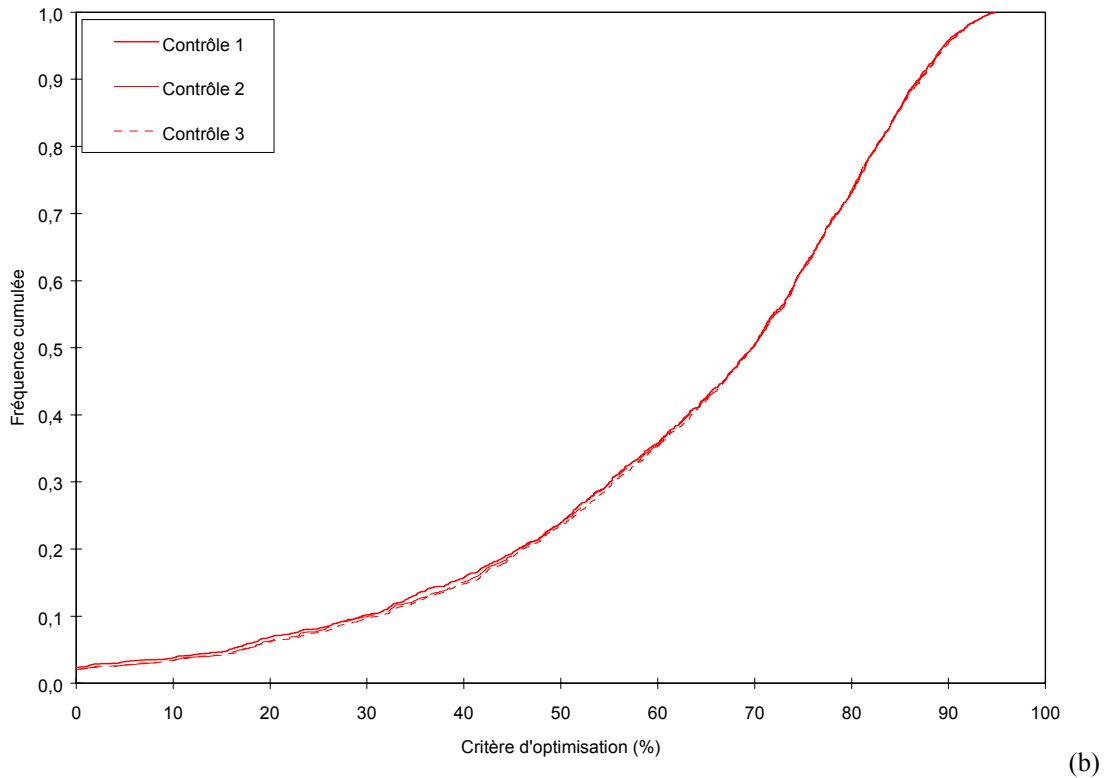


Figure 3.5: Distributions des critères de performance en contrôle pour les structures (a) GR4J et (b) TOPM

Les performances moyennes sur l'échantillon en contrôle sont très stables pour les structures GR4J et IHAC. Une petite amélioration est constatée pour la structure HBV0 lorsque l'on ajoute un deuxième point de départ au niveau de l'optimisation. Les gains les plus sensibles (0,9 points) sont obtenus pour la structure TOPM.

La Figure 3.5 illustre l'évolution des distributions des performances pour les structures GR4J et TOPM. Les courbes restent pratiquement inchangées pour GR4J (et également pour IHAC). En revanche, on retrouve graphiquement les différences plus nettes constatées précédemment pour TOPM et HBV0.

Comme cela a été fait pour les résultats des calages, nous avons regroupé dans le Tableau 3.9 les résultats des tests en simulation (1780 contrôles). N1, N2 et N3 ont les mêmes significations que pour le Tableau 3.7. Dans le Tableau 3.10, nous avons rapporté les nombres N1 et N3 (représentant le nombre de bassins-périodes où il y a amélioration ou dégradation de la valeur du critère) obtenus en calage et contrôle entre les séries de test 1 et 2, au nombre total de tests effectués (856 calages et 1780 contrôles).

	GR4J		IHAC		TOPM		HBV0	
	Opt1->Opt2	Opt2->Opt3	Opt1->Opt2	Opt2->Opt3	Opt1->Opt2	Opt2->Opt3	Opt1->Opt2	Opt2->Opt3
Minimum	-20.1	-178.6	-33.4	-17.1	-19.6	-31.5	-46.9	-80.8
Maximum	208.4	6.7	19.9	3.1	448.4	83.8	84.6	90.9
Moyenne	0.2	-0.2	0.0	0.0	0.9	0.1	0.3	0.0
N1 ↘	382 (2)	247 (5)	508 (12)	456 (3)	457 (9)	499 (10)	652 (40)	650 (18)
N2 =	1056	1320	684	889	770	808	333	469
N3 ↗	342 (9)	213 (0)	588 (5)	435 (0)	553 (37)	473 (18)	795 (38)	661 (23)

Tableau 3.9: Différences de performances entre les trois séries de contrôles (différences maximales, minimales et moyennes; N1: nombre de bassins-périodes avec diminution de la fonction objectif; N2 nombre de bassins-périodes sans changement de la fonction objectif; N3: nombre de bassins-périodes avec augmentation de la fonction objectif)

Les proportions d'améliorations et de dégradations du critère sont très différentes entre le calage et le contrôle. Il existe notamment une beaucoup plus forte proportion des dégradations du critère en contrôle qu'en calage. La différence entre améliorations et dégradations montre que, si la procédure de calage avec deux points de départ apporte un assez net avantage en phase d'optimisation, cet avantage est beaucoup moins net en contrôle. Cette tendance est ici d'autant plus marquée que le modèle a moins de paramètres. Ceci tend à indiquer que les améliorations induites au niveau du calage par les efforts supplémentaires consentis dans la procédure d'optimisation ne se retrouvent pas du tout en proportions équivalentes au contrôle.

		GR4J	IHAC	TOPM	HBV0
Calage	Amélioration (%)	6	20	23	41
	Dégradation (%)	1	2	7	13
	Différence (%)	5	18	16	28
Contrôle	Amélioration (%)	19	33	31	45
	Dégradation (%)	21	29	26	37
	Différence (%)	-2	4	5	8

Tableau 3.10: Proportions des calages (sur un total de 856) et de contrôles (sur un total de 1780) pour lesquels il y a eu amélioration et dégradation du critère de performance en calage et en contrôle, entre la première et deuxième série de tests

Ces résultats vont dans le sens d'un problème de sur-calage des modèles trop complexes: l'inspection d'une plus grande région de l'espace permet d'améliorer légèrement la qualité des optima trouvés en terme de valeur de la fonction critère, mais il n'y a pas parallèlement une meilleure robustesse de ces optima. Les jeux de paramètres optimisés deviennent en effet trop spécifiques aux périodes de calage et ne garantissent pas de meilleures validations.

Par ailleurs, dans le cas de TOPM, il y a deux bassins-périodes pour lesquels le gain de performance est supérieur à 300 points. Dans ces deux cas, les critères obtenus à la deuxième série de tests demeurent cependant très largement négatifs, ce qui signifie que les paramètres identifiés ne sont pas davantage capables d'expliquer la variance des débits observés. Ces deux gains de performance influencent néanmoins lourdement l'amélioration moyenne des performances de ce modèle lorsque l'on passe de un à deux points de départ. Ces observations viennent donc tempérer le jugement que les gains de ce modèle sont assez nettement meilleurs que ceux enregistrés par les autres modèles au contrôle lorsque l'on passe des phases de test 1 à 2.

3.6.7. Conclusion

Nous pouvons dégager quelques points importants de cette évaluation de l'influence du nombre et du choix des points de départ de l'optimisation avec la méthode 'pas-à-pas':

- en calage, les améliorations moyennes constatées restent faibles lorsque l'on passe d'un à deux points de départ dans le processus de calage. L'amélioration de la fonction objectif est loin d'être systématique puisqu'elle est constatée au maximum dans 41 % des cas pour le modèle HBV0 et pour seulement 6 % des cas pour GR4J. Ces améliorations sont généralement d'importance limitée et ne dépassent qu'exceptionnellement 10 % en valeur absolue;
- la convergence de la méthode vers des optima similaires lorsque l'on passe d'un à deux points de départ est d'autant plus vérifiée que le modèle est simple. Pour les modèles avec le plus grand nombre de paramètres, la proportion de bassins-périodes où la fonction critère est restée inchangée est bien supérieure à celle où les paramètres identifiés sont

restés inchangés, ceci indiquant qu'il existe des optima différents correspondant à des valeurs identiques de la fonction critère;

- cette absence d'unicité et d'identifiabilité de l'optimum induit des difficultés pour la méthode de calage à converger vers un optimum stable de l'espace des paramètres;
- cette instabilité se traduit au niveau du contrôle par un manque de robustesse. Les améliorations apportées au calage par l'utilisation de deux points de départ au lieu d'un seul ne se retrouvent pas du tout en même proportion en contrôle. Les performances moyennes au contrôle varient peu lorsque l'on perfectionne la stratégie d'optimisation;
- les diminutions de la fonction objectif observées en calage (de 1 % à 13 % des cas suivant les structures) montrent que, lorsque l'on passe d'un point de départ à deux points de départ choisis de façon à permettre une meilleure couverture de la zone probable de convergence des paramètres, la méthode peut quand même, dans un nombre limité de cas, être piégée sur des optima de moins bonne qualité.

Nous pourrions prolonger cette étude en testant l'effet de l'utilisation d'un plus grand nombre de points de départ (3, 4, ..., n) en adoptant la même méthodologie. Or il paraît probable, étant donné les résultats déjà obtenus, que les performances en contrôle se stabiliseraient très rapidement et atteindraient un palier peu éloigné des performances rapportées ici. En effet, pour les modèles se calant aisément (GR4J et IHAC), les améliorations constatées sont déjà faibles. Pour les modèles plus complexes à optimiser (TOPM et HBV0), les performances d'une quelconque méthode de calage sont rapidement limitées par les problèmes de robustesse ou de stabilité, comme le mentionnent également Gan et Biftu (1996).

Notons enfin que les problèmes cités ci-dessus peuvent varier en fonction des critères utilisés (ou imposés). Certains critères peuvent donner lieu à une surface de réponse moins chaotique.

3.7. Effet de la longueur de la période de calage

Dans la dernière partie de ce chapitre, nous avons voulu compléter l'évaluation de la procédure de calage en regardant l'effet de la longueur de la période de calage sur les résultats des modèles. En effet, même si nous avons vu dans les parties précédentes que ce sont essentiellement des problèmes d'identifiabilité qui semblent limiter les performances de la méthode d'optimisation, il paraît concevable que le contenu des séries hydrologiques utilisées pour le calage ont une influence sur la détermination des paramètres. Certains auteurs, tels que Sorooshian *et al.* (1983), Allred et Haan (1991) et Yapo *et al.* (1996), se sont intéressés à cette problématique, tant au niveau des performances que des paramètres du modèle. Dans le cas du modèle de Sacramento par exemple, Yapo *et al.* (1996) préconisent qu'une durée minimale de huit ans doit être utilisée lors de l'optimisation pour avoir des résultats relativement indépendants de la période sélectionnée.

Les degrés de complexité différents des modèles que nous avons testés dans la comparaison sont susceptibles d'induire des niveaux d'exigence différents, en terme d'information nécessaire pour l'identification des paramètres. Il paraît en effet possible qu'un modèle à neuf paramètres ait besoin de plus d'information (nombre d'années de calage) qu'un modèle à trois paramètres pour fournir des résultats satisfaisants.

Nous considérons ici l'ensemble des bassins français et étrangers. Dans la série de tests effectués sur les modèles, la longueur des 1284 périodes de calage de l'échantillon des 429 bassins varie entre une et huit années comme nous l'avons vu au chapitre 2.

Pour évaluer une possible influence de la longueur de la période de calage sur les performances des modèles, la procédure la plus adaptée serait, pour chaque bassin, de découper la série de données successivement en périodes de différentes longueurs et tester tous les modèles sur ces périodes, puis calculer des performances moyennes sur l'ensemble des bassins pour chaque longueur de calage. Par cette procédure, chaque moyenne dépendrait des résultats sur tous les bassins. De façon pratique, cette solution est cependant difficilement réalisable. D'une part elle demande une masse très importante de calcul si l'on veut traiter de façon homogène tout l'échantillon de bassins. D'autre part, elle suppose que l'on dispose sur chaque bassin d'une série de données suffisamment longue pour réaliser ces découpages en sous-périodes de différentes longueurs. Ce n'était pas notre cas sur les 429 bassins: seuls 32 bassins (dont 28 aux Etats-Unis) présentent des séries de longueur supérieure à 30 ans, comme le montrait la Figure 2.10 du chapitre 2. Pour répondre à notre question, une solution alternative aurait été de travailler uniquement sur ces bassins.

Pour continuer à travailler sur un grand nombre de bassins versants et ne pas induire une trop importante charge de calculs supplémentaires, nous avons préféré réaliser une évaluation moins détaillée et plus qualitative, en allongeant la longueur moyenne des périodes testées. Pour cela, une procédure simple a été utilisée: nous avons agrégé, pour les bassins comportant plus de deux périodes de calage, certaines des périodes entre elles: par exemple, pour un bassin comportant cinq périodes de quatre ans, nous avons obtenu ainsi trois périodes, deux de huit ans et une de quatre ans. Cette agglomération n'a été possible que pour un peu plus de la moitié des bassins versants, puisque initialement 198 bassins ne présentaient que deux périodes test. Nous avons donc effectué la comparaison de l'évolution des performances avec l'allongement de la longueur de la période sur un sous-échantillon de 231 bassins versants (164 en France, 46 aux Etats-Unis, 10 en Australie, 4 au Brésil et 7 en Côte d'Ivoire).

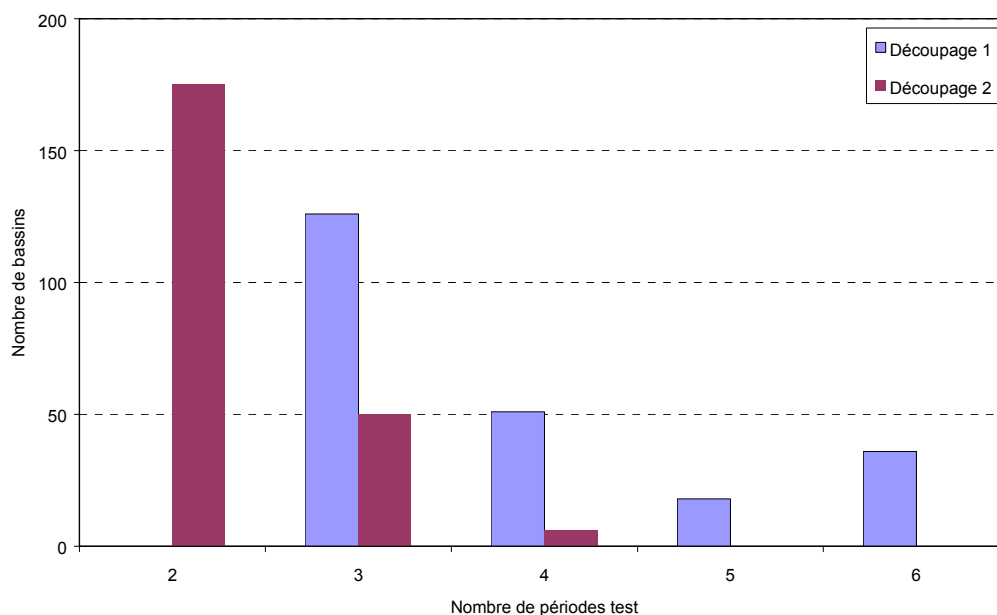


Figure 3.6: Nombre de périodes de calage par bassin

Avec le découpage initial en sous-périodes, il y a sur cet échantillon 888 périodes de calage et 2808 tests en contrôle. Pour le deuxième découpage (après agglomération), on obtient 524 périodes de calage et 722 tests en contrôle. L'importante diminution du nombre des tests en contrôle provient du fait que la majorité des bassins sont testés seulement sur deux ou trois périodes dans le second cas (voir Figure 3.6). La Figure 3.7 montre les différences entre les deux découpages en sous-périodes. La longueur moyenne des périodes est de 5,1 ans dans le

premier cas et de 8,6 ans dans le second. Sur l'échantillon total des 429 bassins, la longueur moyenne des périodes sur l'échantillon total est de 4,8 ans, avec proportionnellement un peu plus de courtes périodes.

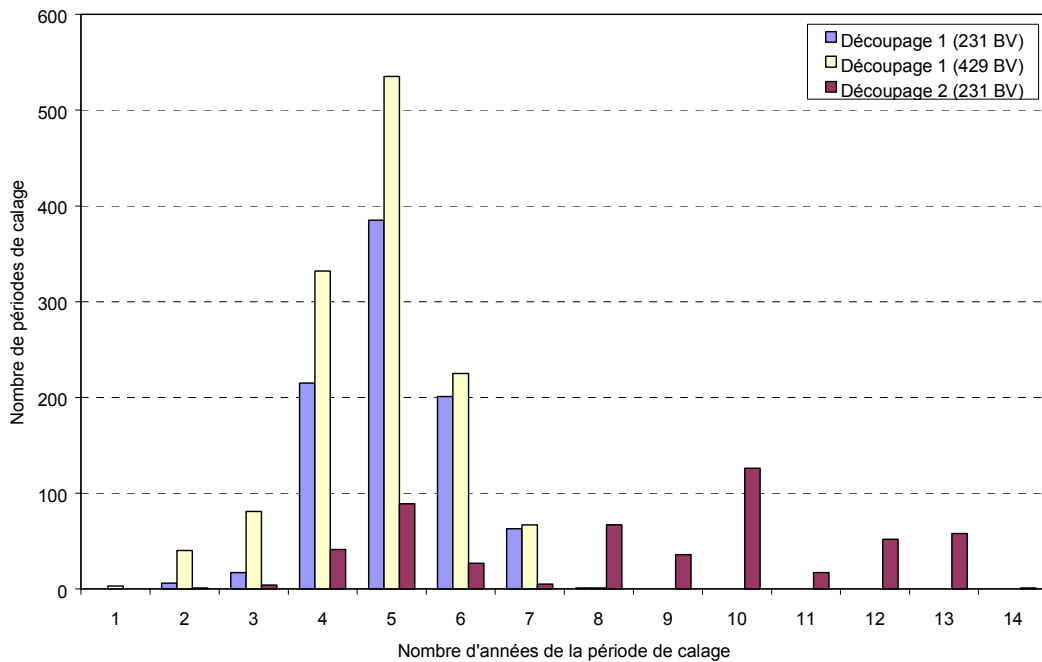
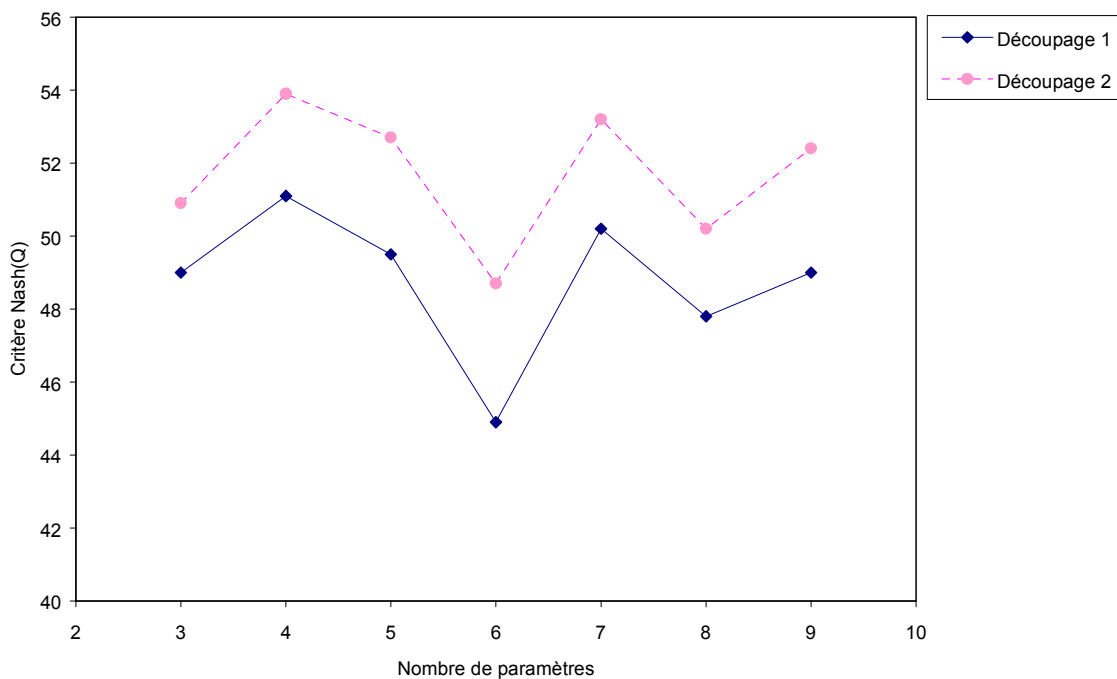
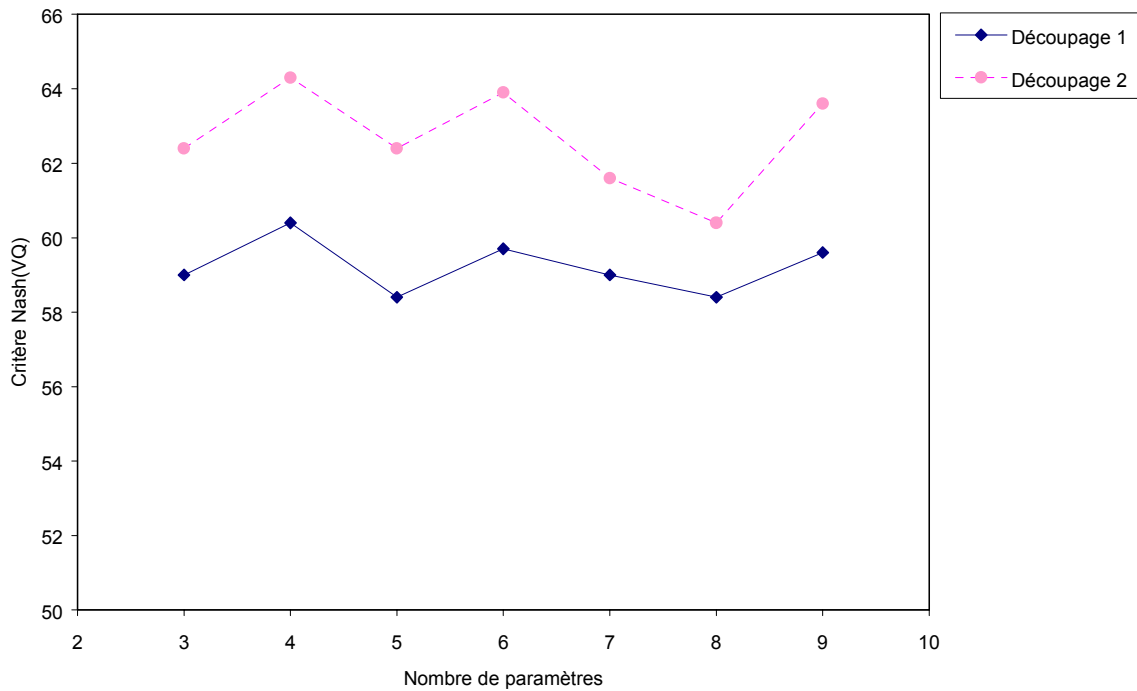


Figure 3.7: Répartition des périodes en fonction de leur longueur

Pour alléger les calculs, nous avons conservé sept modèles seulement, un pour chaque paramétrage entre trois et neuf paramètres. Les résultats moyens obtenus en contrôle sur l'échantillon des 231 bassins sont présentés à la Figure 3.8 pour deux critères de performance: le critère de Nash calculé sur les débits et le critère de Nash calculé sur les racines carrées des débits.



(a)



(b)

Figure 3.8: Performances des modèles en contrôle en fonction du nombre de paramètres pour (a) le critère Nash(Q), et (b) le critère Nash(\sqrt{Q})

Bien que les performances soient légèrement plus élevées pour les périodes plus longues, les différences de performances entre les modèles ne semblent pas être affectées par la longueur des séries de données: les changements de performances moyennes sont très similaires d'un modèle à un autre et semblent indépendants du nombre de paramètres du modèle. Ceci tendrait à montrer que le découpage initial en sous-périodes n'affecte pas les performances des modèles les plus complexes par un défaut d'information disponible lors du calage, et qu'en moyenne 5 années semblent suffisantes pour ne pas affecter la qualité de simulation des modèles complexes.

3.8. Conclusion

Nous avons présenté quelques réflexions générales sur la problématique de l'optimisation des paramètres de modèles pluie-débit. Il existe aujourd'hui de très nombreuses méthodes capables de gérer les difficultés induites par les caractéristiques fortement non-linéaires des modèles pluie-débit, parmi lesquelles on peut distinguer des méthodes locales qui explorent un voisinage d'un point de l'espace des paramètres, et les méthodes globales qui opèrent une recherche dans une zone beaucoup plus vaste de cet espace.

Les comparaisons existantes de différentes méthodes locales et globales de calage montrent que l'avantage des méthodes globales dans des cas théoriques (absence d'erreurs de données et de modélisation) ne se retrouve pas aussi nettement sur des cas concrets, du fait de la présence d'optima équivalents sur la surface de réponse. Dans la pratique, les techniques locales semblent donc autant en mesure de fournir des jeux de paramètres efficaces que les méthodes globales. Cependant, pour de telles méthodes, une bonne connaissance du modèle est indispensable pour bien choisir le point de départ de l'optimisation, qui influe sur les

résultats de la procédure de calage. De plus, il apparaît que les méthodes locales sont plus rapides.

La méthode 'pas-à-pas', méthode locale d'optimisation, a montré son efficacité dans le cas d'un modèle simple à quatre paramètres (Nascimento, 1995). Dans le cadre du présent travail, nous avons étudié l'influence du choix du point de départ sur les performances de la méthode. Notre approche a consisté à comparer les résultats obtenus en utilisant dans un premier temps un unique point de départ de l'optimisation puis dans un second temps deux points de départ conjointement. Cette analyse a permis de souligner l'importance de la paramétrisation des modèles et l'influence des problèmes d'identifiabilité des optima sur l'efficacité de la méthode de calage.

Le présent travail montre que les optima multiples sont essentiellement équivalents et que la recherche du meilleur d'entre eux, pour une période donnée, ne présente que peu d'intérêt: c'est généralement un autre jeu de paramètres qui donnera les meilleurs résultats sur une autre période. L'éventuelle multiplicité des optima n'aura de conséquences vraiment néfastes que lorsqu'il s'agira de rechercher des explications a priori des paramètres.

Lorsqu'un point de départ supplémentaire est utilisé, de légères améliorations des performances moyennes sont constatées au niveau du calage. Mais ces améliorations ne se répercutent pas systématiquement au niveau du contrôle. Ponctuellement, des améliorations significatives peuvent cependant être obtenues.

La sophistication de la méthode de calage par ajout d'un point de départ supplémentaire ne bouleverse en rien les résultats généraux de comparaison des modèles obtenus avec un seul point de départ pour le calage. Elle permet seulement de les affiner. Nous avons montré que les problèmes d'optimisation rencontrés étaient davantage dus à des problèmes internes du modèle qu'à de réelles faiblesses de la méthode d'optimisation, qui semble fiable à 90 % pour localiser des optima robustes. De plus la méthode est rapide et semble donc bien adaptée à la taille de l'échantillon test.

Par ailleurs, nous avons vérifié que le découpage des séries de données en sous-périodes relativement courtes (en moyenne cinq ans environ) n'affecte pas les performances de la méthode de calage. Il apparaît notamment que les modèles avec un plus grand nombre de paramètres ne souffrent pas d'un manque d'information par rapport aux modèles plus simples. Ainsi, la longueur de la série de calage ne semble pas être un facteur limitant la performance des modèles lorsque le nombre de paramètres augmente.

L'utilisation de la méthode 'pas-à-pas', méthode locale de calage, dans le cadre de la comparaison de structures de modèles pluie-débit ne nuit pas à la validité des conclusions générale que nous pourrions tirer au niveau de l'échantillon test, malgré les limites inhérentes à de telles procédures locales d'optimisation. La méthode paraît capable d'identifier des optima satisfaisants quelle que soit le modèle considéré ou le nombre de paramètres de ce dernier. Elle n'introduit donc pas de biais dans le test des modèles au détriment de certains d'entre eux.

Dans notre travail de test des modèles, nous veillerons cependant à bien choisir, pour chacun d'eux, le point de départ de l'optimisation, pour qu'il corresponde au centre approximatif de la zone de convergence la plus probable (il correspondra souvent au jeu de paramètres médians obtenus sur l'échantillon). Pour déterminer ce point de départ pour chaque modèle, nous avons procédé de manière itérative, en partant de valeurs initiales a priori et en affinant la localisation du point de départ en fonction des résultats d'une ou deux séries d'optimisations.

Chapitre 4

Chapitre 4

Sélection de critères d'évaluation des performances des modèles pluie-débit

4.1. Introduction

Une comparaison de modèles hydrologiques comme celle que nous voulons réaliser nécessite le choix de critères d'évaluation des performances. On peut en distinguer deux types. Les critères quantitatifs font appel à des évaluations numériques. Les critères qualitatifs, eux, s'appuient généralement sur des observations graphiques. Ces derniers critères sont largement utilisés en hydrologie (WMO, 1975; Weeks et Hebbert, 1980; Loague et Freeze, 1985; Chiew *et al.*, 1993; Michaud et Sorooshian, 1994; Refsgaard et Knudsen, 1996; Ye *et al.*, 1997) et sont, au cas par cas, des outils d'analyse appréciables pour juger de la qualité de simulation d'un modèle. Ils demeurent néanmoins des estimateurs subjectifs. Houghton-Carr (1999) montre notamment que deux personnes peuvent porter sur des mêmes illustrations graphiques (des hydrogrammes ou des courbes de débits classés par exemple) des jugements différents, attribuant aux modèles testés des valeurs relatives différentes. Il paraît donc difficile de faire reposer l'évaluation des performances de modèles uniquement sur de tels critères². La taille de notre échantillon test est un autre facteur limitant l'utilisation systématique de critères graphiques. Les critères quantitatifs ont donc été privilégiés dans notre travail comparatif et nous nous limiterons dans la suite de la discussion à cette classe de critères.

Certains auteurs, notamment Weglarczyk (1998), ont souligné qu'il n'existe pas de critère d'évaluation universel permettant de juger de la qualité d'un ajustement d'un modèle hydrologique. Ceci provient notamment du fait que les gammes de variation des événements hydrologiques sur un bassin sont souvent assez étendues (les débits pouvant varier typiquement d'un rapport de 1 à 100) et que chaque classe de débits ne présente pas le même intérêt pour l'utilisateur du modèle. Au même titre que le choix de la forme d'une fonction objectif pour le calage du modèle, celui de critères quantitatifs d'évaluation reste lié aux objectifs de l'utilisateur (prévision des crues, simulation des étiages, gestion de la ressource en eau, etc.).

L'évaluation du modèle demeure donc une opération partielle lorsqu'un seul critère d'appréciation est utilisé (Diskin et Simon, 1977; Sefè et Boughton, 1982). En mettant l'accent sur une caractéristique spécifique des hydrogrammes, il est alors difficile d'avoir un jugement global de la qualité des simulations (WMO, 1975; WMO, 1986; Houghton-Carr,

² Par ailleurs, la phase d'optimisation automatique des paramètres du modèle nécessite l'usage d'un critère numérique unique (dérivant éventuellement de plusieurs autres critères)

1999; Legates et McCabe, 1999). Si certains auteurs (par exemple Martinec et Rango, 1989) recommandent de ne pas multiplier de façon abusive le nombre de critères d'évaluation, l'utilisation de quelques critères complémentaires permet d'étendre le spectre des qualités évaluées du modèle (Legates et McCabe, 1999).

Les travaux comparatifs de test de modèles menés jusqu'à présent font appel à une large gamme de critères d'évaluation. Le Tableau 4.1 donne un aperçu des critères quantitatifs utilisés dans certaines de ces études. Nous avons sélectionné ici les études comparatives réalisées sur des modèles en fonctionnement continu. D'autres études comparatives (Sarma *et al.*, 1973; Singh, 1976; Loague et Freeze, 1985; Michaud et Sorooshian, 1994; Park *et al.*, 1999) se sont appuyées sur des évaluations par événement, avec des critères plus spécifiques à ce type de modélisation (tels que le temps de montée de l'hydrogramme ou la valeur du pic de crue). Les dix critères d'évaluation présentés dans le Tableau 4.1 sont assez couramment utilisés mais ne représentent que quelques-uns des dizaines de critères numériques proposés en modélisation hydrologique, tant au niveau du calage du modèle comme fonction objectif que lors d'évaluations a posteriori (voir notamment diverses formes de critères proposées par Fortin *et al.*, 1971; Diskin et Simon, 1977; Sefe et Boughton, 1982; Servat et Dezetter, 1991).

	Erreur quadratique	Erreur quadratique sur débits transformés	Critère de Nash-Sutcliffe	Coefficient de corrélation	Coefficient de détermination	Coefficient de variation des résidus	Erreur absolue	Erreur de bilan ou moyenne des débits	Indice sur la courbe des débits classés	Ecart-type ou variance
Moore et Mein (1975)	*			*				*	*	
WMO (1975)						*	*	*		
Weeks et Hebbert (1980)						*		*		
WMO (1986)			*			*	*	*		*
Moussavi et Feyen (1990)	*				*					
Wilcox <i>et al.</i> (1990)			*					*		
Franchini et Pacciani (1991)			*	*					*	
Vandewiele <i>et al.</i> (1992)										
Chiew <i>et al.</i> (1993)	*	*								
Reefsgard et Knudsen (1996)			*					*	*	
Zhang et Lindström (1996)			*	*				*		
Ye <i>et al.</i> (1997)			*				*	*		
Gan <i>et al.</i> (1997)	*		*					*		
Hugues et Metzler (1998)			*		*			*		*

Tableau 4.1: Critères d'évaluation utilisés dans des études comparatives de modèles pluie-débit.

Dans les paragraphes suivants, nous présentons les critères sélectionnés dans notre étude, qui sont construits à partir de trois types d'erreurs du modèle. Pour l'un de ces critères, des transformations préalables des débits ont également été utilisées. Une analyse des inconvénients de certains de ces critères nous a conduits à proposer de nouvelles formulations jugées plus satisfaisantes. Enfin, nous discutons l'intérêt d'adopter plusieurs critères d'évaluation des performances.

4.2. Première sélection de critères et transformation sur les débits

Comme nous l'avons mentionné précédemment, il existe de très nombreux critères. Dans notre travail, nous nous sommes ramenés à des critères numériques faisant appel à trois types de prise en compte de l'erreur du modèle: l'erreur quadratique, l'erreur absolue et l'erreur cumulée. Nous proposons ici une première sélection de trois critères faisant appel respectivement à ces trois erreurs.

4.2.1. Critère d'erreur quadratique

Les critères mesurent le degré d'adéquation entre les valeurs de la variable simulée par le modèle (ici le débit) et de la variable mesurée. Les critères en moindres carrés se basent sur une fonction suggérée par la régression linéaire (Nash et Sutcliffe, 1970), somme des erreurs quadratiques du modèle définie par:

$$F^2 = \sum_{i=1}^n (Q_{obs,i} - Q_{calc,i})^2 \quad \text{Eq. (4.1)}$$

où $Q_{obs,i}$ et $Q_{calc,i}$ sont respectivement les débits observé et simulé au pas de temps i , et n le nombre total de pas de temps de la période d'étude. F^2 est analogue à la variance résiduelle d'une régression linéaire. Il est cependant difficile en utilisant F^2 , de pouvoir comparer les performances du modèle d'une période à l'autre ou d'un bassin à l'autre, la valeur de cette fonction étant généralement d'autant plus élevée que le niveau des débits est fort (les erreurs absolues étant en moyenne plus élevées). Nash et Sutcliffe (1970) proposent donc de normer la fonction en effectuant une comparaison de F^2 avec ce même critère appliqué à un modèle élémentaire. Ils ont pris comme modèle $Q = cste$ et ont calé ce modèle élémentaire, ce qui donne $cste = \overline{Q_{obs}}$, la moyenne des débits observés. Pour ce modèle élémentaire, utilisé comme référence³, la fonction F^2 devient F_0^2 définie par:

$$F_0^2 = \sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2 \quad \text{Eq. (4.2)}$$

ce qui correspond à la variance des débits observés. Nash et Sutcliffe (1970) construisent ainsi un critère d'efficacité R défini comme la proportion de la variance initiale des débits expliquée par le modèle. Il est défini par:

$$R = 1 - F^2 / F_0^2 = 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{calc,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2} \quad \text{Eq. (4.3)}$$

Ce critère (que nous appellerons critère de Nash-Sutcliffe ou critère de Nash), variant dans l'intervalle $]-\infty;1]$ a l'avantage d'être d'interprétation facile. Il estime l'amélioration d'ajustement que l'on obtient en utilisant le modèle pour simuler les débits par rapport à un modèle 'zéro' (modèle de référence) qui donnerait sur toute la période considérée un débit constant égal au débit moyen. Un critère de valeur 1 signifie que l'erreur du modèle est nulle (modèle parfait). Une valeur de R inférieure à zéro signifie que le modèle n'explique pas mieux le comportement du bassin que le modèle de débit constant. Ce critère d'évaluation des

³ Dans la suite, nous désignerons par modèle de référence le modèle élémentaire utilisé dans l'expression de F_0^2 .

modèles est très largement utilisé en hydrologie, probablement en raison de sa simplicité et de son lien intrinsèque avec la régression linéaire.

Le rapport F/F_0 , que l'on peut désigner par ε , correspond en fait à l'erreur relative moyenne du modèle. Le critère de Nash est alors donné par $1-\varepsilon^2$. Lorsque ε vaut 0,2 par exemple, ceci signifie que l'erreur quadratique moyenne correspond à 20 % de l'écart-type des débits observés. Dans ce cas, le critère de Nash vaut 96 %. De même, pour une erreur relative de 40 %, le critère de Nash vaut encore 84 %. Ces valeurs élevées du critère pour des valeurs assez élevées de l'erreur relative donnent en fait une idée un peu optimiste de la qualité réelle du modèle. L'association abusive que l'on peut faire du critère de Nash à un coefficient de détermination tend à faire croire que le modèle est bon alors que ce n'est pas vraiment le cas. Une image plus réaliste de la qualité du modèle pourrait être donnée en prenant la racine carrée du critère. Il faut noter toutefois que les mesures en milieu naturel conduisent, par leur seule présence, à un ε de l'ordre de 0,2. Obtenir pour la transformation pluie-débit une erreur relative du même ordre de grandeur confirme bien que la qualité du modèle est très élevée. Ici nous nous en tiendrons à la formulation du critère de Nash classiquement adoptée en hydrologie.

4.2.2. Critère d'erreur absolue

L'erreur moyenne absolue A d'un modèle peut être définie par :

$$A = \frac{1}{n} \sum_{i=1}^n |Q_{obs,i} - Q_{calc,i}| \quad \text{Eq. (4.4)}$$

Elle représente la déviation absolue du débit simulé par rapport à l'observé en moyenne à chaque pas de temps. Ce critère peut être intéressant dans le contexte de la prévision de débit, où l'on veut être aussi proche que possible de la valeur observée à chaque pas de temps. L'avantage de cette erreur par rapport à l'erreur quadratique est qu'elle n'accorde pas un poids proportionnellement plus important aux fortes erreurs. A l'instar de l'erreur quadratique définie à l'Eq. (4.1), A n'est cependant pas normée et ne permet pas une comparaison aisée des performances du modèle d'une période à l'autre ou d'un bassin à l'autre. Pour cette raison, nous pouvons utiliser une forme équivalente à la forme du critère de Nash, définie par:

$$AD = 1 - \frac{\sum_{i=1}^n |Q_{obs,i} - Q_{calc,i}|}{\sum_{i=1}^n |Q_{obs,i} - \overline{Q_{obs}}|} \quad \text{Eq. (4.5)}$$

AD est un critère relatif. Il varie ainsi également entre $]-\infty;1]$ et permet de réaliser des moyennes sur les performances des modèles sur un échantillon de bassins. En fait, le critère de Nash et le critère AD pourraient être rapprochés de la forme plus générale d'une norme N donnée par:

$$N = 1 - \left[\frac{\sum_{i=1}^n |Q_{obs,i} - Q_{calc,i}|^m}{\sum_{i=1}^n |Q_{obs,i} - \overline{Q_{obs}}|^m} \right]^{1/m} \quad \text{Eq. (4.6)}$$

où m est un exposant entier. AD est la norme d'ordre 1 et le critère de Nash est fondé sur la norme d'ordre 2. Lorsque m devient très grand, on retrouve la norme de Tchebychev représentée par le rapport des écarts maximaux.

4.2.3. Critère d'erreur cumulée

L'erreur moyenne cumulée EC d'un modèle peut être définie par :

$$EC = \frac{1}{n} \sum_{i=1}^n (Q_{obs,i} - Q_{calc,i}) \quad \text{Eq. (4.7)}$$

$$= \overline{Q_{obs,i}} - \overline{Q_{calc,i}}$$

Elle indique, par comparaison des débits moyens observés et calculés, la capacité du modèle à reproduire le volume d'eau total observé sur la période étudiée. Egalement appelée erreur de bilan ou biais du modèle, cette erreur peut être positive ou négative suivant que le modèle surestime ou sous-estime les flux sur la période. A la différence des deux erreurs précédentes, celle-ci ne fait plus référence à une adéquation temporelle à chaque pas de temps entre le débit calculé et le débit observé (de la même façon que certains critères graphiques tels que la courbe des débits classés éludent cette référence temporelle). Elle ne peut donc être utilisée comme seul critère d'optimisation, un débit constant égal au débit moyen donnant un critère parfait. Cette erreur peut aussi être écrite sous forme relative, par :

$$B = \frac{\sum_{i=1}^n Q_{calc,i}}{\sum_{i=1}^n Q_{obs,i}} \quad \text{Eq. (4.8)}$$

ou
$$B' = \frac{\sum_{i=1}^n Q_{calc,i}}{\sum_{i=1}^n Q_{obs,i}} - 1 \quad \text{Eq. (4.9)}$$

Une valeur de 1 pour B ou 0 pour B' indiquera un bilan parfait. Une valeur supérieure à 1 pour B ou supérieure à 0 pour B' indiquera une surestimation du bilan. Ce critère permet de comparer les performances du modèle d'un bassin à l'autre ou d'une période à l'autre. Cependant, cette formulation présente quelques inconvénients que nous détaillerons dans la suite de ce chapitre. Remarquons que s'il ne peut être utilisé seul en optimisation, certains auteurs comme Lindström *et al.* (1997) et Quesney (1999) ont en revanche utilisé ce critère combiné au critère de Nash sous la forme $R - \alpha |B'|$, α étant un coefficient de l'ordre de 0,1 ou 0,2. Cette pondération, opération rendue possible par le fait que les critères sont adimensionnels, permet d'attribuer un poids plus important à la bonne simulation des bilans lors de l'optimisation, sans pour autant revêtir un sens physique particulier.

4.2.4. Transformations préalables des débits

Un des avantages des trois critères R , AD et B (ou B') réside dans le fait que l'on obtient une appréciation globale de la performance du modèle sur l'ensemble de la période. D'autres critères s'intéressent à des classes particulières de débit, tel que l'indice de débit de base ou Base Flow Index (BFI) (Gustard *et al.*, 1992) ou des critères spécifiques sur les crues ou les étiages tels que ceux proposés par Nascimento (1995).

Certains critères accordent, par la nature de leur formulation, une importance relative plus grande à certaines classes de débit. Une solution pour donner plus de poids à telle ou telle classe de débits dans un critère global est d'utiliser des transformations préalables sur les débits. Ceci provient du fait que les résidus du modèle ne sont généralement pas homoscedastiques, c'est-à-dire que leur variance n'est pas indépendante de la valeur du débit. Ainsi, les erreurs sur les forts débits sont généralement plus élevées que celles sur les faibles

débits. Cette caractéristique (ainsi que l'autocorrélation) des erreurs des modèles en hydrologie a été la source d'importantes recherches, notamment en lien avec des problèmes d'optimisation des paramètres. Certains auteurs (Sorooshian et Dracup, 1980; Kuczera, 1983) ont notamment montré que les résidus ne répondaient pas aux hypothèses d'utilisation des critères simples en moindres carrés dans la théorie du maximum de vraisemblance. Ils ont alors proposé des critères (e.g. Heteroscedastic Maximum Likelihood Estimator – HMLE) permettant d'obtenir des résidus répondant mieux à ces hypothèses, critères utilisant des transformations préalables sur les débits telles que celles proposées par Box et Cox (1964) (transformations logarithmiques et puissance). Ces critères sont largement utilisés en optimisation (voir par exemple Ibbitt et Hutchinson, 1984; Gan *et al.*, 1997; Abdulla *et al.*, 1999). Nascimento (1995) a également testé l'effet de l'utilisation de ces fonctions objectifs sur la détermination des paramètres du modèle GR4, montrant qu'elles sont particulièrement intéressantes lorsque l'on cherche à reproduire certaines caractéristiques spécifiques des écoulements.

Les transformations préalables sur les débits que nous avons utilisées n'ont cependant pas été choisies avec pour premier objectif d'avoir des résidus en adéquation avec des hypothèses mathématiques. Nous nous sommes plutôt placés d'un point de vue de l'utilisateur qui choisit un critère (de calage et/ou d'évaluation) en fonction de ses objectifs. Certaines transformations choisies, néanmoins, diminuent en conséquence l'hétéroscédasticité des résidus. Les transformations choisies ont été utilisées conjointement au critère de Nash précédemment décrit.

Une caractéristique de ce critère est de donner une importance prépondérante aux périodes de hautes eaux. Pour avoir un critère donnant un poids plus grand aux périodes de basses eaux, nous avons utilisé une première transformation logarithmique sur les débits (auxquels on ajoute une constante faible – par exemple $\text{Module}/40$ – pour éviter des problèmes numériques dans le cas de débits nuls). Cette transformation a été utilisée par exemple par Ambroise *et al.* (1995). Elle nivelle les valeurs des débits, les erreurs du modèle variant alors dans un même ordre de grandeur pour toutes les classes de débits. Les périodes de hautes eaux étant généralement plus courtes que celles de basses eaux, on accorde ainsi une importance prépondérante à ces dernières.

La deuxième transformation utilisée est une transformation puissance (transformation avec une puissance $\frac{1}{2}$, c'est-à-dire en racine carrée), qui permet d'avoir un critère intermédiaire entre le critère de Nash calculé sur les débits et celui calculé sur les logarithmes des débits. Chiew *et al.* (1993) ont utilisé une telle transformation puissance (avec une puissance $1/5$). C'est ce critère que nous utiliserons comme fonction objectif pour l'optimisation des paramètres des structures de modèles lors de la comparaison.

4.2.5. Relations entre critères

La Figure 4.1 montre que l'appréciation des performances d'un modèle sur un bassin peut être tout à fait différente lorsque la formulation du critère change, avec différentes formes analytiques de l'erreur ou différentes prises en compte de la variable cible. S'il semble y avoir une relative corrélation entre les différents critères, il est en revanche possible de trouver des cas où le modèle est jugé très satisfaisant suivant un critère et beaucoup moins suivant un autre.

Ces remarques renforcent la nécessité d'avoir plusieurs critères d'évaluation des performances permettant de juger les simulations sur différents aspects.

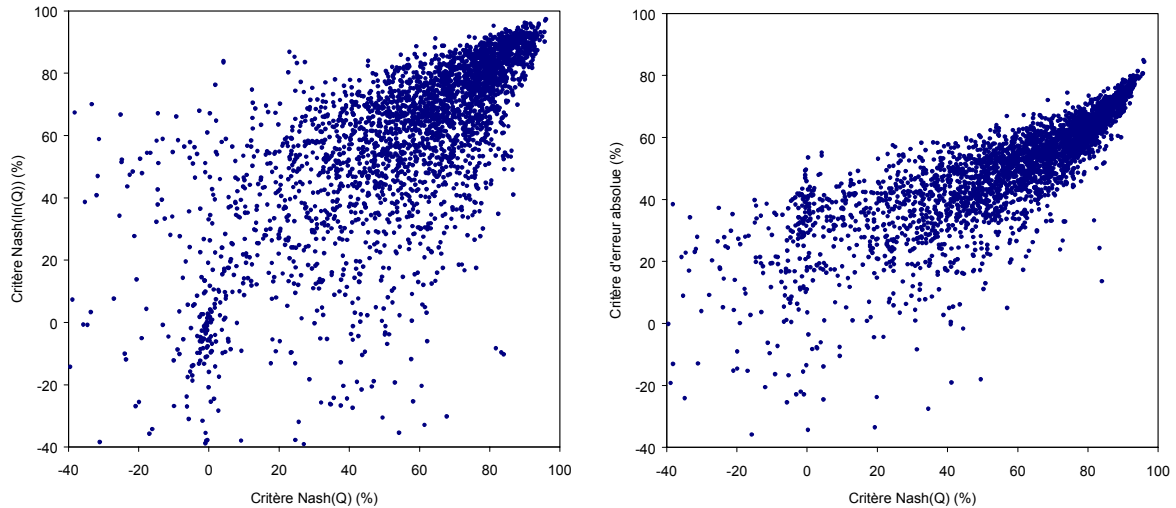


Figure 4.1: Critères obtenus par le modèle GR4K en contrôle sur l'échantillon de bassins après calage en utilisant le critère $\text{Nash}(\sqrt{Q})$: comparaison (a) des critères $\text{Nash}(Q)$ et $\text{Nash}(\ln(Q))$, et (b) du critère $\text{Nash}(Q)$ et du critère d'erreur absolue défini à l'Eq. (4.5)

4.3. Critère de Nash: analyse et proposition d'une nouvelle formulation

4.3.1. Quelques réflexions sur le critère de Nash-Sutcliffe

Les critères en moindres carrés sont parmi les plus utilisés en modélisation pluie-débit. Certains ont fait l'objet de critiques contestant leur applicabilité au contexte hydrologique. Par exemple Willmott (1981) remet en cause l'utilisation du coefficient de corrélation: des variables observées et simulées peuvent être parfaitement corrélées, mais avec une surestimation systématique des sorties modélisées par exemple. Nous allons dans cette partie examiner un peu plus en détail le critère en moindres carrés que nous avons choisi, le critère de Nash-Sutcliffe. S'il est très utilisé en modélisation, ce critère a cependant fait l'objet de réflexions soulignant dans certains cas des faiblesses dans sa pertinence.

Martinec et Rango (1989), à la suite de l'intercomparaison de modèles réalisée par l'Organisation Mondiale de la Météorologie (WMO, 1986), ont montré que le choix du débit moyen de la période étudiée comme modèle de référence n'est pas toujours adapté pour évaluer la qualité d'un ajustement, notamment pour les périodes de comportement extrême (périodes beaucoup plus sèches ou plus humides qu'en moyenne). Dans ces cas, le module interannuel des débits peut être préféré. Nous pouvons donner un exemple tiré de notre échantillon, dans le cas du bassin versant du Loing à Chalette-sur-Loing sur la période 1989-1992. Le module interannuel sur la période 1976-1993 est de 0,502 mm/j et seulement 0,172 mm/j sur 1989-1992, période particulièrement sèche (voir illustration à la Figure 4.2). En prenant le module interannuel comme modèle de référence, le critère de Nash obtenu sur 1989-1992 avec GR4K⁴ est de 80,5 % alors qu'il est seulement de 51,3 % lorsque l'on prend le débit moyen de la période comme modèle de référence. Par comparaison, le modèle obtient un critère de Nash de 88,9 % sur la période 1983-1988 plus arrosée.

⁴ Dans toute notre réflexion sur les critères, nous utiliserons les résultats du modèle GR4K pour illustrer notre démarche

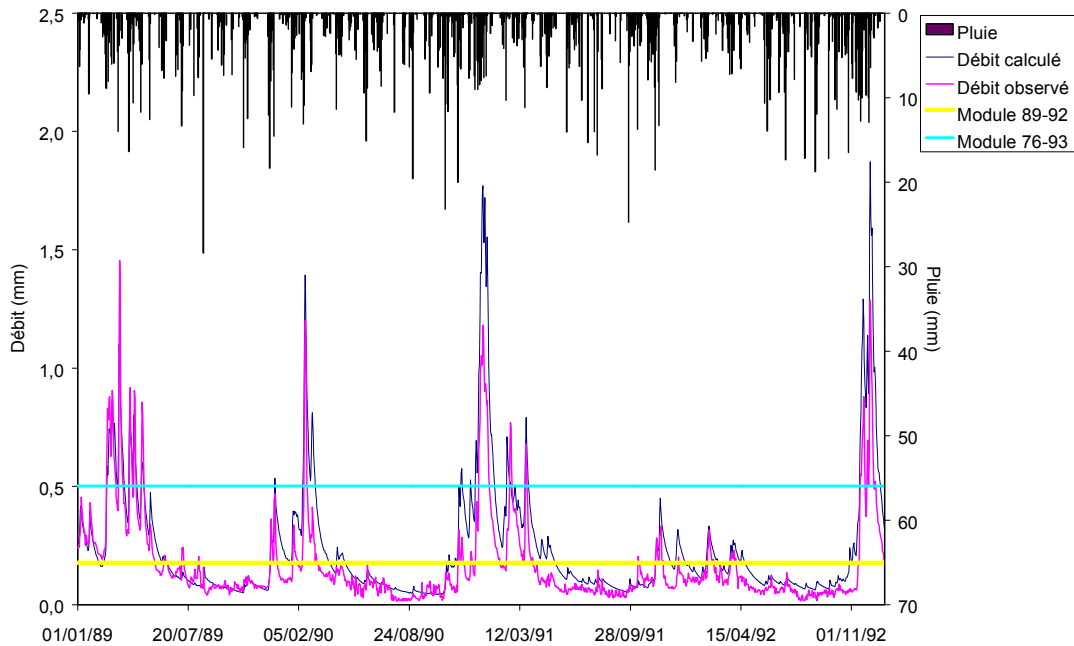


Figure 4.2: Hydrogrammes observés et simulés sur la période 1989-1992 pour le bassin versant du Loing à Chalette-sur-Loing

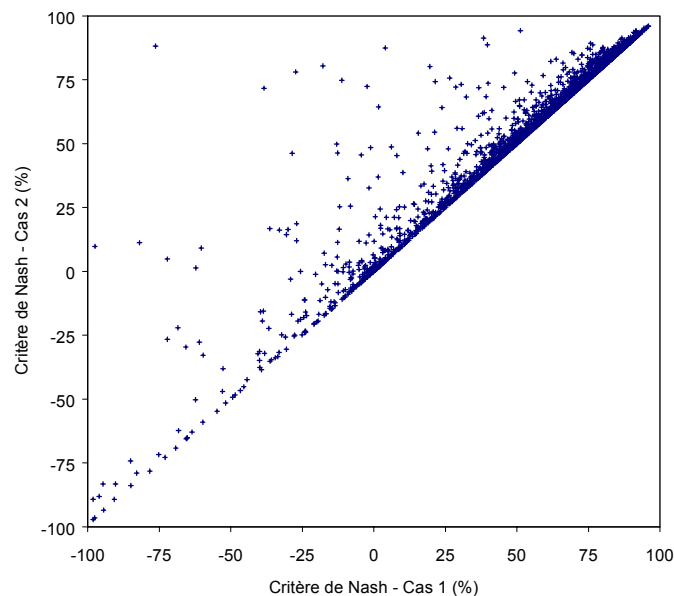


Figure 4.3: Comparaison des valeurs du critère de Nash avec comme modèle de référence le débit moyen de la période considérée (cas 1) et le débit moyen de la période de calage (cas 2) sur les 3204 tests en contrôle

De façon similaire, on peut prendre en compte comme modèle de référence dans le critère de Nash non plus le débit moyen de la période considérée mais le débit moyen de la période de calage. Ceci revient à dire que l'on cale le modèle de référence sur la période utilisée pour le calage des paramètres et que l'on conserve ensuite la valeur de ce modèle de référence quelque soit la période utilisée. La Figure 4.3 montre que, s'il y a globalement accord entre les deux formulations, des simulations sur certains bassins-périodes peuvent être jugées médiocres suivant le critère de Nash classique (cas 1) et assez satisfaisantes suivant le critère modifié (cas 2).

L'évaluation relative (la 'note') des performances du modèle dépend donc clairement de la référence choisie, c'est-à-dire de l'amélioration de la qualité de simulation demandée au modèle par rapport à l'utilisation d'un modèle très simplifié (équivalent à une hypothèse 'zéro'). Remarquons que dans le contexte comparatif dans lequel nous nous plaçons, l'utilisation de critères relatifs plutôt qu'absolus est indispensable pour pouvoir comparer les performances d'un bassin à l'autre. De manière plus générale, une erreur relative est un objectif global plus aisé à prendre en compte dans un contexte opérationnel: on ne peut pas par exemple viser une même précision en valeur absolue sur des débits d'étiage ou des débits de crue.

Ce problème se retrouve également dans le cas des bassins intermittents où le critère de Nash prend très facilement des valeurs fortement négatives. Cette tendance est illustrée à la Figure 4.4 où l'on a tracé les distributions des critères de Nash obtenus en contrôle par le modèle GR4K (i) sur l'échantillon total de bassins, (ii) sur les 15 % des bassins à plus faible rendement (majoritairement intermittents) et (iii) sur les 15 % des bassins à plus fort rendement (le rendement Cr étant défini ici comme le rapport entre lame d'eau écoulée et pluie cumulée). Le modèle a des performances nettement moins élevées sur les bassins à faible rendement, avec beaucoup de valeurs inférieures à zéro. De la même façon, nous avons représenté sur la Figure 4.5 les distributions des critères de Nash obtenus en contrôle sur les cinq classes de bassins établies au chapitre 2. Pour les classes 4 et 5, qui regroupent les bassins les moins arrosés et avec les plus faibles rendements, il y a d'importantes proportions de critères négatifs. Même si un modèle peut s'avérer être moins performant sur ce type de bassins, l'obtention de critères négatifs très grands en valeur absolue, sans justification évidente, est un obstacle difficilement surmontable à un calcul de moyenne de ce critère sur plusieurs bassins versants. L'interprétation de critères négatifs reste mal-aisée. En fait, nous allons voir dans ce qui suit que la forte proportion de critères négatifs tient en partie du fait que le modèle de référence choisi correspond au débit moyen observé.

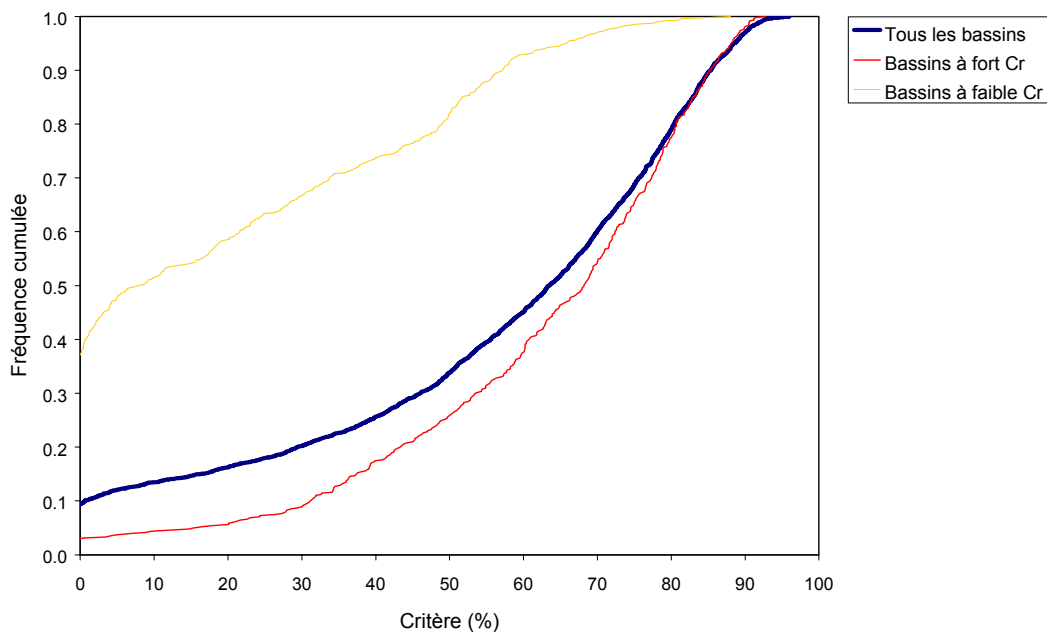


Figure 4.4: Distributions des critères de Nash (%) obtenus en contrôle par le modèle GR4K sur l'échantillon de 429 bassins, avec distinction des bassins à faible et fort rendements Cr .

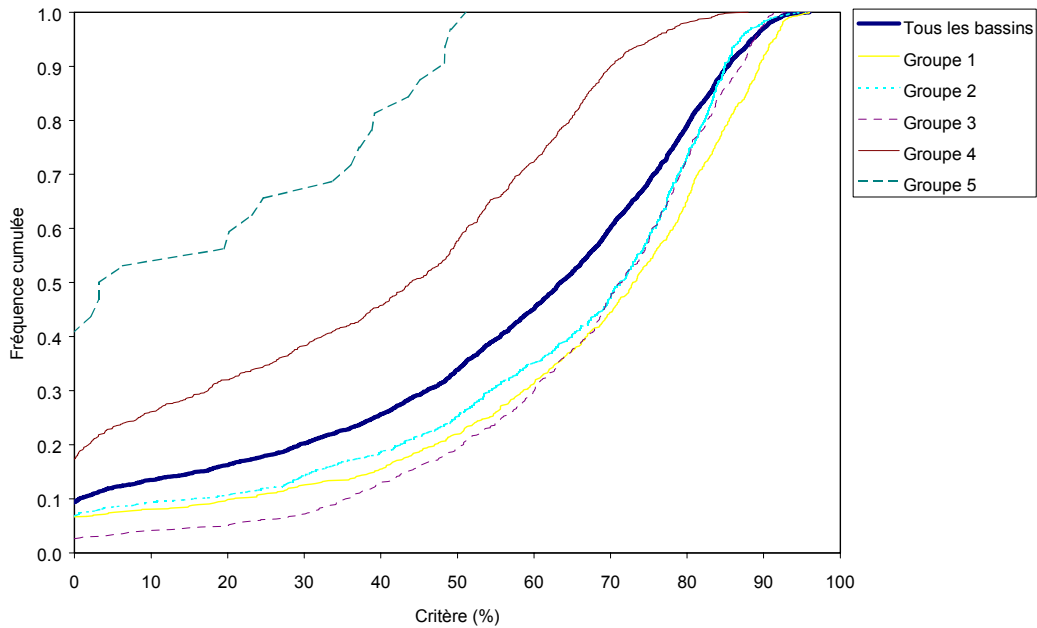


Figure 4.5: Distributions des valeurs du critère de Nash obtenues par le modèle GR4K sur les cinq classes de bassins en contrôle

Ceci illustre le problème de la subjectivité du choix du modèle de référence, qui semble dans certains cas être trop exigeant vis-à-vis du modèle. Inversement, ce modèle basique apparaît trop 'primitif' pour Garrick *et al.* (1978) ou Martinec et Rango (1989) qui ont jugé l'obtention de critères élevés trop aisée. Garrick *et al.* (1978) ont alors suggéré de mieux exploiter la connaissance initiale du bassin et ont proposé comme alternative dans le choix du modèle de référence de prendre pour chaque jour le débit moyen interannuel de ce jour, modèle qui ajoute au débit moyen constant une composante de saisonnalité. Il est ainsi plus difficile d'obtenir des critères élevés avec ce nouveau modèle de référence, meilleur simulateur des débits observés que le débit moyen constant. Cependant, en contrepartie, l'utilisation de ce critère entraîne l'obtention de nombreuses valeurs négatives, dont l'interprétation reste peu aisée.

Il semble donc difficile de choisir un modèle de référence pertinent pour l'ensemble des bassins, qui permette d'une part d'avoir une illustration objective de la performance d'un modèle sur un bassin particulier et d'autre part de comparer les performances d'un modèle obtenues sur différentes périodes ou différents bassins.

4.3.2. Objectifs de la mise au point d'un nouveau critère

L'objectif principal sous-tendant la recherche d'une nouvelle formulation d'un critère est de trouver un critère pour lequel il y aurait peu de valeurs négatives, et en tout état de cause pas de valeurs très grandes en valeur absolue, car cette détérioration du critère ne correspond à rien de réel. Par ailleurs, elle peut biaiser les moyennes calculées sur plusieurs bassins, particulièrement importantes dans notre étude comparative pour avoir une appréciation d'ensemble des performances des modèles. Il serait donc intéressant d'avoir un critère qui prenne moins facilement des valeurs très négatives et qui permette ainsi d'assurer une bonne significativité des moyennes calculées sur plusieurs tests.

La Figure 4.6 illustre l'évolution du critère de Nash en fonction de la qualité 'réelle' des simulations du modèle, dont l'échelle serait entre 0 et 1 (la définition de cette qualité 'réelle'

reste ici théorique et n'est utilisée qu'à des fins d'illustration). Lorsque la qualité 'réelle' diminue, le critère de Nash tend vers une asymptote et s'écarte de plus en plus d'une échelle d'évolution linéaire en fonction de la qualité réelle. Il en résulte une trop forte sensibilité du critère: une faible diminution de qualité 'réelle' induit alors de très fortes chutes du critère de Nash, dont la valeur n'a plus alors de réelle signification. Par conséquent, il est beaucoup plus difficile de constater une amélioration des hydrogrammes simulés lorsque que l'on passe par exemple d'un critère de Nash de -500% à -100% que lorsque l'on passe de 70% à 90% . Nous cherchons donc à nous rapprocher d'une évolution plus 'linéaire' du critère de Nash en fonction de la qualité 'réelle' et à limiter l'existence de valeurs fortement négatives.

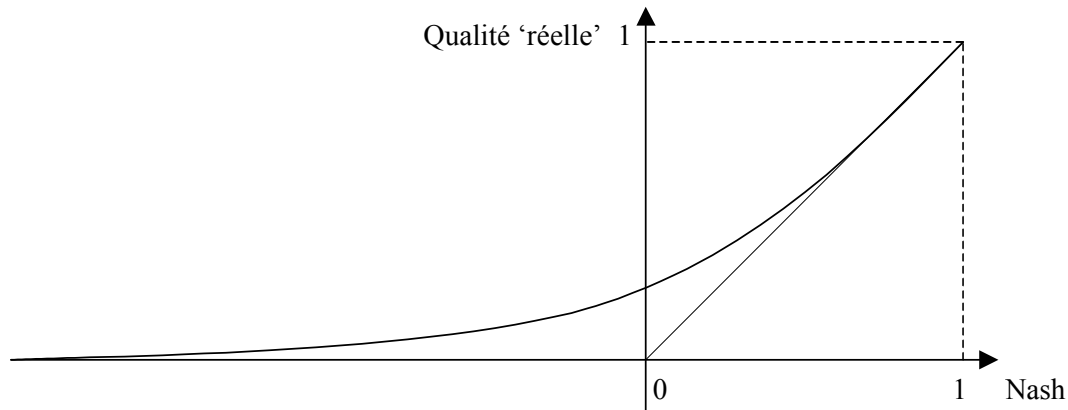


Figure 4.6: Comportement schématique du critère de Nash en fonction de la qualité 'réelle' des simulations du modèle

Il serait intéressant, par ailleurs, de trouver un critère qui soit le moins dépendant possible du type de bassin étudié. En effet, une trop forte dépendance de la valeur du critère au type de bassin favoriserait artificiellement certains bassins, donnant la fausse impression que le modèle est mieux adapté à certains types de bassins qu'à d'autres. Il devient dans ce cas impossible de comparer les performances du modèle d'un bassin à un autre en raison de la formulation inadéquate du critère d'évaluation. Il paraît donc commode d'avoir un critère pour lequel il y ait statistiquement à peu près autant de chances d'obtenir des valeurs de critère élevées (ou faibles) quel que soit le type de bassins auquel on s'intéresse.

Enfin, il faut remarquer que le modèle de référence utilisé dans le critère de Nash n'est pas réellement un modèle puisqu'il est calculé à partir du résultat attendu et non à partir des entrées du modèle pluie-débit. Prendre en compte ces entrées permettrait donc d'avoir un réel modèle de référence.

Les objectifs sous-tendant la mise au point d'un nouveau critère ne doivent pas occulter le fait qu'il n'existe pas de critère universel, comme nous l'avons mentionné en introduction de ce chapitre. Nous voulons donc ici rechercher une formulation qui minimise les problèmes évoqués précédemment, sans la considérer pour autant comme universelle.

4.3.3. Analyse de la problématique

Pour parvenir à proposer une solution satisfaisante à la question soulevée précédemment, nous pouvons tout d'abord essayer de répondre à la question suivante: quelle doit être la qualité requise essentielle d'un critère destiné à évaluer les performances d'un modèle de simulation des débits? La qualité essentielle est de pouvoir rendre compte de la similitude entre hydrogrammes observés et simulés. Cela est rendu possible grâce à l'utilisation de l'erreur du modèle (absolue ou quadratique) vue aux paragraphes précédents. Pour rendre cette erreur interprétable, il faut choisir un modèle de référence (hypothèse 'zéro'). Ce dernier est

construit, comme les modèles pluie-débit étudiés, à partir de fonctions agissant d'une part sur le rendement (calcul des volumes écoulés), et d'autre part sur le transfert (distribution dans le temps de ces volumes). C'est l'association de ces deux composantes qui permet au modèle de simuler les débits, le degré de satisfaction étant fonction à la fois de la qualité du module de rendement et de celle du module de routage. Une deuxième qualité du critère d'évaluation serait donc d'utiliser un modèle de référence qui traite de façon assez homogène ces deux fonctions dans le degré de vraisemblance qu'il leur accorde. Remarquons aussi qu'il est préférable que ce modèle de référence ne soit pas calé.

Au niveau du rendement, le modèle de référence à choisir se situe entre un opérateur 'identité', qui donnerait en sortie un volume de pluie nette égal au volume des pluies tombées (rendement du bassin égal à un), et un opérateur 'zéro' qui donnerait en sortie un débit nul (rendement nul du bassin). Le modèle de référence parfait donne exactement le débit moyen observé.

Au niveau du routage, le modèle de référence à choisir se situe entre un opérateur 'instantané' qui restituerait immédiatement la pluie efficace, et un opérateur 'infini' qui nivellerait les débits au point de ne plus avoir de référence temporelle, c'est-à-dire qui donnerait une constante (égale à la moyenne de la pluie nette). Ici, l'opérateur parfait serait celui qui permettrait de restituer un hydrogramme exactement proportionnel à l'hydrogramme observé.

Ces domaines de variations des opérateurs de transformation sont illustrés à la Figure 4.7. Dans le cas du critère de Nash, la référence étant le débit moyen sur la période considérée, on se place dans le cas d'un opérateur idéal au niveau du rendement (qui donne les volumes exacts observés) et dont le routage serait assuré par un opérateur 'infini'. Il s'agit donc d'un modèle parfait pour le rendement et très mauvais pour le routage dans le contexte d'une modélisation pluie-débit (un tel routage peut néanmoins être relativement satisfaisant dans le cas de cours d'eau alimentés par des nappes). On compare donc les performances du modèle, qui est une représentation simplifiée du bassin réel, à un autre bassin hypothétique qui donnerait tout au long de l'année un débit constant indépendamment de la répartition temporelle de la pluie.

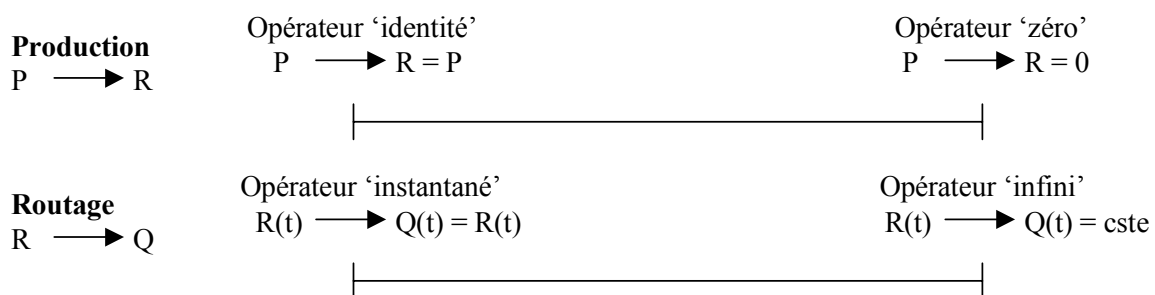


Figure 4.7: Schéma des plages de variation des opérateurs de production et de routage pour le choix d'un modèle de référence (P: pluie brute; R: pluie nette; Q: débit; cste: constante égale à la pluie nette moyenne)

Pour essayer de pallier les inconvénients liés à l'utilisation de ce critère, nous avons cherché à utiliser d'autres combinaisons d'opérateurs de rendement et de routage pour le modèle de référence. Pour le rendement, nous avons utilisé diverses possibilités. Il nous a paru intéressant d'essayer de faire intervenir les entrées du modèle (pluie et ETP). En effet, jusqu'à présent, les modèles de référence sont toujours construits sur le débit observé, jamais sur les entrées du système. La prise en compte de ces entrées permet indirectement de tenir compte de la transformation opérée par le bassin pour donner les débits et donc du travail demandé au

modèle sur ces entrées. On pourra notamment explorer la possibilité d'utiliser un opérateur proche de l'opérateur 'identité'. Les opérateurs de rendement suivants ont été utilisés:

$$\begin{array}{ll}
 P & \longrightarrow & R = 0 \\
 P & \longrightarrow & R = \alpha \cdot P \\
 P & \longrightarrow & R = \alpha \cdot \max(0, P - E) \\
 P & \longrightarrow & R = \alpha \cdot \left(\sqrt{P^2 + E^2} - E \right)
 \end{array}$$

où P représente la pluie, E l'évapotranspiration potentielle et α un coefficient inférieur à 1. Les deux derniers opérateurs sont des combinaisons simples des deux variables d'entrée. Ils ont été suggérés par les travaux en cours de Mouelhi (2000) sur le modèle pluie-débit interannuel. Ce sont des modèles d'estimation du module interannuel des débits à partir de la pluie et de l'ETP. La quatrième formulation (sans le coefficient α) correspond au modèle proposé par Tixeront (1964). La Figure 4.8 montre que ce modèle est un relativement bon estimateur du module interannuel.

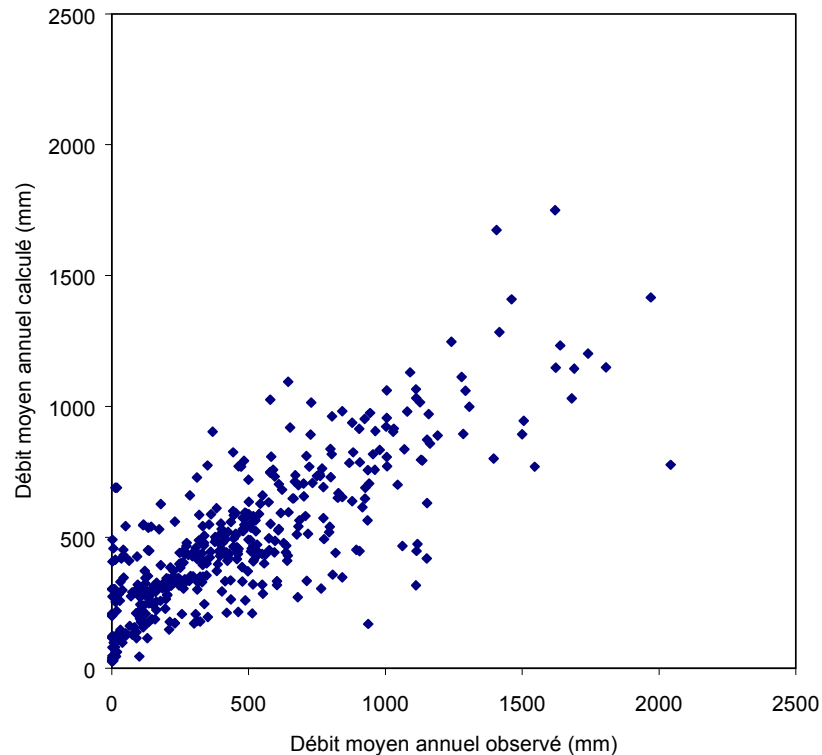


Figure 4.8: Comparaison des modules observés et simulés d'après le modèle de Tixeront (1964) pour les 429 bassins versants (d'après Mouelhi, 2000)

Pour l'opérateur de routage, nous nous sommes limités aux deux opérateurs 'instantané' et 'infini', d'autres opérateurs étant trop complexes et/ou trop arbitraires pour servir de référence.

4.3.4. Evaluation de différents critères et des modèles de référence associés

Les sept critères testés ici sont tous construits à partir de la même forme analytique de l'erreur du modèle, erreur quadratique utilisée dans le critère de Nash. Des modèles testés sur une période obtiendront donc avec l'un de ces sept critères le même classement qu'avec le critère

de Nash, puisque le numérateur, somme des carrés des erreurs, n'est pas modifié d'un critère à l'autre. Seul le modèle de référence change ici.

Critère 1 (C1)

Une première possibilité est de prendre comme référence une fraction de la pluie du jour considéré. Cela revient à considérer le modèle de référence comme un bassin produisant instantanément une fraction constante de la pluie tombée.

La formule de ce critère C1 s'écrit:

$$C1 = 1 - \left(\frac{\sum_{i=1}^n (Q_{obs,i} - Q_{calc,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \alpha \cdot P_{obs,i})^2} \right)^{\beta} \quad \text{Eq. (4.10)}$$

où $P_{obs,i}$ est la pluie observée au jour i et α et β sont deux coefficients à déterminer. L'exposant β , inférieur à 1, permet d'obtenir des ordres de grandeurs similaires à ceux du critère de Nash.

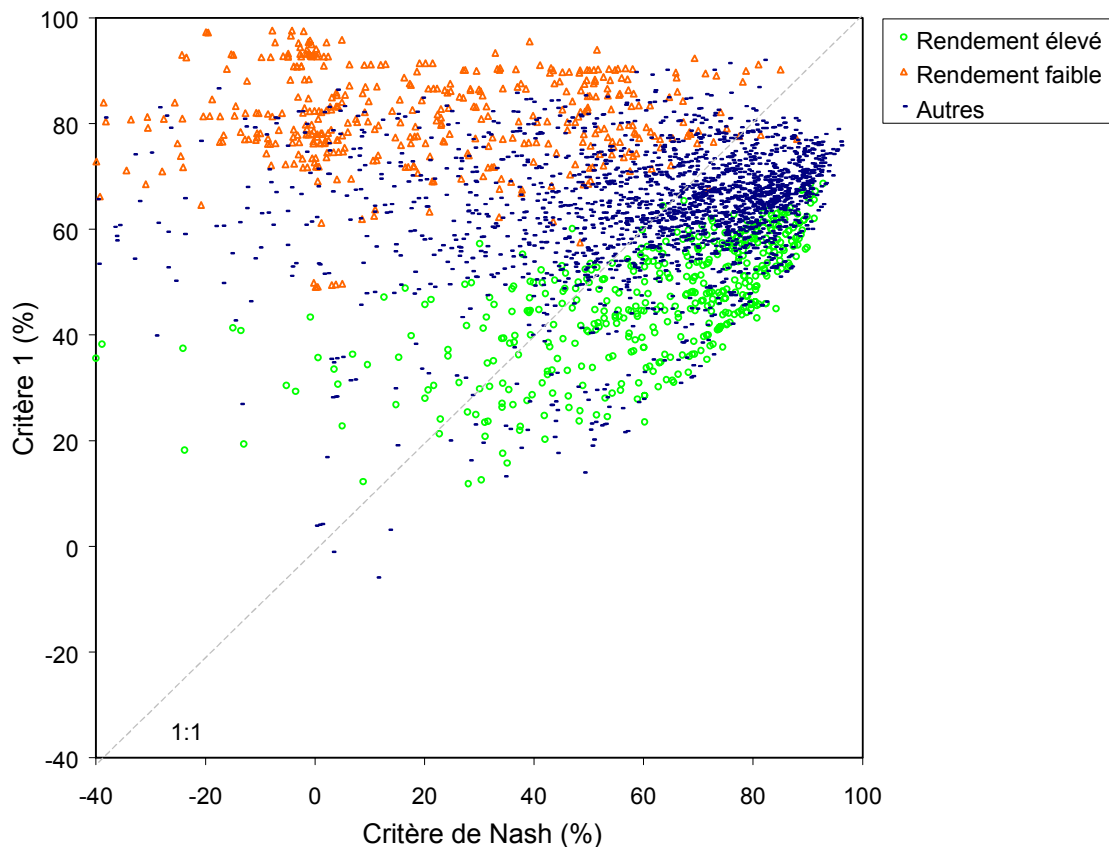


Figure 4.9: Comparaison des valeurs du critère C1 et du critère de Nash sur l'échantillon de bassins en contrôle ($\alpha = 1$ et $\beta = 1/4$)

L'avantage du critère C1, comme le montre la Figure 4.9, est que la proportion de contrôles négatifs devient négligeable en comparaison du critère de Nash. Cependant, lorsque α est voisin de 1, un problème inverse à celui du critère de Nash se pose, avec des bassins à faible rendement obtenant de bien meilleurs résultats que les bassins à fort rendement (voir Figure 4.10). En effet, le critère C1 est fortement influencé par le rendement du bassin, des bassins à faible rendement ayant tendance à avoir des valeurs de critère élevées lorsque α est proche

de 1. Pour ces bassins, les valeurs élevées proviennent en partie du fait de la prédominance de la pluie par rapport au débit au dénominateur de l'Eq. (4.10).

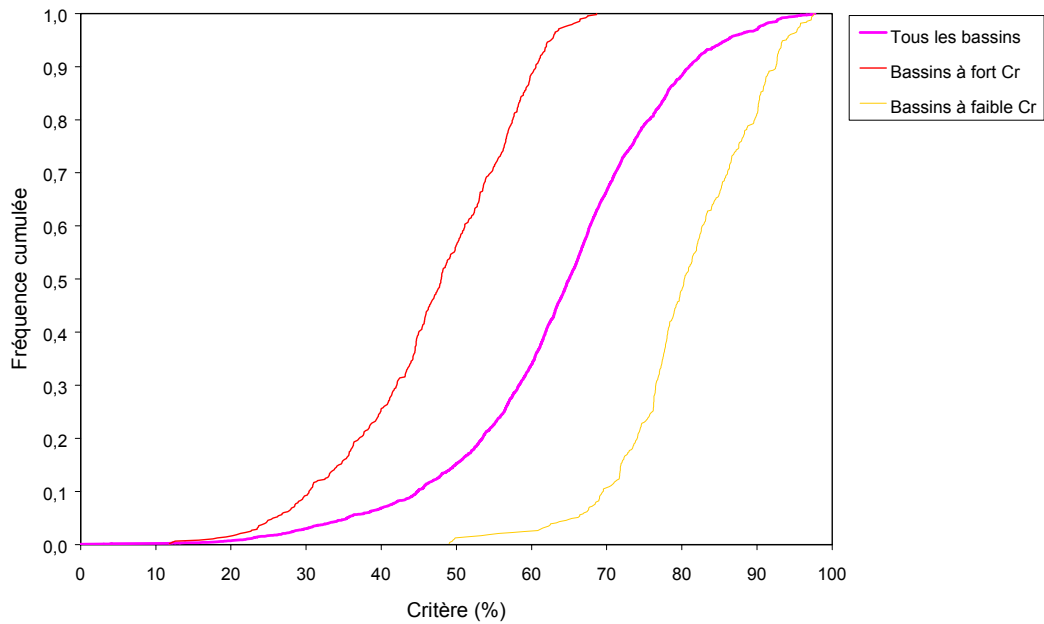


Figure 4.10: Distribution des valeurs du critère 1 ($\alpha = 1$ et $\beta = 1/4$) obtenues en contrôle, avec distinction des bassins à faible et fort rendements Cr

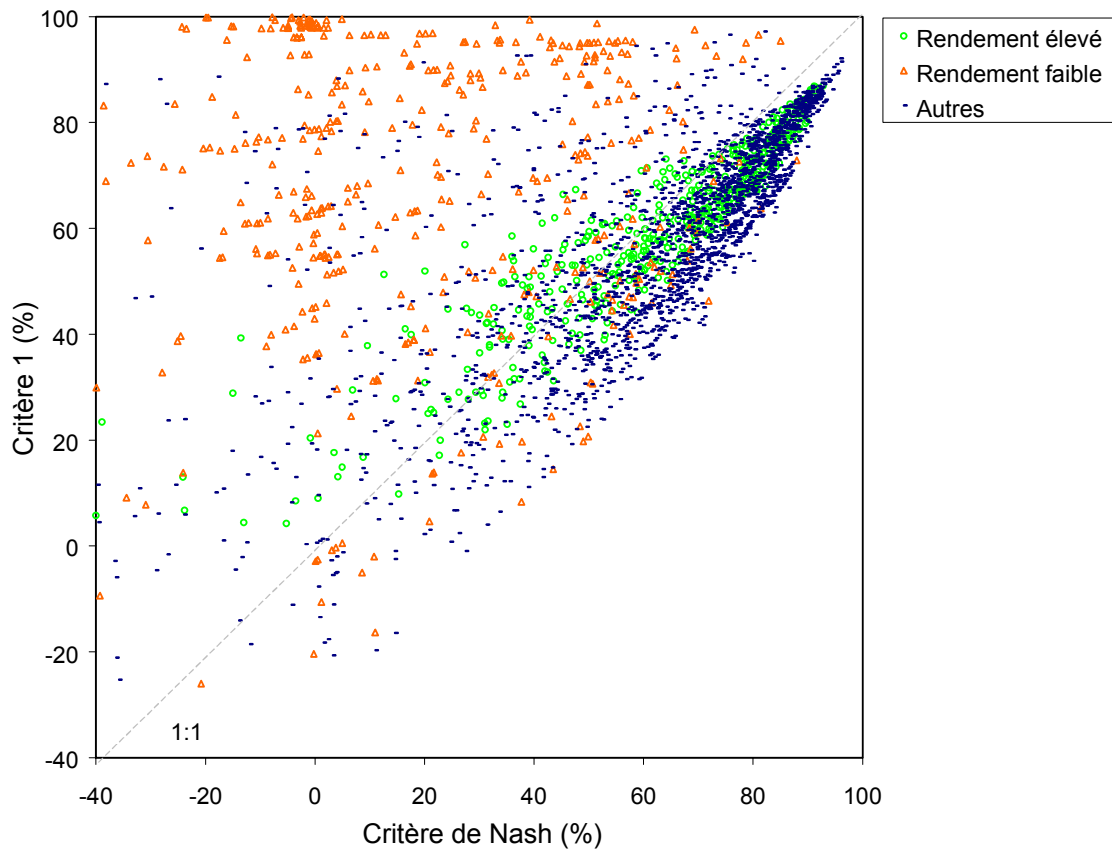


Figure 4.11: Comparaison des valeurs du critère $C1$ et du critère de Nash sur l'échantillon de bassins en contrôle ($\alpha = 1/10$ et $\beta = 2/3$)

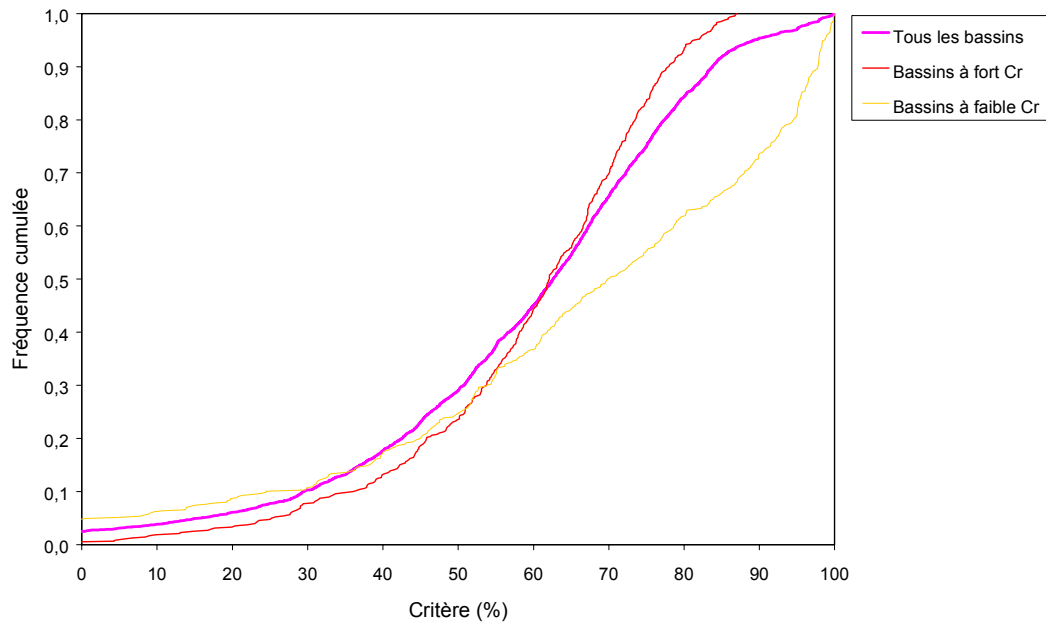


Figure 4.12: Distribution des valeurs du critère 1 ($\alpha = 1/10$ et $\beta = 2/3$) obtenues en contrôle, avec distinction des bassins à faible et fort rendements Cr

La Figure 4.9 montre par ailleurs la grande indépendance des deux critères, confirmant ainsi que le jugement porté sur le modèle par un critère dépend du modèle de référence choisi. En utilisant une valeur de α plus faible ($\alpha = 1/10$), on obtient une situation plus équilibrée, comme le montre la Figure 4.11 et la différence entre bassins à fort et faible rendement est minimisée (voir Figure 4.12). Les distributions des valeurs du critère sur les classes de bassins (Figure 4.13) sont alors relativement voisines, ceci permettant une comparaison plus aisée des performances du modèle d'un bassin à l'autre.

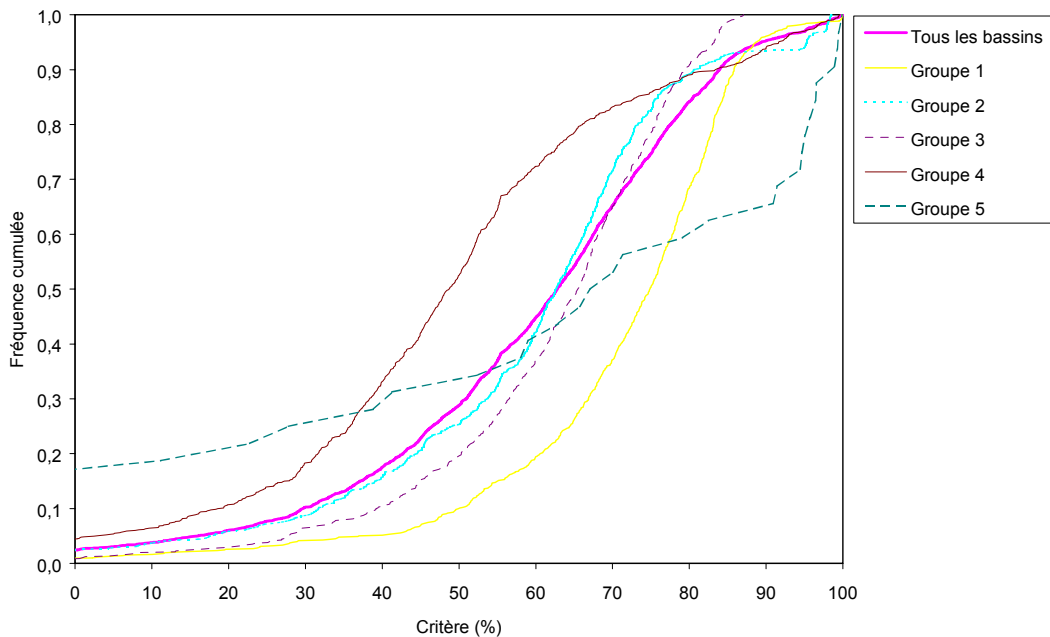


Figure 4.13: Distributions des valeurs du critère 1 obtenues sur les cinq classes de bassins en contrôle ($\alpha = 1/10$ et $\beta = 2/3$)

En utilisant un modèle de référence issu de la combinaison de l'opérateur de routage 'instantané' et d'une fraction de l'opérateur de rendement 'identité', on obtient donc une formulation qui présente des qualités intéressantes. Elle suppose néanmoins de choisir une faible valeur de α .

Critère 2 (C2)

Le modèle de référence du critère 2 fait correspondre à la pluie tombée un débit nul, représentant ainsi un bassin hypothétique très perméable où toute la pluie serait évaporée ou s'infiltrerait vers des nappes profondes pour réapparaître plus en aval. Ce critère est donné par l'expression:

$$C2 = 1 - \left(\frac{\sum_{i=1}^n (Q_{obs,i} - Q_{calc,i})^2}{\sum_{i=1}^n Q_{obs,i}^2} \right)^\beta \quad \text{Eq. (4.11)}$$

Une puissance $\beta = 3/4$ donne des ordre de grandeurs comparables au critère de Nash. S'il permet une diminution du nombre de valeurs négatives, ce critère conserve néanmoins les inconvénients du critère de Nash pour les bassins à faible rendement. Les distributions des valeurs du critère sont très similaires à celles obtenues aux Figures 4.4 et 4.5 pour le critère de Nash. La formulation C2 ne sera donc pas retenue dans la suite.

Critère 3 (C3)

Le critère C3 correspond à une combinaison linéaire des deux modèles de référence précédents. Il est donné par :

$$C3 = 1 - \left(\frac{\sum_{i=1}^n (Q_{obs,i} - Q_{calc,i})^2}{\delta \cdot \sum_{i=1}^n (Q_{obs,i} - \alpha \cdot P_{obs,i})^2 + (1 - \delta) \cdot \sum_{i=1}^n Q_{obs,i}^2} \right)^\beta \quad \text{Eq. (4.12)}$$

où δ est un troisième coefficient inférieur à 1. En prenant $\alpha = 1$, on utilise ainsi les deux cas extrêmes de l'opérateur de rendement ('identité' et 'zéro'). Ce critère pourrait permettre une compensation des inconvénients de ces deux opérateurs. Diverses combinaisons des paramètres α , β et δ ont été essayées. Des résultats relativement satisfaisants ont été obtenus avec le triplet (1/10, 2/3, 1/5). Ces valeurs indiquent que l'importance accordée à la prise en compte de la pluie doit être faible. Cette solution n'apporte pas de nette amélioration par rapport à la solution du critère C1. Ce dernier, dont la formulation est plus simple, sera donc préféré à C3.

Critère 4 (C4)

Pour l'indice C1, un des problèmes majeurs était la trop grande différence existant entre le débit observé et la pluie pour les bassins à faible rendement. Ces bassins sont majoritairement des bassins situés dans des zones soumises à une ETP relativement forte. Une façon de diminuer de façon réaliste la prépondérance de la pluie, plutôt que d'utiliser une faible valeur de α , est de lui faire subir une neutralisation préalable par l'ETP. La pluie journalière est donc diminuée de l'ETP correspondante. La deuxième entrée du modèle est ainsi utilisée dans la construction du modèle de référence. Le critère C4 correspondant est donné par:

$$C4 = 1 - \left(\frac{\sum_{i=1}^n (Q_{obs,i} - Q_{calc,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \alpha \cdot \max(0., P_{obs,i} - E_{obs,i}))^2} \right)^\beta \quad \text{Eq. (4.13)}$$

où $E_{obs,i}$ est l'ETP du jour i . Les résultats obtenus avec ce critère sont très proches de ceux obtenus avec le critère 1. On ne peut éviter ici de prendre une valeur de α différente de 1. Des valeurs faibles de α ($\alpha = 1/10$) comme dans le cas du critère 1 semblent le mieux convenir. La grande similarité avec le critère 1 est en fait facilement compréhensible : la soustraction de l'ETP à la pluie ne diminue que faiblement la variance des pluies observées, en particulier pour les bassins soumis à des régimes de pluie fortes et peu fréquentes.

Critère 5 (C5)

Pour limiter les effets de la grande variance de la pluie journalière, on peut choisir d'utiliser comme opérateur de routage non pas l'opérateur 'instantané' mais l'opérateur 'infini'. Le critère 5 ainsi construit s'écrit:

$$C5 = 1 - \left(\frac{\sum_{i=1}^n (Q_{obs,i} - Q_{calc,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \alpha \cdot \bar{P})^2} \right)^\beta \quad \text{Eq. (4.14)}$$

où \bar{P} est la pluie moyenne sur la période considérée. Cette solution donne de bons résultats en référence aux objectifs que nous nous sommes fixés. L'utilisation d'une valeur de α différente de 1 est néanmoins nécessaire. Plusieurs couples de valeurs de (α, β) donnent des résultats assez proches, dont les plus satisfaisants sont obtenus avec $(1/6, 5/6)$, $(1/5, 4/5)$, $(1/4, 3/4)$ ou $(1/3, 2/3)$. Assez curieusement, pour ces couples de valeurs, α et β semblent liés par la relation $\beta = 1 - \alpha$.

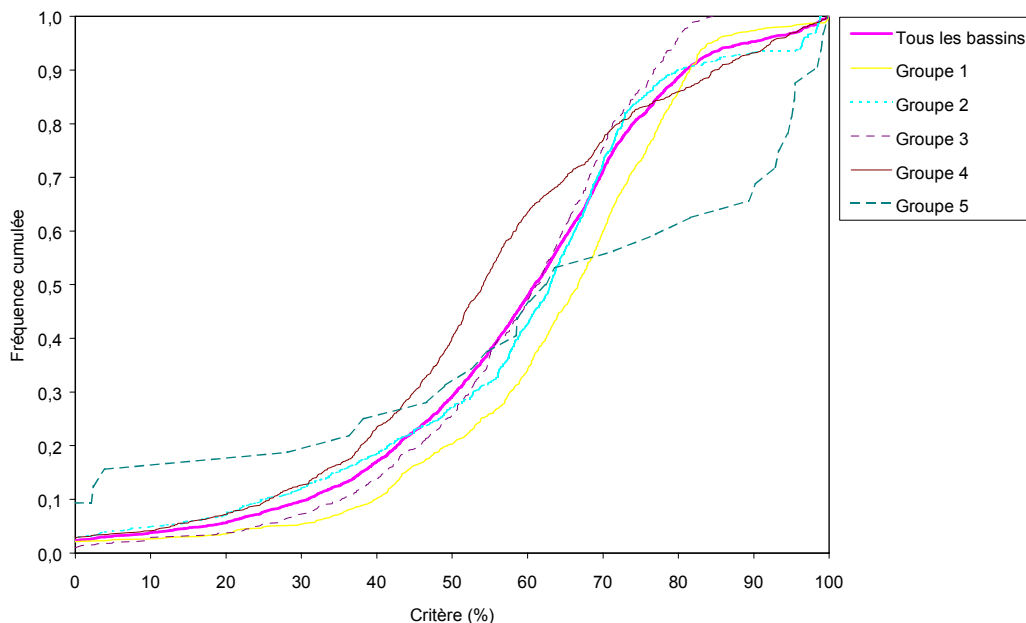


Figure 4.14: Distributions des valeurs du critère 5 obtenues sur les cinq classes de bassins en contrôle ($\alpha = 1/3$ et $\beta = 2/3$)

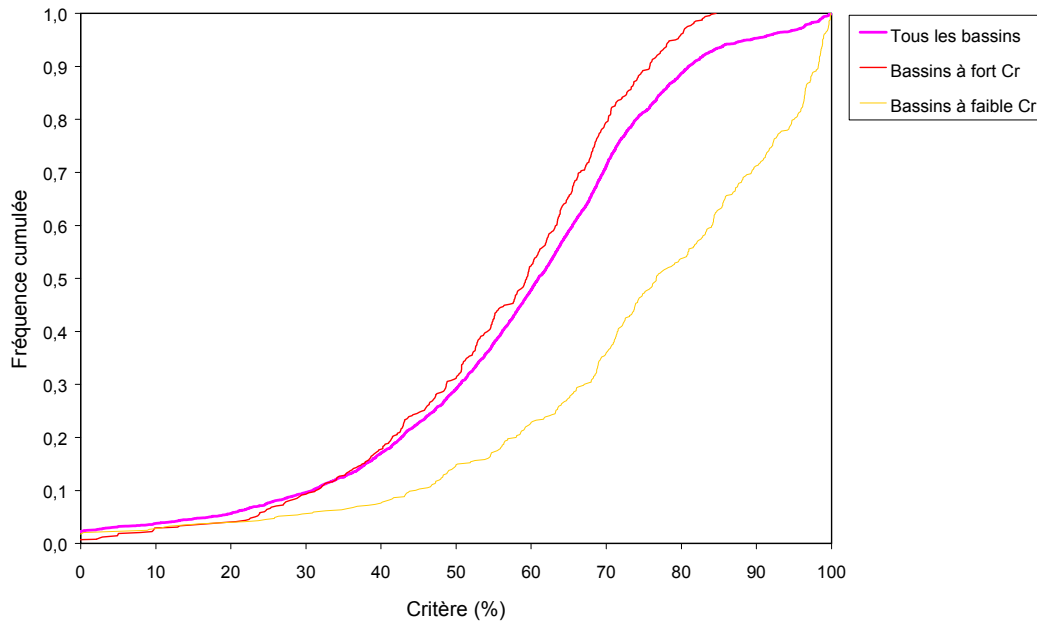


Figure 4.15: Distribution des valeurs du critère 5 ($\alpha = 1/3$ et $\beta = 2/3$) obtenues en contrôle, avec distinction des bassins à faible et fort rendements Cr.

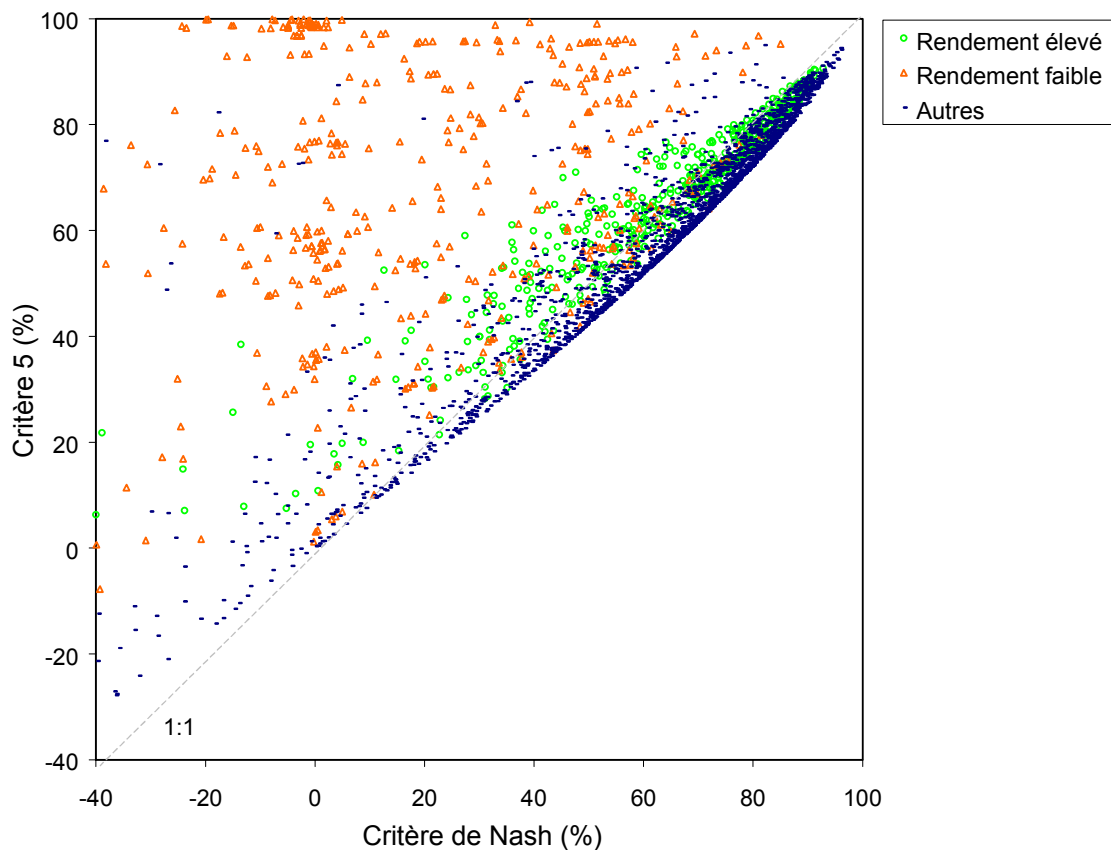


Figure 4.16: Comparaison des valeurs du critère 5 et du critère de Nash sur l'échantillon de bassins en contrôle ($\alpha = 1/5$ et $\beta = 4/5$)

Lorsque α vaut $1/3$, les distributions de résultats sur les différents groupes de bassins sont très proches les unes des autres, comme le montre la Figure 4.14. En revanche, la Figure 4.15 révèle que ce coefficient semble privilégier les bassins à faible coefficient d'écoulement.

Inversement, en prenant la valeur $1/6$, il y a moins de différence entre les distributions des critères pour les bassins à fort et faible coefficients d'écoulement, mais il y a alors moins de valeurs élevées du critère pour les bassins les plus secs (groupes 4 et 5).

Nous avons adopté une solution intermédiaire, en prenant le couple $(1/5, 4/5)$, dont les résultats sont illustrés aux Figures 4.16, 4.17 et 4.18. Pour tous les groupes de bassins, on obtient des plages de variations du critère similaires et la différence entre bassins à faible et fort rendement est atténuée. En comparaison du critère de Nash, il y a beaucoup moins de valeurs négatives, même si ces dernières ne sont pas totalement supprimées. Pour les groupes de bassins 4 et 5, il devient statistiquement plus plausible d'obtenir des valeurs élevées du critère que dans le cas du critère de Nash.

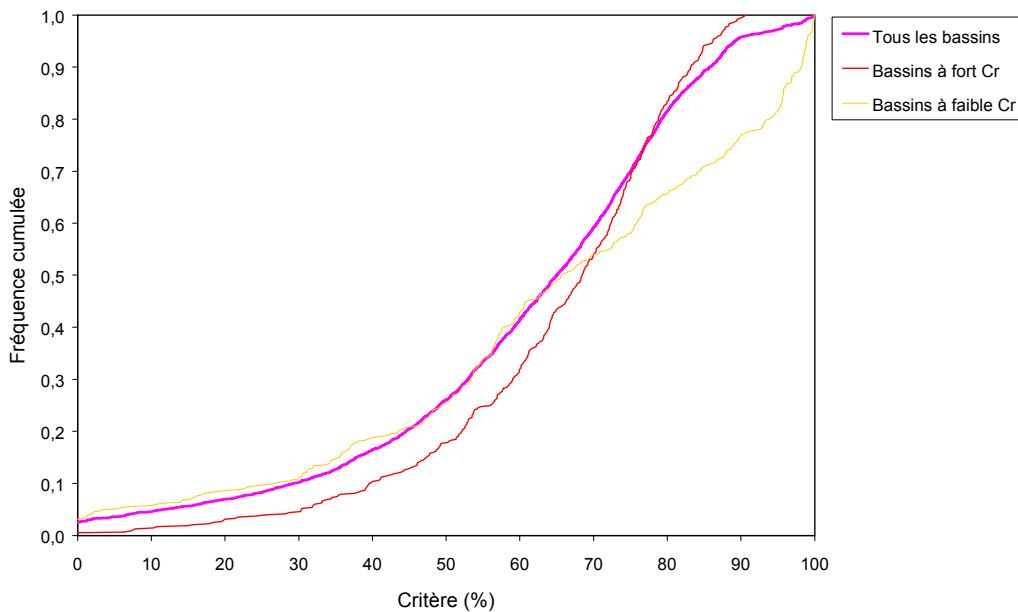


Figure 4.17: Distribution des valeurs du critère 5 ($\alpha = 1/5$ et $\beta = 4/5$) obtenues en contrôle, avec distinction des bassins à faible et fort rendements Cr.

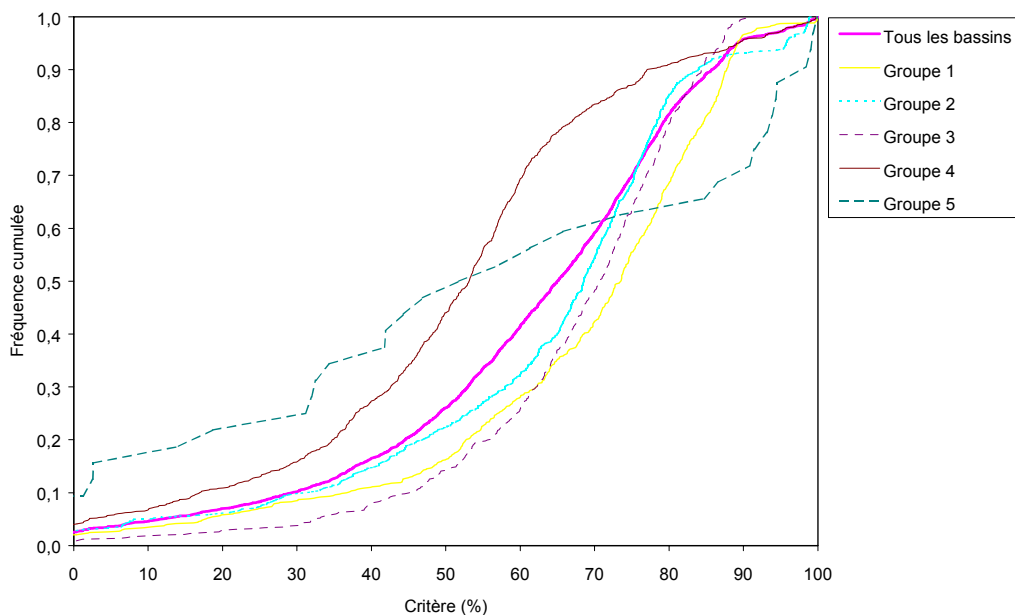


Figure 4.18: Distributions des valeurs du critère 5 obtenues sur les cinq classes de bassins en contrôle ($\alpha = 1/5$ et $\beta = 4/5$)

Critère 6 (C6)

Pour ce critère, nous avons adopté la même attitude que précédemment dans le passage du critère 1 au critère 5 (changement d'opérateur de transfert), à partir cette fois de l'opérateur de rendement $\max(0, P-ETP)$ du critère 4. Ce sixième critère proposé est de la forme :

$$C6 = 1 - \left(\frac{\sum_{i=1}^n (Q_{obs,i} - Q_{calc,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \alpha \cdot \overline{\max(0, P - E)})^2} \right)^\beta \quad \text{Eq. (4.15)}$$

où $\overline{\max(0, P - E)}$ est la moyenne des pluies journalières après neutralisation par l'ETP sur la période considérée. Les résultats trouvés ici sont proches de ceux du critère précédent (en prenant des valeurs de α légèrement plus élevées), pour la même raison que les résultats du critère 4 étaient proches de ceux du critère 1. Ceci est illustré à la Figure 4.19.

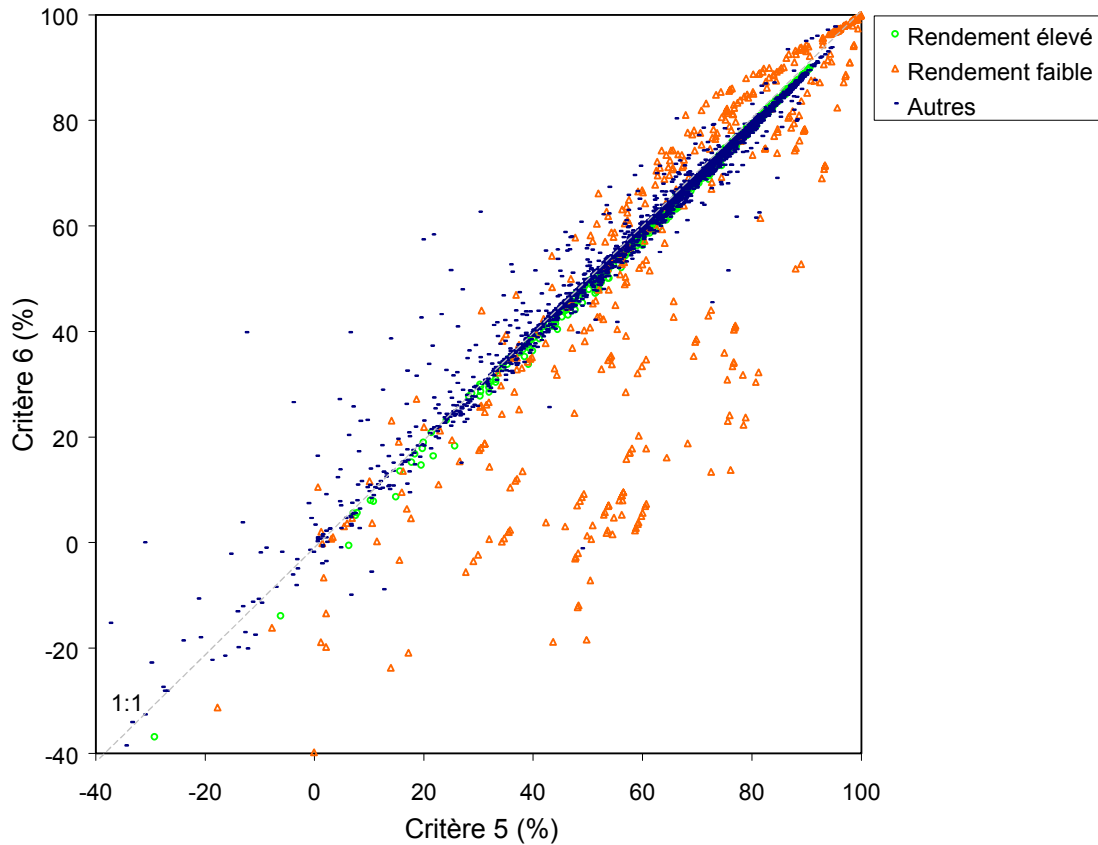


Figure 4.19: Comparaison des valeurs du critère 6 ($\alpha = 1/3$ et $\beta = 4/5$) au critère 5 sur l'échantillon de bassins en contrôle ($\alpha = 1/5$ et $\beta = 4/5$)

Critère 7 (C7)

Devant les faibles bénéfices d'utiliser évapotranspiration de manière soustractive par rapport à la pluie, nous avons essayé d'utiliser un dernier opérateur de rendement, fonction rationnelle de la pluie et de l'ETP tirée des travaux de Mouelhi (2000). Le critère C7 ainsi construit est donné par:

$$C7 = 1 - \left(\frac{\sum_{i=1}^n (Q_{obs,i} - Q_{calc,i})^2}{\sum_{i=1}^n \left(Q_{obs,i} - \alpha \cdot (\sqrt{P^2 + E^2} - E) \right)^2} \right)^\beta \quad \text{Eq. (4.16)}$$

En prenant $\alpha = 1$ et $\beta = 1$, on trouve des résultats relativement satisfaisants, comme le montrent les Figures 4.20 et 4.21: il y a moins de valeurs négatives du critère; on retrouve la similarité des distributions des résultats suivant les cinq classes de bassins observées pour certaines des formulations précédentes des critères et une répartition plus équilibrée des valeurs du critère pour les bassins à fort et à faible coefficient de rendement. Par rapport aux formulations précédentes, l'avantage de celle-ci est de ne pas trop favoriser les bassins à très faible rendement comme ceux de la classe 5.

Une valeur de 0,8 pour β permet de retrouver des niveaux de performance similaires à ceux obtenus avec le critère de Nash. Cependant, par souci de simplicité et pour conserver une correspondance de formulation avec le critère de Nash, nous avons préféré conserver un exposant égal à 1. Les valeurs du critère obtenues seront donc légèrement supérieures à celles du critère de Nash en moyenne.

La formulation du critère 7 présente donc les avantages de prendre en compte conjointement les deux entrées du modèle, d'être simple, d'utiliser une formulation bien établie du modèle interannuel (celle de Tixeront, 1964) et de répondre de façon satisfaisante aux objectifs fixés. Ce sera finalement cette formulation que nous retiendrons comme alternative au critère de Nash.

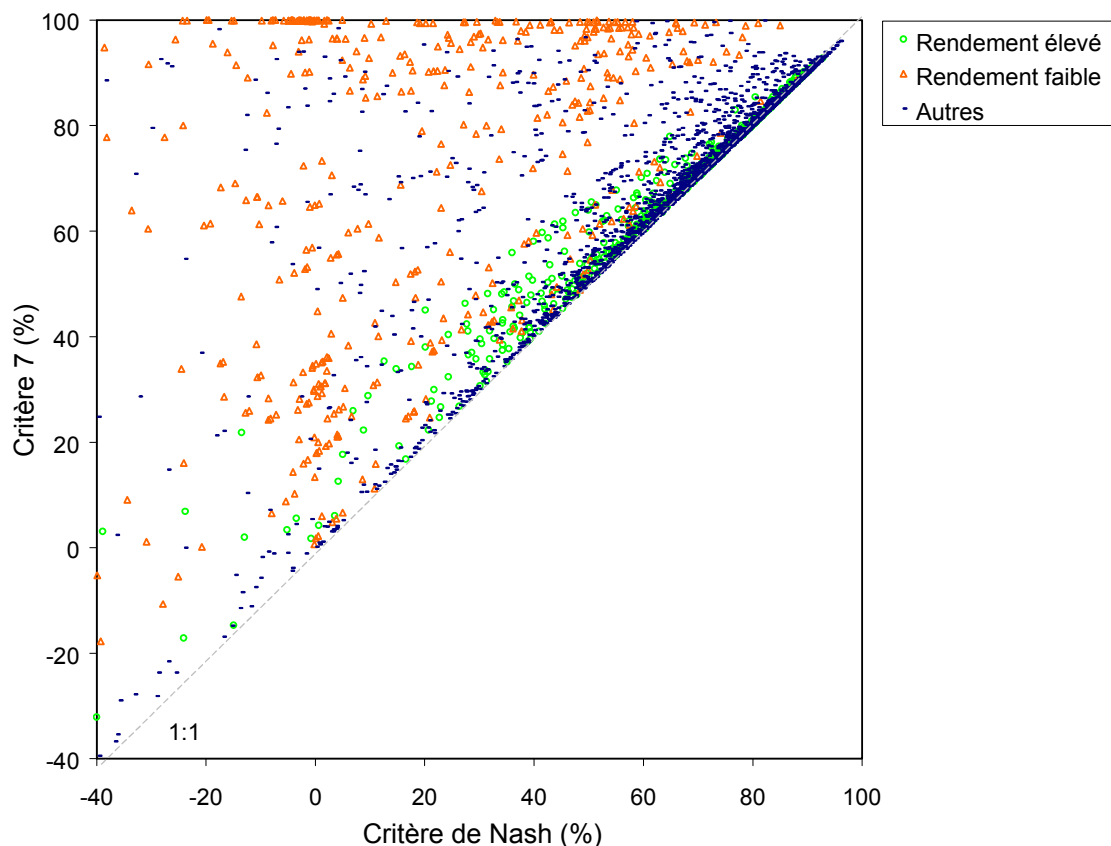


Figure 4.20: Comparaison des valeurs du critère $C7$ et du critère de Nash sur l'échantillon de bassins en contrôle ($\alpha = 1/2$ et $\beta = 1$)

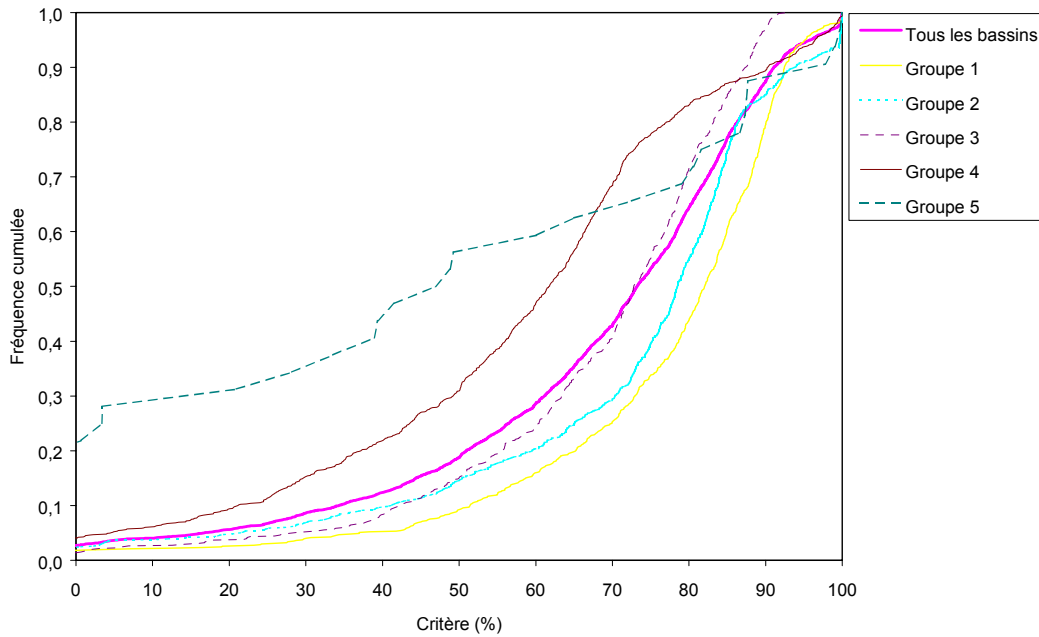


Figure 4.21: Distributions des valeurs du critère 7 obtenues sur les cinq classes de bassins en contrôle ($\alpha = 1/2$ et $\beta = 1$)

4.3.5. Quelques conclusions

Nous avons pu mettre en évidence, dans la recherche d'un nouveau critère, que l'interprétation numérique des résultats d'un modèle au travers d'une mesure de performance relative est très dépendante du modèle de référence choisi. Le critère de Nash utilise comme modèle de référence un débit constant égal au débit moyen journalier sur la période observée. Ce dernier semble être un modèle de référence trop exigeant dans le cas des bassins semi-arides ou à faible rendement. Sur ce type de bassins, le critère prend donc très facilement des valeurs négatives voire très fortement négatives qui enlèvent tout sens à une moyenne de plusieurs critères. Par ailleurs, les valeurs du critère plafonnant à des valeurs relativement basses par rapport aux autres groupes de bassins, il semble que le modèle ne puisse jamais donner satisfaction sur ce type de bassins.

En combinant des opérateurs de rendement ou de routage pour construire des modèles de référence différents mais toujours très simples, nous avons constaté qu'une grande indépendance avec le critère de Nash peut être obtenue, notamment pour les bassins intermittents ou à faible rendement. L'utilisation des entrées du modèle permet de pallier les inconvénients du critère de Nash, notamment en prenant une fraction de la pluie journalière ou de la pluie moyenne comme modèle de référence, ou encore la formulation d'un modèle interannuel qui permet d'exprimer le module des débits en fonction de la pluie et de l'ETP.

Ceci permet de mieux rendre compte du travail de transformation demandé au modèle sur les entrées. Parmi les formulations proposées, celle du critère 7 utilisant le modèle de Tixeront (1964) est sans doute l'une des plus satisfaisantes: elle prend en compte explicitement les deux entrées du modèle, elle permet de diminuer substantiellement la proportion de valeurs négatives du critère et elle permet d'obtenir des gammes de variation des critères similaires sur les différents types de bassins. La formulation choisie, avec l'exposant β valant 1, a cependant tendance à produire des valeurs pour ce critère légèrement plus élevées que celles communément obtenues avec le critère de Nash, une valeur de β de 0,8 permettant un réajustement de ces niveaux de valeurs.

4.4. Proposition d'un nouveau critère de bilan

Nous allons discuter ici de la forme à donner au critère de bilan, les formulations des critères de bilan B ou B' données aux Eq. (4.8) et (4.9) présentant des inconvénients que nous exposons maintenant. Rappelons tout d'abord que ce critère est très différent des critères précédents puisqu'un bilan exact peut être obtenu pour une infinité de jeux de paramètres (environ \mathbb{R}^n si n est le nombre de paramètres).

Nous considérons dans la suite la forme du critère B comme base de notre réflexion. Du fait de la possibilité d'avoir des valeurs du critère de bilan inférieures ou supérieures à 1, une moyenne sur un échantillon de bassins perd tout son intérêt car des compensations entre valeurs peuvent indiquer que le modèle fournit un bon bilan alors que ce n'est pas le cas. Pour contourner ce problème, il est intéressant de se ramener à un critère pour lequel l'adéquation parfaite entre les valeurs de bilan observée et calculée est indiquée par un maximum du critère et non par une valeur centrale. Nous avons donc cherché à nous ramener à un critère variant dans l'intervalle $]-\infty;1]$, comme c'est le cas pour le critère de Nash par exemple.

Le premier critère simple envisagé est donné par :

$$\begin{aligned} B1 &= 1 - |1 - B| \\ &= 1 - |B'| \end{aligned} \quad \text{Eq. (4.17)}$$

Ce critère a l'avantage d'être très proche de la formulation du critère B' . Cependant, comme le montre la Figure 4.23, $B1$ s'annule pour des valeurs de B de 0 et 2. Ceci signifie que $B1$ ne peut pas prendre de valeurs négatives lorsque les débits sont sous-estimés, alors que des valeurs négatives existent lorsque la surestimation des débits excède un facteur 2. La symétrie de la décroissance de $B1$ de part et d'autre de 1 a donc pour effet de pénaliser davantage les surestimations du débit au delà d'un facteur 2 que la simulation d'un débit nul. Pour remédier à ce déséquilibre, nous avons donc cherché une expression permettant d'avoir des intervalles de variation identiques suivant que l'on surestime ou sous-estime les bilans.

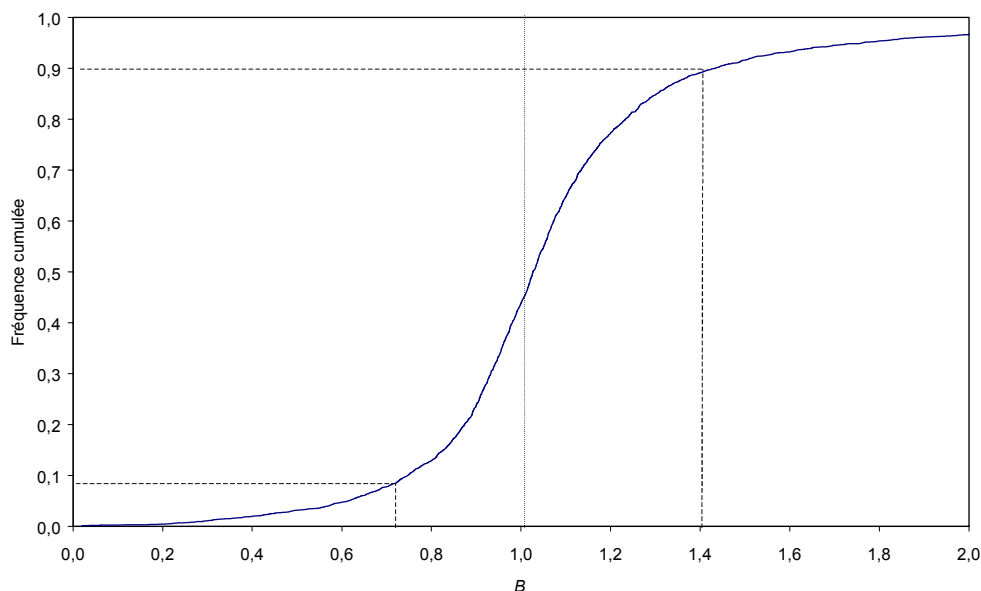


Figure 4.22: Distribution des valeurs de B au contrôle sur l'échantillon de bassins versants pour le modèle GR4K.

Par ailleurs, la distribution des valeurs de B au contrôle sur l'échantillon de bassins (voir Figure 4.22) montre que l'on a à peu près une même proportion de cas où la surestimation des bilans par le modèle excède un facteur a que de cas où la sous-estimation des bilans est en deçà d'un rapport $1/a$ (a étant un nombre supérieur à un). Ainsi par exemple, il y a environ 10 % des contrôles pour lesquels B est plus grand que 1,4 et également 10 % des contrôles pour lesquels B est plus petit que $1/1,4 = 0,71$. Il y a donc statistiquement à peu près autant de chance d'être au dessus de a qu'en dessous de $1/a$ sur notre échantillon de bassins.

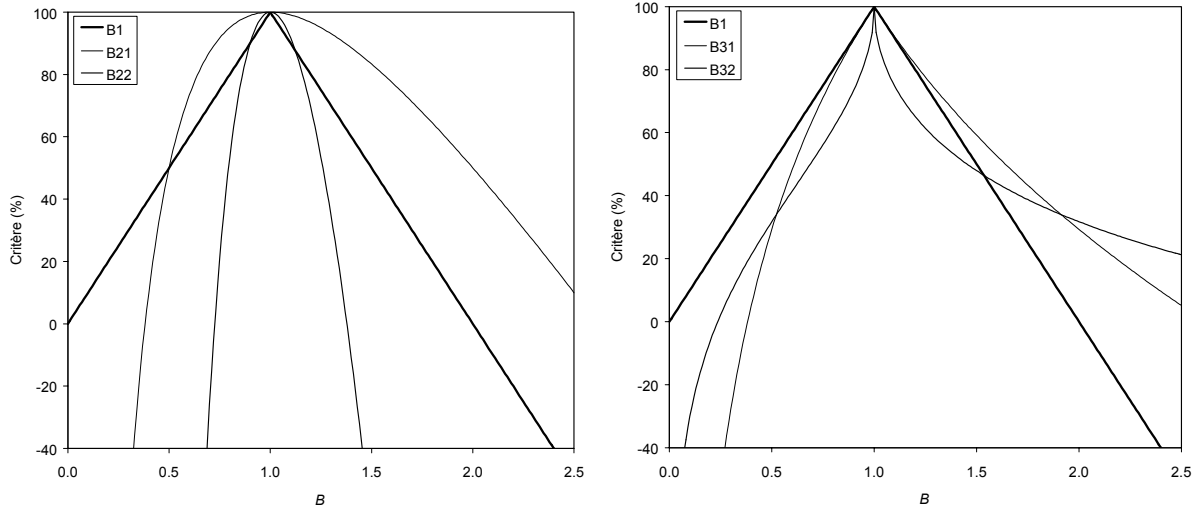


Figure 4.23: Courbes des critères $B1$, $B2$ et $B3$ en fonction de B ($B21 = B2(1,1)$, $B22 = B2(1,3)$, $B31 = B3(1,1/2,1)$, $B32 = B3(1,1/3,1/2)$, avec entre parenthèses les paramètres des Eq. (4.18) et Eq. (4.19))

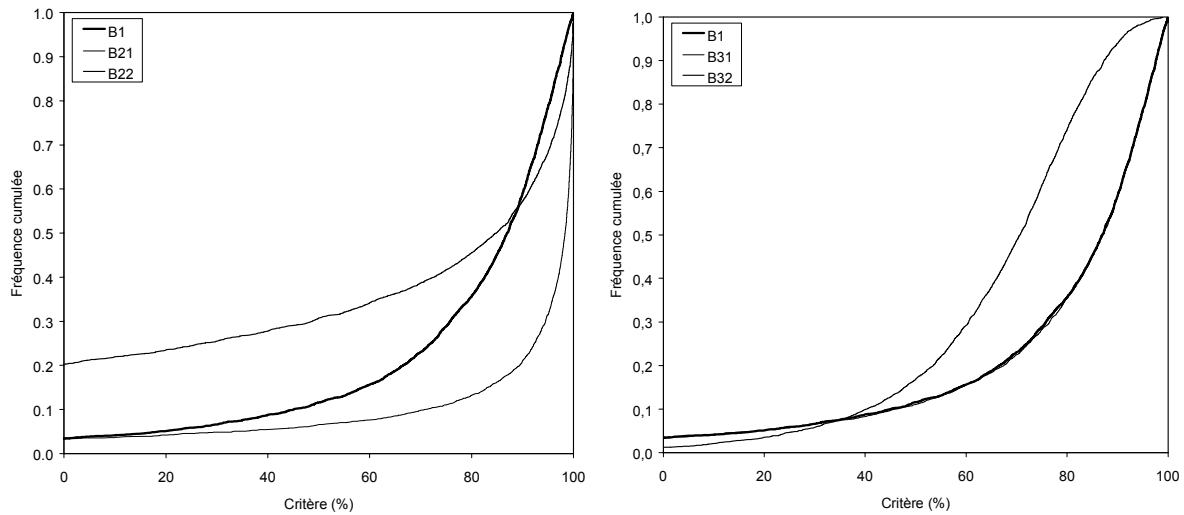


Figure 4.24: Distributions des critères $B1$, $B2$ et $B3$ obtenus en contrôle sur l'échantillon test ($B21 = B2(1,1)$, $B22 = B2(1,3)$, $B31 = B3(1,1/2,1)$, $B32 = B3(1,1/3,1/2)$, avec entre parenthèses les paramètres des Eq. (4.18) et Eq. (4.19))

Nous avons donc souhaité garder cette symétrie multiplicative dans la formulation du critère de bilan, en faisant en sorte d'avoir une même valeur du critère suivant que le bilan est surestimé d'un facteur a ou sous-estimé d'un facteur $1/a$.

Pour satisfaire ces objectifs, nous proposons les deux expressions suivantes:

$$B2 = (2 \cdot x_1 + 1) - x_1 \cdot \left(B^{x_2} + \frac{1}{B^{x_2}} \right) \quad \text{Eq. (4.18)}$$

$$B3 = 1 - x_1 \cdot \left| B^{x_2} - \frac{1}{B^{x_2}} \right|^{x_3} \quad \text{Eq. (4.19)}$$

où x_1 , x_2 et x_3 sont des nombres positifs. Les expressions de $B1$ et $B2$ répondent bien aux objectifs fixés précédemment: $B2$ et $B3$ ont une valeur de 100 % lorsque B vaut 1 et prennent des valeurs inférieures pour des valeurs de B différentes de 1; les intervalles de variations sont identiques pour des valeurs de B supérieures ou inférieures à 1; les conditions $B2(a)=B2(1/a)$ et $B3(a)=B3(1/a)$ sont vérifiées.

La Figure 4.23 illustre les variations de $B2$ et $B3$ en fonction de B . Nous avons tracé quelques formulations simples de $B2$ et $B3$: deux courbes de $B2$ ($B21$ et $B22$) avec (x_1, x_2) égal à (1,1) et (1,3) respectivement, ainsi que deux courbes de $B3$ ($B31$ et $B32$) avec pour triplet (x_1, x_2, x_3) respectivement (1,1/2,1) et (1,1/3,1/2). Les distributions correspondantes pour les résultats du modèle au contrôle sont illustrées à la Figure 4.24.

Le critère $B2$ présente une tangente horizontale en son maximum en 1. Ceci induit un nombre élevé de fortes valeurs du critère, dans le cas de $B21$ notamment. En accroissant les valeurs de x_1 et/ou x_2 , on peut augmenter la courbure en 1 et diminuer ainsi le nombre de valeurs élevées du critère. Cependant, il y a alors parallèlement une diminution rapide du critère lorsque l'on s'écarte de 1, ce qui se traduit par un nombre plus élevé de valeurs négatives du critère dans la distribution des résultats de la Figure 4.24. Par conséquent, il semble difficile de retenir la formulation du critère $B2$. La formulation du critère $B3$ permet de dépasser les limitations du critère $B2$: ainsi, comme le montre la Figure 4.23, on peut réussir avec $B31$ à avoir des valeurs équivalentes à celles de $B1$ lorsque l'on est proche de $B = 1$, sans avoir le problème de la trop rapide décroissance du critère lorsque l'on s'écarte de 1. Dans ce cas, la distribution des résultats est quasiment similaire à celle de $B1$. En prenant x_3 différent de 1, on peut obtenir (cas de $B32$) une plus rapide décroissance autour de $B = 1$, ce qui rend les fortes valeurs du critère plus difficile à atteindre.

Les critères $B31$ et $B32$ semblent tous deux intéressants à adopter. Afin de conserver les mêmes ordres de grandeur du bilan qu'avec les critères classiques lorsque l'on est proche de $B = 1$, nous préférons conserver pour la suite de notre étude le critère $B31$. Cette nouvelle formulation du critère de bilan répond bien à nos attentes. Remarquons qu'avec ce critère, on perd néanmoins l'information de sur- ou sous-estimation du bilan qui, si elle est importante pour une étude particulière, devra être enregistrée lors du calcul intermédiaire de B .

4.5. Optimisation et évaluation multi-critère

Nous avons exposé dans ce qui précède différents critères qui peuvent être utiles pour juger des qualités des simulations des modèles pluie-débit. Dans un contexte opérationnel, il est intéressant d'utiliser plusieurs critères complémentaires. En effet, les problèmes hydrologiques qui se posent aujourd'hui aux gestionnaires sont de plus en plus complexes, car ils doivent souvent prendre en compte des usages multiples de la ressource dans une gestion 'intégrée' du bassin. Les modèles d'aide à la gestion doivent par conséquent de plus en plus être capables d'apporter des réponses pertinentes sur des aspects différents. Ainsi, on pourra exiger du modèle qu'il puisse estimer avec la même qualité les crues et les étiages, qu'il puisse représenter, en plus des débits, des flux de polluants ou de matières en suspension ou encore qu'il permette de simuler les pertes par évaporation et les niveaux de la nappe. Le modèle devient alors un outils multi-objectif. Dans ces exemples, deux cas différents

apparaissent, celui où l'on s'intéresse à plusieurs variables (cas d'un modèle à plusieurs sorties) et celui où l'on s'intéresse à plusieurs qualités d'une unique variable de sortie.

Le premier cas suppose que l'on ait un modèle à plusieurs sorties. On s'intéresse alors à mettre en évidence le réalisme du modèle vis-à-vis de ces différentes variables. Cette validation multi-critères est souvent difficile à mener. De relatifs succès ont été obtenus par exemple par Ambroise *et al.* (1995) sur le bassin versant de la Fecht dans les Vosges, avec une version du Modèle Couplé conceptuel semi-distribué de Girard *et al.* (1981), qui permet de prendre en compte notamment les stocks hydriques ou les couverts neigeux. Ma *et al.* (1990) ont également pu, à partir d'un modèle hydrologique pluie-débit, modéliser des flux de nitrates à l'échelle du bassin versant. En revanche, des limitations de la formulation conceptuelle de certains modèles ont été mises en évidence par Lamb *et al.* (1998) et Güntner *et al.* (1999) dans le cas de TOPMODEL ou par Mroczkowski *et al.* (1997) dans le cas du modèle CATPRO: ils ont montré la difficulté de relier des variables internes de ces modèles à des variables observées sur le terrain. Le problème d'avoir un modèle à plusieurs sorties apparaît ainsi beaucoup plus compliqué que d'avoir un modèle à une seule sortie, la formulation de ce dernier étant rarement adaptée a priori pour simuler plusieurs variables.

Dans le cas d'une unique variable, ici le débit, la démarche est différente. Le choix d'une fonction objectif pour l'optimisation des paramètres du modèle est essentiellement dépendante de l'utilisateur et de l'objectif assigné au modèle. La procédure de calage classique fournit un jeu de paramètres satisfaisants relativement à cet objectif particulier. Cependant, devant la mise en évidence du caractère non-unique des jeux de paramètres satisfaisants, certains auteurs, notamment Beven (1987), soulignent la nécessité de mettre en place une nouvelle approche, un nouveau paradigme permettant de tenir compte de manière probabiliste du caractère incertain de la détermination des paramètres de modèles hydrologiques. Ainsi ont été proposées des démarches probabilistes permettant d'identifier non plus un jeu de paramètres satisfaisants mais un ensemble de jeux de paramètres également satisfaisants, avec des notions d'incertitudes associées (Van Straten et Keesman, 1991; Beven et Binley, 1992; Kuczera et Parent, 1998). Dix ans après Beven (1987), Gupta *et al.* (1998) plaident de nouveau pour un changement de paradigme, indiquant qu'il peut apparaître plus satisfaisant d'adopter une démarche multi-critères pour aborder le calage des modèles et déterminer les jeux de paramètres acceptables.

En effet, il peut être intéressant pour l'utilisateur d'avoir un modèle robuste sur plusieurs objectifs complémentaires, par exemple pour une bonne simulation des crues et des étiages. De ce fait, certains modélisateurs ont adopté pour le calage de leurs modèles des approches essentiellement manuelles, dans lesquelles on est amené à choisir les jeux de paramètres par compromis entre plusieurs statistiques d'évaluation de la qualité. C'est le cas par exemple dans la méthodologie de modélisation IHACRES (Jakeman *et al.*, 1990; Littlewood *et al.*, 1997) ou dans la démarche de calage du modèle de Sacramento (Brazil et Hudlow, 1980).

La démarche multi-critères en optimisation est nettement plus complexe qu'une démarche classique avec un seul critère, puisqu'elle implique, comme nous venons de le mentionner, une situation de compromis: on accepte une diminution de la qualité d'un critère pour avoir en compensation une meilleure qualité sur un autre. La détermination d'un jeu de paramètres suppose que l'on puisse déterminer les importances relatives des différents objectifs. Les travaux de Yan et Haan (1991a, 1991b) soulignent la difficulté de parvenir à des pondérations satisfaisantes. Les résultats qu'ils présentent montrent que ces pondérations peuvent être dépendantes du bassin considéré. Seibert (1997), lui, propose des combinaisons de critères fondées sur des règles floues et non sur des pondérations arithmétiques. Du fait de la difficulté de choisir des règles combinatoires satisfaisantes, l'automatisation d'une procédure de calage

multi-critères n'est pas aisée. Quelques auteurs ont néanmoins proposé des méthodes automatiques de calage multi-objectifs, comme par exemple Yapo *et al.* (1998).

Dans notre étude, nous nous sommes contentés d'adopter une démarche d'optimisation classique avec une fonction objectif reposant sur seul critère de qualité, le critère de Nash calculé sur les racines carrées des débits. Ce critère semble suffisamment central pour constituer une situation de compromis entre simulation des crues et des étiages

4.6. Conclusion

Nous avons présenté ici une réflexion sur les critères d'appréciation des performances des modèles que nous allons tester. Pour des évaluations au cas par cas dans des études ponctuelles, des critères graphiques peuvent être utilisés. Ils permettent de confronter, par des illustrations, les simulations à la réalité observée et de détecter certaines anomalies mal décelables par les critères numériques classiques. Cependant, de tels critères graphiques restent subjectifs et sont donc peu appropriés à des études comparatives. Par ailleurs, étant donné la taille de notre échantillon de bassins, il n'était pas possible d'envisager de procédures d'inspection graphique systématique. Ce sont donc des critères numériques qui ont été préférés.

Les critères globaux d'appréciation des performances des modèles de simulation des débits diffèrent sur trois points essentiels: la forme analytique de l'erreur (quadratique, absolue, cumulée...), la forme de la variable cible (transformée ou non) et le choix d'un modèle de référence (pour avoir une mesure relative). Nous avons choisi de retenir dans la suite de notre étude les six critères suivants:

- **CR1**: critère de Nash calculé sur les débits

$$CR1 = 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{calc,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2} \quad \text{Eq. (4.20)}$$

où $Q_{obs,i}$ et $Q_{calc,i}$ sont respectivement les débits observé et simulé au pas de temps i , n le nombre total de pas de temps de la période d'étude. La barre de surlignement indique la moyenne sur les n pas de temps de la période considérée.

- **CR2**: critère de Nash calculé sur les racines carrées des débits

$$CR2 = 1 - \frac{\sum_{i=1}^n (\sqrt{Q_{obs,i}} - \sqrt{Q_{calc,i}})^2}{\sum_{i=1}^n (\sqrt{Q_{obs,i}} - \sqrt{\overline{Q_{obs}}})^2} \quad \text{Eq. (4.21)}$$

- **CR3**: critère de Nash calculé sur les logarithmes des débits

$$CR3 = 1 - \frac{\sum_{i=1}^n (\ln(Q_{obs,i} + \varepsilon) - \ln(Q_{calc,i} + \varepsilon))^2}{\sum_{i=1}^n (\ln(Q_{obs,i} + \varepsilon) - \ln(\overline{Q_{obs}} + \varepsilon))^2} \quad \text{Eq. (4.22)}$$

où ε est une valeur faible (par exemple Module/40).

- **CR4**: critère 'Nash-bis'

$$CR4 = 1 - \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{calc,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \hat{M})^2} \quad \text{avec } \hat{M} = \sqrt{\overline{P^2 + E^2} - \overline{E}} \quad \text{Eq. (4.23)}$$

où P et E sont la pluie et l'ETP respectivement

- **CR5**: critère relatif d'erreur absolue

$$CR5 = 1 - \frac{\sum_{i=1}^n |Q_{obs,i} - Q_{calc,i}|}{\sum_{i=1}^n |Q_{obs,i} - \overline{Q_{obs}}|} \quad \text{Eq. (4.24)}$$

- **CR6**: critère de bilan

$$CR6 = 1 - \left| \frac{\sum_{i=1}^n Q_{calc,i}}{\sum_{i=1}^n Q_{obs,i}} - \frac{\sum_{i=1}^n Q_{obs,i}}{\sum_{i=1}^n Q_{calc,i}} \right| \quad \text{Eq. (4.25)}$$

Les quatre premiers critères utilisent l'erreur quadratique du modèle. Comme nous l'avons montré, le critère $CR4$ pallie certains inconvénients du critère de Nash ($CR1$) et nous l'utilisons donc à ce titre comme complément pour l'analyse des résultats de la comparaison. Les transformations des critères $CR2$ et $CR3$ apportent un éclairage complémentaire en accordant plus de poids à différentes classes de débit. Le critère $CR2$, n'accordant pas une trop forte importance aux événements de crues en comparaison du critère $CR1$, a été préféré à ce dernier comme fonction objectif lors de la phase d'optimisation des paramètres. Les critères $CR5$ et $CR6$ représentent respectivement des mesures sur l'erreur absolue et sur l'erreur de volume global écoulé.

Ces critères globaux prennent tous en compte l'ensemble des débits simulés sur la période d'étude. Ils varient dans l'intervalle $]-\infty; 1]$ et permettent de réaliser des moyennes des performances sur plusieurs périodes ou plusieurs bassins. La diversité de ce jeu de critères couvre relativement bien les différentes qualités qui peuvent être souhaitées pour un modèle de simulation des débits. Il s'agit en fait d'une grille de lecture des performances des modèles qui permettra d'interpréter les résultats de la comparaison au regard de différents objectifs complémentaires de simulation des débits. En ce sens, ils peuvent refléter les préoccupations d'utilisateurs des modèles intéressés dans la gestion de la ressource en eau, la simulation des étiages ou celle des crues par exemple.

Cependant, ces critères ont aussi leurs limites, car ils ne mettent pas en évidence tous les types d'erreurs, comme l'expliquent Garrick *et al.* (1978), en particulier dans le cas d'erreurs systématiques (Aitken, 1973). Des critères appliqués sur certaines parties des hydrogrammes sont également pertinents pour l'analyse des qualités des modèles. Enfin, si nous nous sommes restreints à une évaluation au pas de temps journalier, des évaluations aux pas de temps mensuel ou annuel peuvent également être envisagées, notamment vis-à-vis de questions de ressources en eau en évaluant la capacité des modèles à bien gérer les bilans en eau (voir par exemple Weeks et Hebbert, 1980; Chiew *et al.*, 1993).

Chapitre 5

Chapitre 5

Présentation et analyse des résultats de la comparaison de structures

5.1. Introduction

Après avoir exposé au chapitre 1 les modalités de comparaison des différentes structures de modèles, nous présentons dans ce chapitre l'intégralité des résultats des tests comparatifs réalisés sur les 38 structures de modèles. Nous avons adopté pour cela une présentation semblable à celle de l'article de présentation des résultats, qui est retranscrit en Annexe 3.

Nous procédons ici à une présentation et une analyse thématique des résultats. Performances, robustesse, fiabilité des modèles sont abordées, ainsi que l'effet de la complexité (nombres de paramètres optimisés) sur ces différents aspects. Nous essayons également de mettre en évidence une éventuelle complémentarité entre structures. Pour cela, nous adoptons deux approches différentes, l'une fondée exclusivement sur les performances des modèles, l'autre s'appuyant sur la recherche d'une typologie entre modèles et types de bassins. Nous nous appuyerons dans la recherche de cette typologie bassins-modèles sur la classification des bassins réalisée au chapitre 2.

5.2. Résultats et discussions

Les 38 structures de modèles présentées dans le premier chapitre, ainsi que la structure de référence (modèle de Tsykin), ont toutes été successivement calées sur chacun des 429 bassins de l'échantillon à l'aide de la méthode de calage 'pas-à-pas' détaillée au chapitre 3. Ainsi, les 1284 calages et les 3204 tests en contrôle ont été réalisés pour chacune d'elles. Dans la présentation des résultats, nous porterons notre attention essentiellement sur les performances des modèles au contrôle, celles au calage étant moins révélatrices des réelles capacités de simulation des modèles.

Dans les paragraphes suivants, le terme 'modèle' sera parfois abusivement employé seul pour désigner la structure du modèle. Nous rappelons que nous donnons des résultats de structures de modèles, la plupart du temps versions modifiées des structures originales, comme cela a été expliqué au chapitre 1.

Nous avons choisi de présenter les résultats sous forme thématique, en distinguant plusieurs questions auxquelles nous voulions, au travers de la comparaison, essayer d'apporter des éléments de réponse. Les résultats sont analysés en fonction des critères d'évaluation *CR1* à *CR6*, dont nous avons discuté les qualités au chapitre précédent.

Etant donné le grand nombre de tests au calage et au contrôle, il était intéressant de pouvoir analyser statistiquement les résultats. Nous avons ainsi cherché à synthétiser les performances pour permettre une comparaison plus aisée entre structures sur notre échantillon de bassins. Les six critères de qualité présentent tous l'avantage de permettre le calcul d'une moyenne significative des performances (c'était d'ailleurs le but de la recherche de la nouvelle formulation pour le critère de bilan $CR6$). C'est donc la moyenne arithmétique des performances, statistique basique, que nous avons choisie en priorité pour analyser nos résultats. Comme nous l'avons vu dans le chapitre précédent, le problème des valeurs fortement négatives a une influence importante sur les moyennes. Nous avons donc voulu corroborer la significativité de la moyenne par l'utilisation d'un quantile de la distribution des résultats. Ces distributions sont obtenues en classant les performances du modèle pour chaque contrôle par ordre croissant et en construisant la distribution expérimentale correspondante. La Figure 5.1 montre un exemple de distribution, avec en ordonnée la fréquence cumulée au non dépassement: ainsi la lecture d'un quantile correspondant à une proportion des tests permet de déterminer la valeur du critère d'évaluation qui majore cette proportion des tests et minore la partie complémentaire. C'est le quantile 0,3 (noté $Cr(0,3)$ sur la figure) qui a été choisi pour compléter l'analyse par la moyenne.

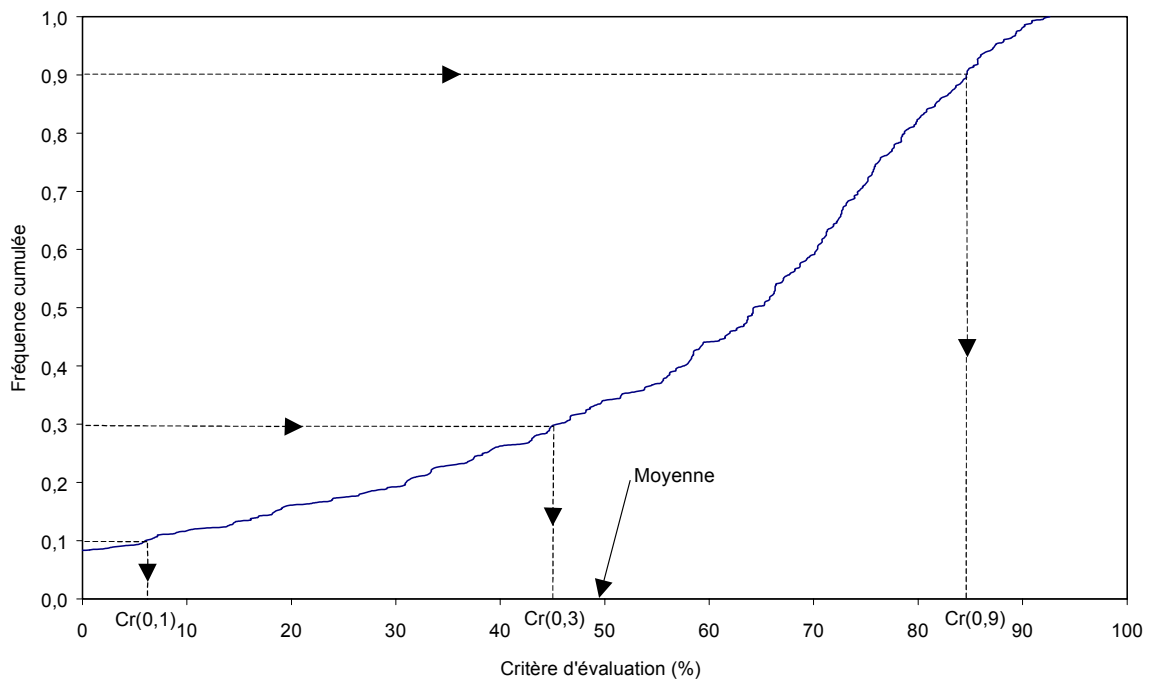


Figure 5.1: Exemple de distribution expérimentale des résultats d'un modèle, avec détermination des quantiles 0,1, 0,3 et 0,9 (dans cet exemple, la moyenne des valeurs des critères, désignée par une flèche, est de 49,3 %)

Pour étudier les valeurs les plus fortes ou les plus faibles prises par le critère, nous aurions pu nous intéresser aux valeurs maximales et minimales. Ces dernières ont paru peu informatives, surtout dans le cas des valeurs les plus faibles, où de larges écarts sont constatés entre les minima. Nous avons donc préféré utiliser les quantiles 0,1 et 0,9, illustrés à la Figure 5.1. L'écart entre ces deux quantiles représente un intervalle de variation contenant 80 % des valeurs prises par le critère.

Dans la suite, nous utilisons la moyenne des résultats et les quantiles des distributions pour analyser les performances des modèles.

5.2.1. Quel niveau de performance peut-on attendre des modèles testés ?

A côté des modèles hydrologiques intégrant une prise en compte de l'humidité du bassin développés depuis le début des années 60, de nombreux modèles n'intégrant pas de telles procédures de suivi de l'humidité ont été proposés, au rang desquels nous pouvons citer notamment les modèles ARMAX, outils maintenant classiques en hydrologie, ou les réseaux de neurones artificiels appliqués récemment dans ce domaine. Nous avons qualifié ces modèles de 'boîtes noires' au début de ce travail. Le modèle que nous avons choisi pour servir d'élément de comparaison fait partie de cette classe de modèles et ne comporte pas non plus de suivi de l'humidité au travers d'une variable interne. Il représente ainsi une alternative simple de modélisation aux structures de modèles à réservoirs que nous évaluons ici. Précisons qu'il n'entend pas être représentatif cependant de tous les modèles que nous avons classés parmi les modèles boîtes noires au chapitre 1. Il était donc intéressant dans un premier temps de comparer les performances des 38 structures à ce modèle simple.

La Figure 5.2 présente les performances moyennes (et quantiles) des modèles obtenues en phase de calage sur les 1284 bassins-périodes pour le critère d'optimisation $CR2$ ($\text{Nash}(\sqrt{Q})$). La Figure 5.3 (a, b, c, d, e, f) donne les résultats des 3204 tests réalisés en contrôle, jugés suivant les critères $CR1$ à $CR6$ (pour mémoire respectivement $\text{Nash}(Q)$, $\text{Nash}(\sqrt{Q})$, $\text{Nash}(\ln(Q))$, Nash-bis, erreur absolue, bilan). Les modèles ont été ordonnés en fonction des performances moyennes. Des classements réalisés sur le quantile 0,3 auraient été généralement assez peu différents. Notons cependant le cas des critères $CR1$ et $CR3$, pour lesquels des différences plus notables sont observées, ceci étant dû principalement à la plus grande part de critères négatifs (indiquée par les valeurs de quantiles 0,1 plus faibles que pour les autres critères).

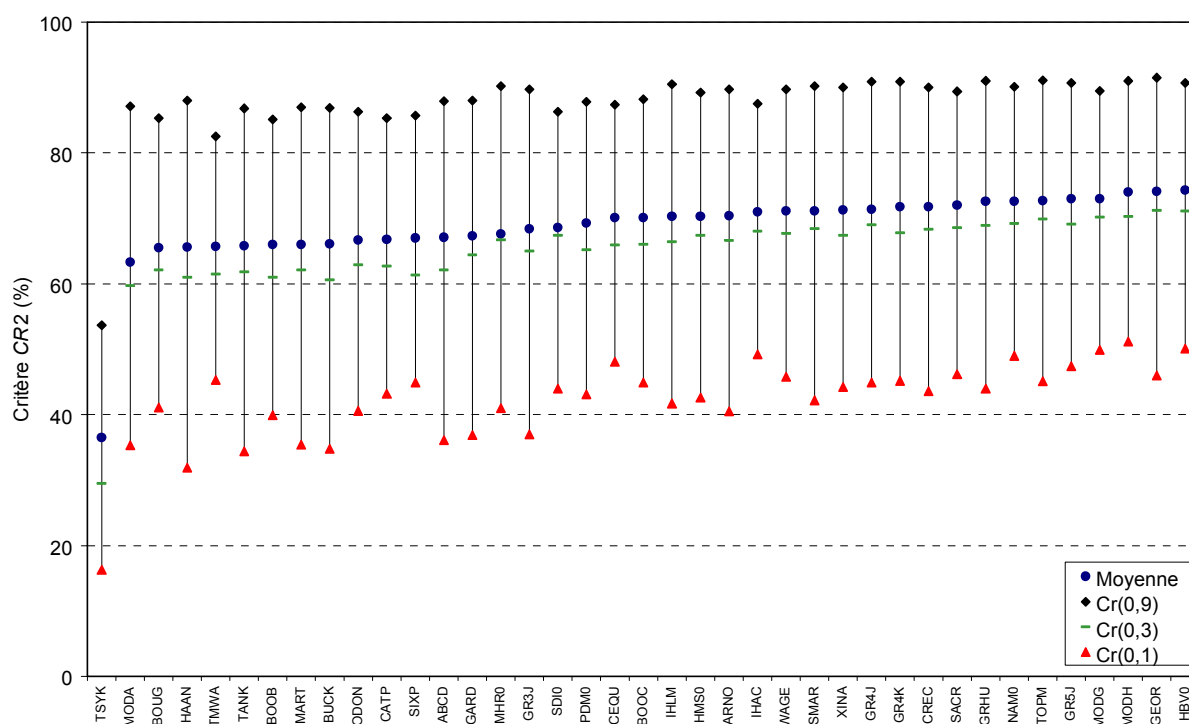


Figure 5.2: Performances des structures de modèles au calage pour le critère de calage $CR2$ ($\text{Nash}(\sqrt{Q})$), avec moyennes et quantiles 0,1, 0,3 et 0,9.

Une première observation qui ressort de ces graphiques est que les 38 structures de modèles sont systématiquement meilleures, en calage et en contrôle, que le modèle TSYK, toujours

classé en dernière position quel que soit le critère d'évaluation considéré. Ce modèle ne semble pas capable, en moyenne, de donner des résultats satisfaisants par rapport aux modèles à réservoirs, ce qui corrobore les remarques faites par Chiew *et al.* (1993) indiquant la même tendance au pas de temps journalier.

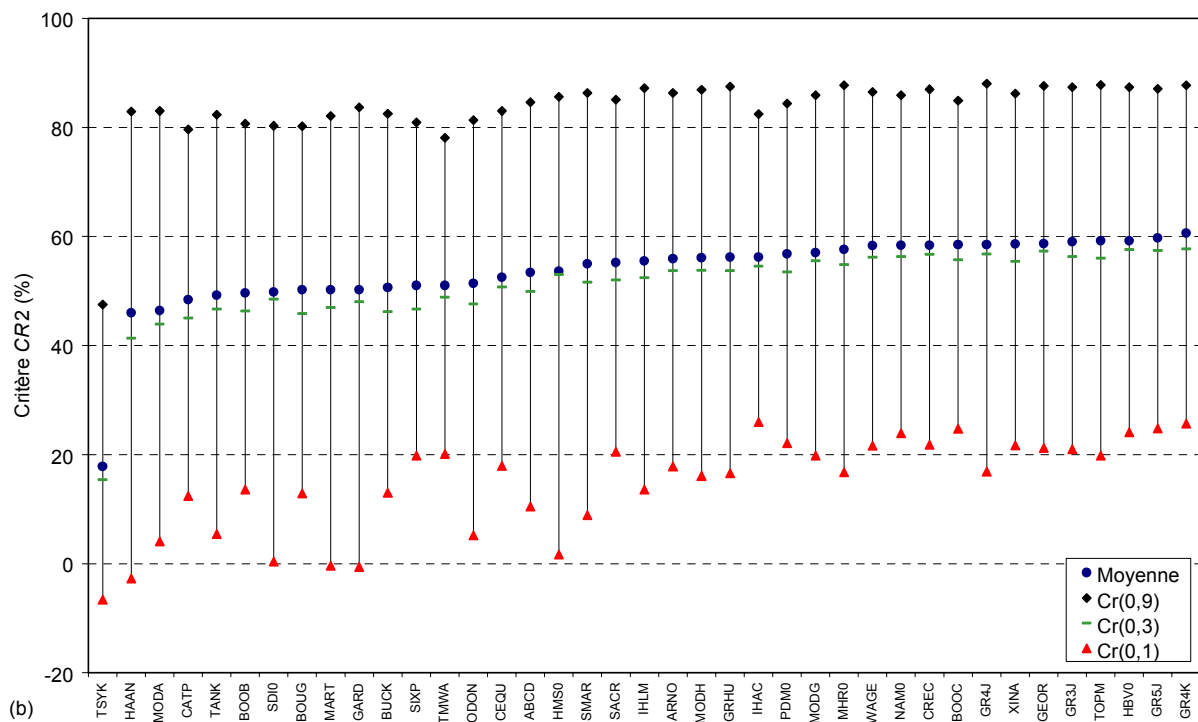
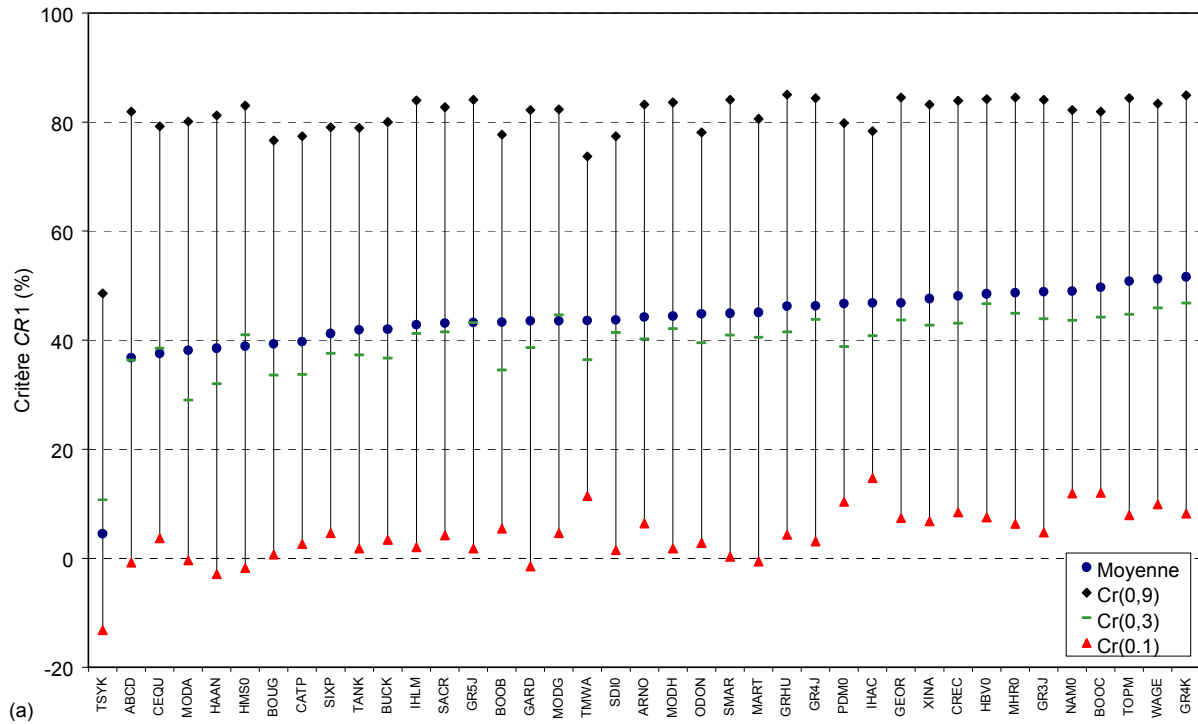
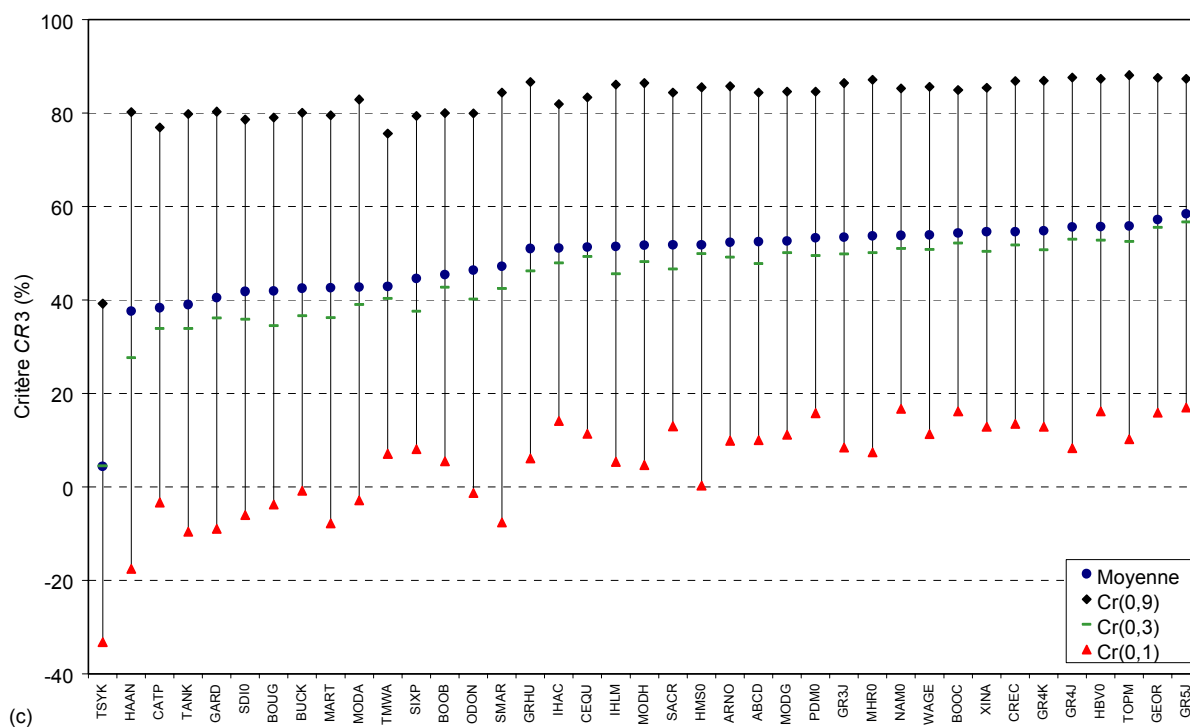
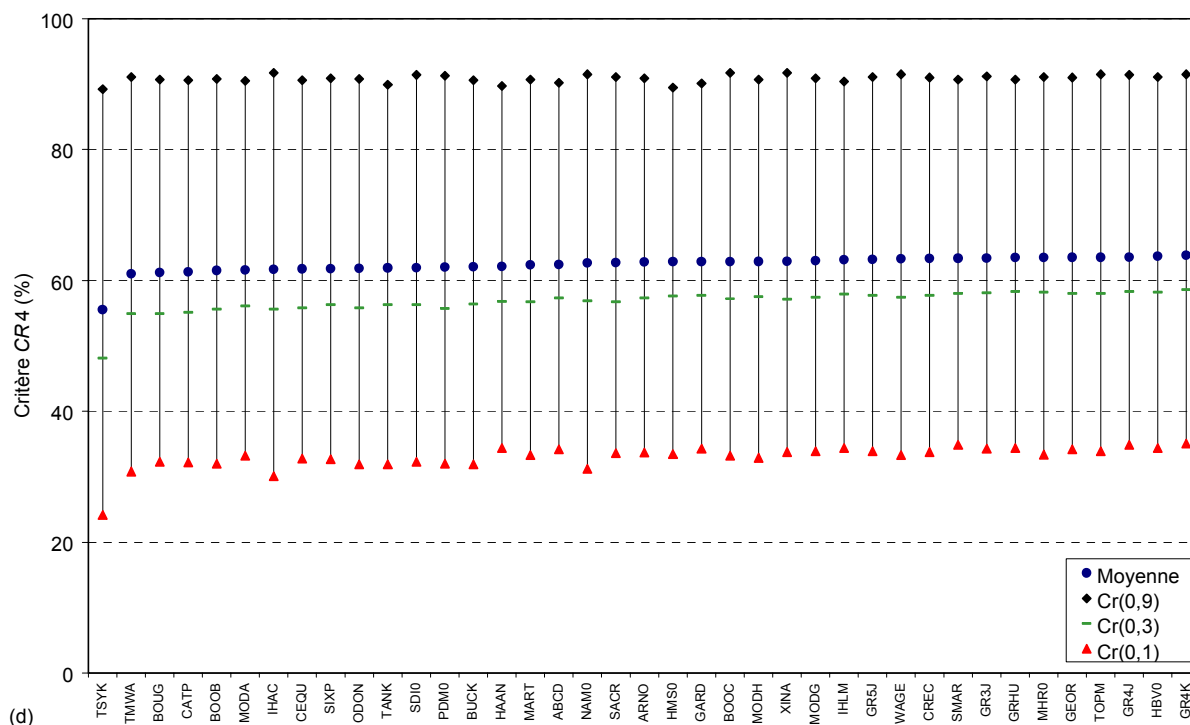


Figure 5.3: Performances des structures de modèles au contrôle pour les critère (a) CR1 (Nash(Q)) et (b) CR2 (Nash(\sqrt{Q})), avec moyennes et quantiles 0,1, 0,3 et 0,9



(c)



(d)

Figure 5.3 (suite): Performances des structures de modèles au contrôle pour les critère (c) $CR3$ (Nash($\ln(Q)$)) et (d) $CR4$ (Nash-bis), avec moyennes et quantiles 0,1, 0,3 et 0,9

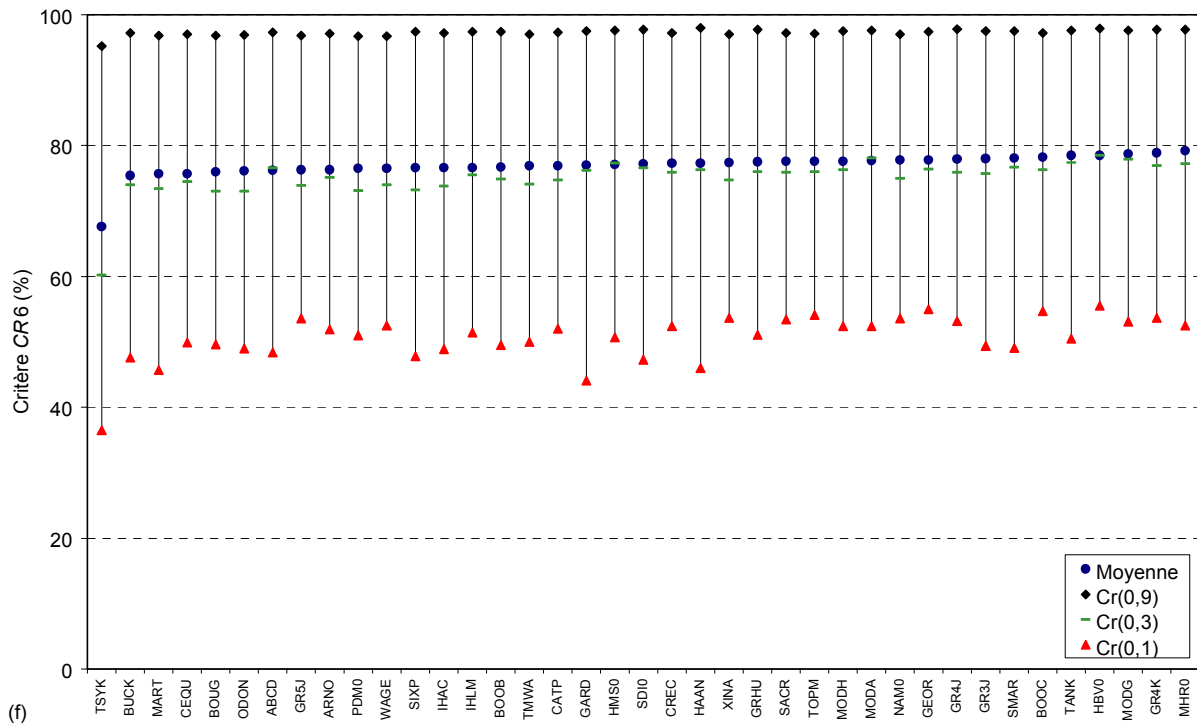
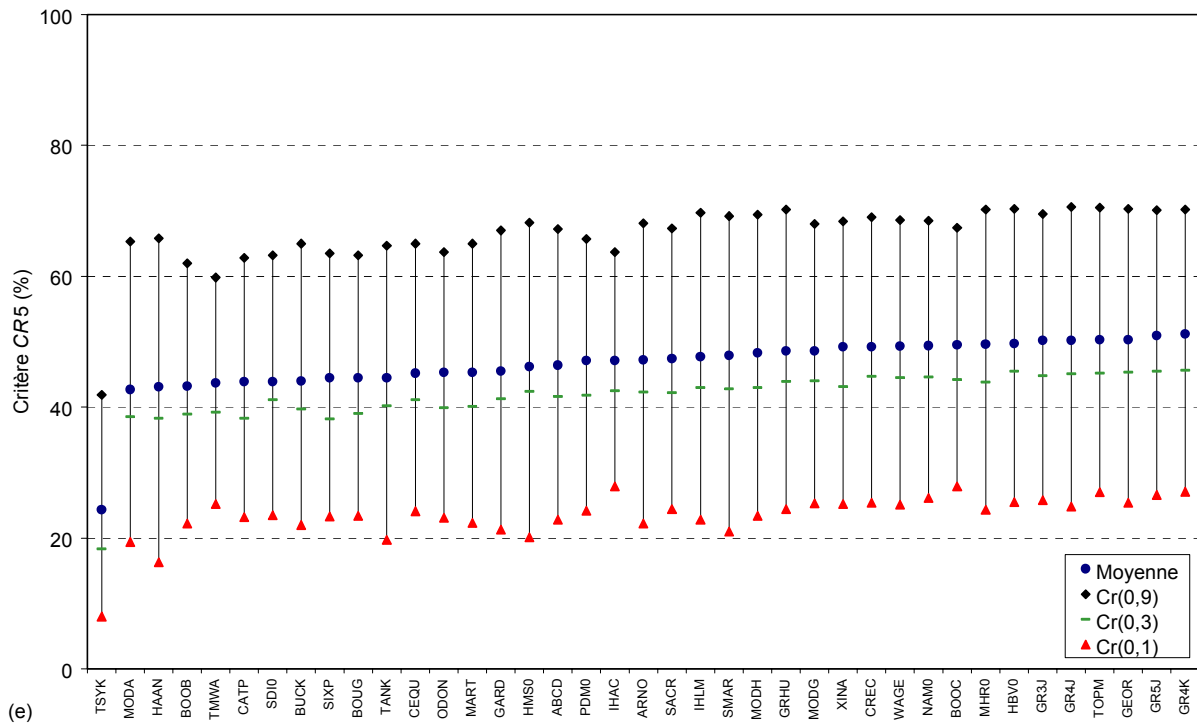


Figure 5.3 (suite): Performances des structures de modèles au contrôle pour les critère (e) CR5 (erreur absolue) et (f) CR6 (bilan), avec moyennes et quantiles 0,1, 0,3 et 0,9

Ce manque de pertinence est imputable à l'absence de réelle fonction de routage et au suivi trop rudimentaire des conditions antérieures d'humidité, qui ne dépendent dans ce modèle que des pluies et ETP antérieures. Toutes les structures de modèles à réservoirs utilisent au contraire une procédure de suivi (explicite ou implicite) des conditions d'humidité sur le bassin au cours du temps. Le recours à une (ou des) variable(s) interne(s) de suivi d'humidité dans le modèle semble donc, en moyenne, prouver une bonne efficacité pour la simulation des débits sur la plupart des bassins testés.

Les Figures 5.2 et 5.3 montrent par ailleurs que les performances des 38 modèles sont proches les unes des autres, quel que soit le critère utilisé, par rapport au modèle TSYK. Ceci apparaît aussi sur le Tableau 5.1 où sont donnés les écarts entre performances extrêmes pour les différents critères de qualité moyens. Les écarts sur la moyenne sont relativement réduits: moins de 10 points pour les critères *CR4*, *CR5* et *CR6*, et moins de 15 points pour les autres critères (excepté *CR3* en contrôle). Les écarts sur le critère *CR4* (Nash-bis) sont particulièrement réduits. Les quantiles 90 % sont encore plus proches, suggérant ainsi que tous les modèles sont capables, de façon très similaire, d'atteindre de hauts niveaux de performance sur des nombres équivalents de bassins-périodes. Ces hautes performances ne sont cependant pas obtenues par tous les modèles sur les mêmes bassins-périodes. Comparativement, les écarts entre les quantiles 0,1 sont plus importants, ce qui indique une plus grande variabilité des taux de forts échecs des modèles que des taux de francs succès. Cet effet est en partie dû à la forme des critères d'évaluation qui n'ont pas de borne inférieure.

		Quantile 0,1 (%)	Moyenne (%)	Quantile 0,9 (%)
<i>CR1</i>	Calage	22,9	12,6	10,5
	Contrôle	17,6	14,8	11,3
<i>CR2</i>	Calage	19,3	11,0	9,0
	Contrôle	28,7	14,6	9,9
<i>CR3</i>	Calage	23,1	12,4	10,3
	Contrôle	34,5	20,8	12,5
<i>CR4</i>	Calage	4,7	3,8	1,4
	Contrôle	5,0	2,8	2,2
<i>CR5</i>	Calage	8,2	8,1	10,7
	Contrôle	11,6	8,5	10,8
<i>CR6</i>	Calage	11,9	5,8	2,3
	Contrôle	11,4	3,8	1,3

Tableau 5.1: Ecart maximum entre modèles pour la moyenne et les quantiles 0,1 et 0,9 en calage et en contrôle sur les six critères de qualité

Même si certaines structures se retrouvent parmi les plus performantes pour une majorité des critères de qualité (c'est par exemple le cas des structures HBV0, TOPM ou GR4K), il apparaît qu'aucun des modèles n'est capable de surclasser tous les autres suivant tous les critères de performance. Ceci est explicité à la Figure 5.4 qui synthétise les rangs des structures dans les classements en contrôle présentés à la Figure 5.3 pour les six critères. Sur cette figure, nous avons ordonné les structures en fonction d'un rang 'moyen', moyenne arithmétique des rangs sur les six critères. On constate par exemple que le modèle GR5J, qui obtient de très bons rangs pour trois des critères, est assez nettement moins bon sur les trois autres. Par ailleurs, le critère de bilan *CR6* paraît le plus indépendant des autres: des modèles obtenant des bons (ou mauvais) résultats suivant ce critère peuvent inversement obtenir des mauvais (ou bons) résultats suivant les autres.

Ici ressort l'intérêt d'utiliser plusieurs critères de qualité complémentaires pour évaluer les performances, ce qui a pour effet d'augmenter le pouvoir de discrimination sur ces modèles.

Les modèles ayant été optimisés suivant le critère *CR2*, c'est en premier lieu suivant ce critère que doivent être jugés les modèles en contrôle. On leur demande en effet, en sélectionnant cette fonction objectif lors de l'optimisation, d'être performants suivant cet objectif. Or, si les modèles réussissent également à obtenir de bonnes performances suivant d'autres critères, cela renforce leur qualité intrinsèque en soulignant leur capacité à répondre également à des objectifs complémentaires, non explicitement mentionnés lors du calage des paramètres.

Les commentaires précédents sur la proximité des performances des structures peuvent être également illustrés par les observations suivantes, s'appuyant sur les performances moyennes par bassin. Tous les modèles obtiennent des valeurs du critère *CR1* en contrôle plus grandes que 60 % pour 97 bassins (23 % de l'échantillon), indiquant dans ces cas une relative aisance de tous les modèles. Inversement, tous les modèles obtiennent des performances moyennes inférieures à 60 % pour 138 bassins (32 % de l'échantillon). Ceci signifie que, pour environ un tiers des bassins, aucun des modèles testés ne réussit à apporter de réponses satisfaisantes quant à la simulation des débits observés. Quelles sont les raisons de cet échec des modèles ? L'inadéquation de leur structure est probablement la principale. Les erreurs dans les données de pluie ou de débit, les effets de nappes ou de fontes neigeuses, les influences d'origine humaine, sont aussi responsables de cette incapacité des modèles à simuler correctement les débits.

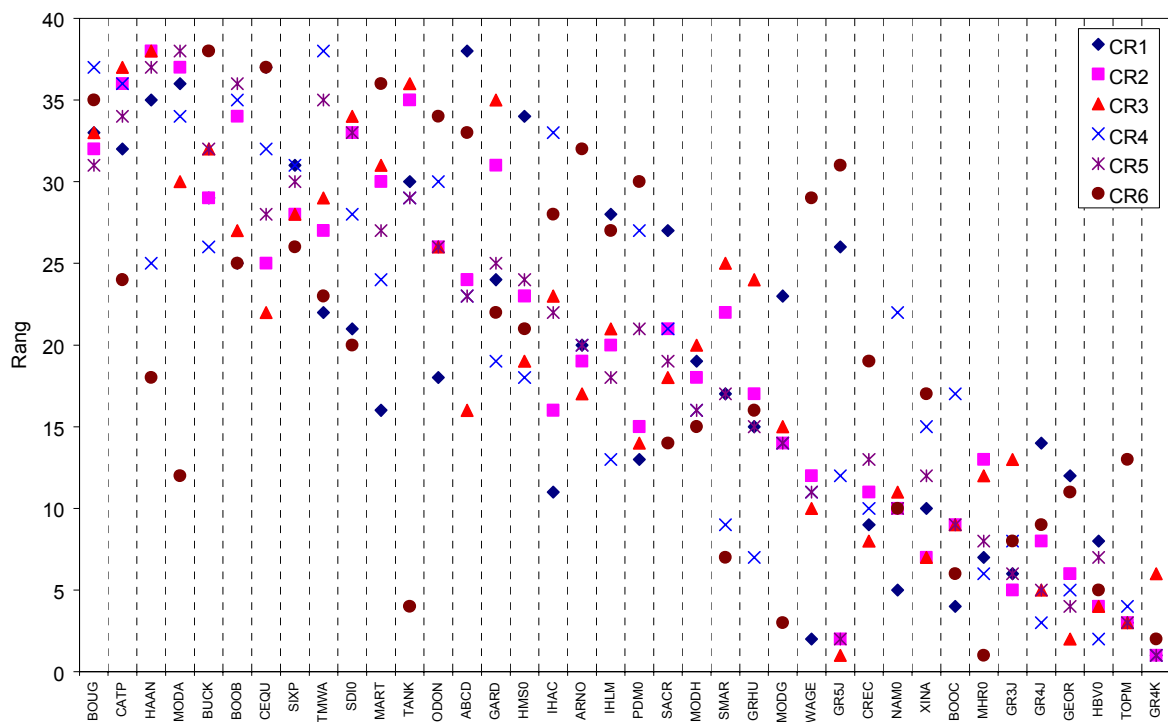


Figure 5.4: Rang des structures suivant les six critères de qualité au contrôle

Ces problèmes, liés à l'absence de sélection des bassins et de critique de données, comme nous l'avons expliqué au chapitre 2, sont également en partie responsables du relativement faible niveau des performances moyennes obtenues en contrôle (inférieures à 52 % dans le cas de *CR1* par exemple). Pour quantifier cet effet, nous avons retiré de notre échantillon les 42 bassins pour lesquels les données de débit avaient été jugées peu propices à la modélisation pluie-débit au chapitre 2. LaFigure 5.5 compare les performances obtenues en contrôle pour les critères *CR1* et *CR2*. Le fait d'éliminer les 42 bassins litigieux rehausse le niveau général des performances moyennes, avec une proportion moins importante de faibles valeurs comme en témoigne le rehaussement des quantiles $Cr(0,1)$. Par ailleurs, on constate que l'absence de

critique des données a peu d'influence sur les performances relatives des modèles. Si dans le cas du critère $CR1$, le classement obtenu sur les performances moyennes serait légèrement différent sur le sous-échantillon de 387 bassins, il serait quasiment identique dans le cas du critère $CR2$. Quel que soit le critère considéré, la tendance générale est cependant la même, indiquant que le choix de ne pas éliminer des bassins a priori ne remet pas en cause les conclusions de notre étude.

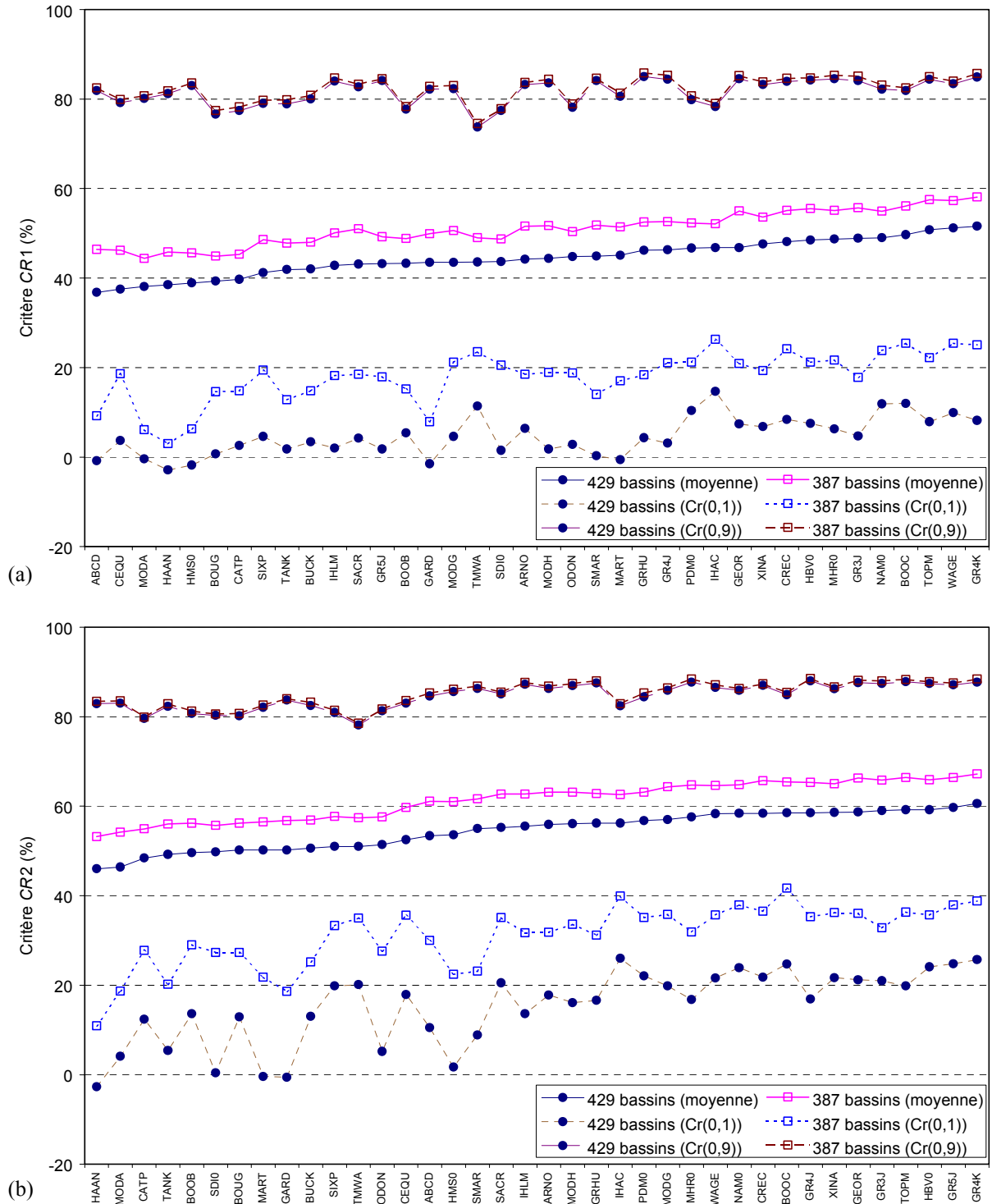


Figure 5.5: Comparaison des performances avec et sans critique préalable des données pour les critères (a) $CR1$ et (b) $CR2$

Remarquons également que l'application des modèles en dehors de leur domaine de validité préconisé par les concepteurs n'a pas paru être un facteur limitant de façon significative leur efficacité numérique. Des modèles développés pour des climats semi-arides ont pu donner de bons résultats en climat humide, et inversement. Enfin, il n'a pu être établi de lien entre la performance des modèles et la superficie des bassins versants testés. La Figure 5.6 montre dans le cas du modèle GR4K que de bonnes performances ($\text{Nash}(\sqrt{Q}) > 80\%$) ont pu être obtenues sur toute la gamme de variation des superficies, même pour les deux plus grands bassins qui dépassent 40000 km².

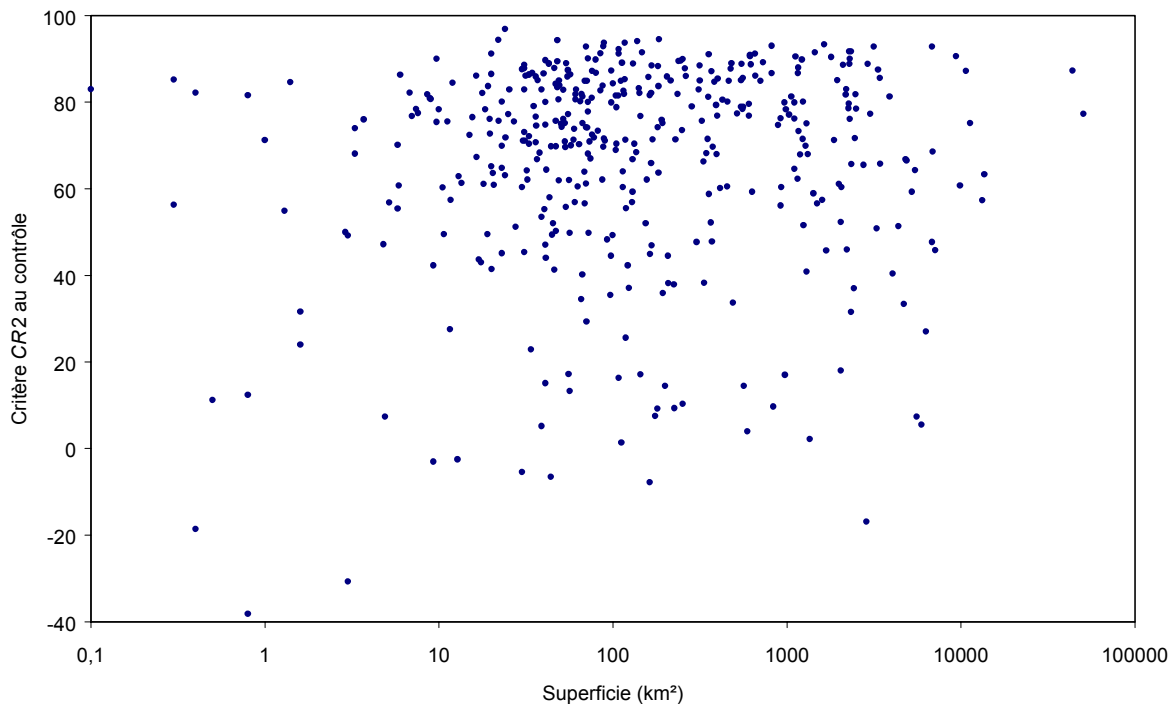


Figure 5.6: Lien entre superficie des 429 bassins versants de l'échantillon et performances moyennes par bassin obtenues en contrôle par le modèle GR4K suivant le critère *CR2*

5.2.2. Comment les résultats en calage et contrôle illustrent la robustesse des modèles ?

A l'instar de la fiabilité dont nous discutons dans la partie suivante, la robustesse est l'une des qualités les plus importantes et les plus souhaitables des modèles permettant de leur conférer une bonne crédibilité. Nous proposons ici de quantifier la robustesse par la diminution des performances moyennes lorsque l'on passe de la phase de calage à la phase de validation. On évalue ainsi la capacité du modèle à fournir en phase de contrôle, sur des données différentes de celles utilisées en calage, des performances de même qualité qu'en phase d'optimisation. Cette diminution est un comportement bien connu des modèles hydrologiques.

Le Tableau 5.2 donne les diminutions de performance moyenne, minimales et maximales que l'on obtient avec les 38 structures lorsque l'on passe du calage au contrôle. Il existe une diminution systématique et significative des performances. Cependant, ce manque de stabilité des résultats n'est pas le même pour toutes les structures. Ainsi, le Tableau 5.2 indique, dans le cas du critère *CR2*, que la chute de performance va de 9,4 % à 19,6 % suivant les modèles. Ceci signifie donc que certains modèles sont plus robustes que d'autres, c'est-à-dire que leurs résultats sont plus stables et qu'ils sont donc davantage capables de fournir des simulations de qualité comparable à celles obtenues en phase de calage.

	CR1 Nash (Q)	CR2 Nash(VQ)	CR3 Nash(ln(Q))	CR4 Nash-bis	CR5 Err. abs.	CR6 Bilan
Différence moyenne(%)	-17,8	-15,0	-17,8	-1,7	-8,8	-9,1
Différence maximale (%)	-25,4	-19,6	-25,6	-3,0	-11,6	-11,4
Différence minimale (%)	-11,9	-9,4	-11,3	-1,1	-5,2	-7,1

Tableau 5.2: Diminutions maximales, minimales et moyennes du critère CR2 pour les 38 structures de la phase de calage à la phase de contrôle

5.2.3. Certaines structures peuvent-elles garantir une meilleure fiabilité ?

Les résultats précédents sont maintenant analysés sous l'angle de la fiabilité des modèles. Nous proposons de quantifier la fiabilité du modèle de manière relative, par le nombre de bassins où le modèle est classé parmi les structures les plus performantes. Pour cela, on utilise les performances moyennes par bassin en contrôle, et on effectue sur chaque bassin le classement des structures. Sur un bassin, un modèle sera dit de classe 1 s'il se place parmi les six structures les plus performantes et de classe 2 s'il se classe entre la septième et la douzième place, c'est-à-dire dans le tiers supérieur des structures.

En utilisant le critère CR2 (critère de Nash-Sutcliffe sur les racines carrées des débits), la Figure 5.7 montre le nombre de bassins pour lesquels le modèle est de classes 1 et 2. Ce nombre va de 25 à 289 (et entre 7 et 204 pour la classe 1 seulement). Cette disparité indique que certains modèles obtiennent comparativement de meilleures performances sur bien plus de bassins que d'autres et sont de ce fait plus fiables. Par ailleurs, nous avons observé qu'il a été possible de trouver, pour 33 des 38 structures, au moins un bassin où l'une de ces structures était la plus satisfaisante suivant le critère CR2. Si l'on considère le critère CR1, pour 36 des 38 structures, il existe au moins un bassin pour lequel une structure donnée est la plus performante.

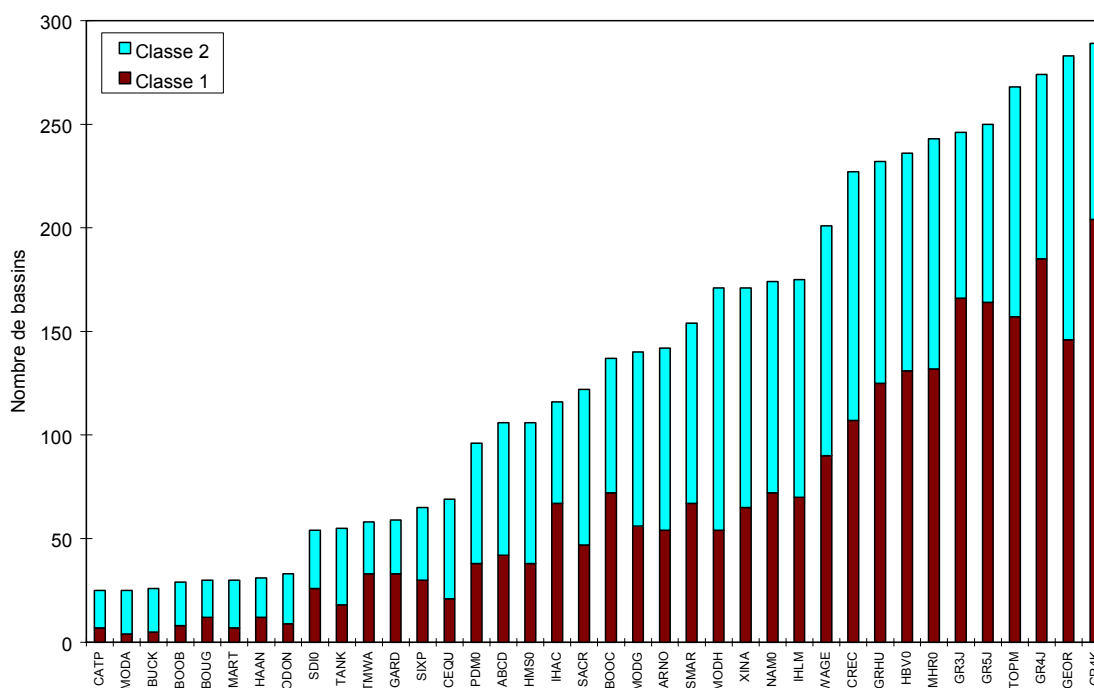


Figure 5.7: Nombres de bassins pour lesquels le modèle est de classes 1 et 2

Ceci tend à indiquer que, si l'on avait choisi un bassin particulier ou un petit nombre de bassins, comme cela a été fait dans les études comparatives antérieures, les conclusions de notre comparaison auraient été beaucoup plus dépendantes de ce choix qu'elles ne risquent de l'être ici. L'avantage de disposer d'un large échantillon de bassins est d'obtenir une significativité statistique des performances.

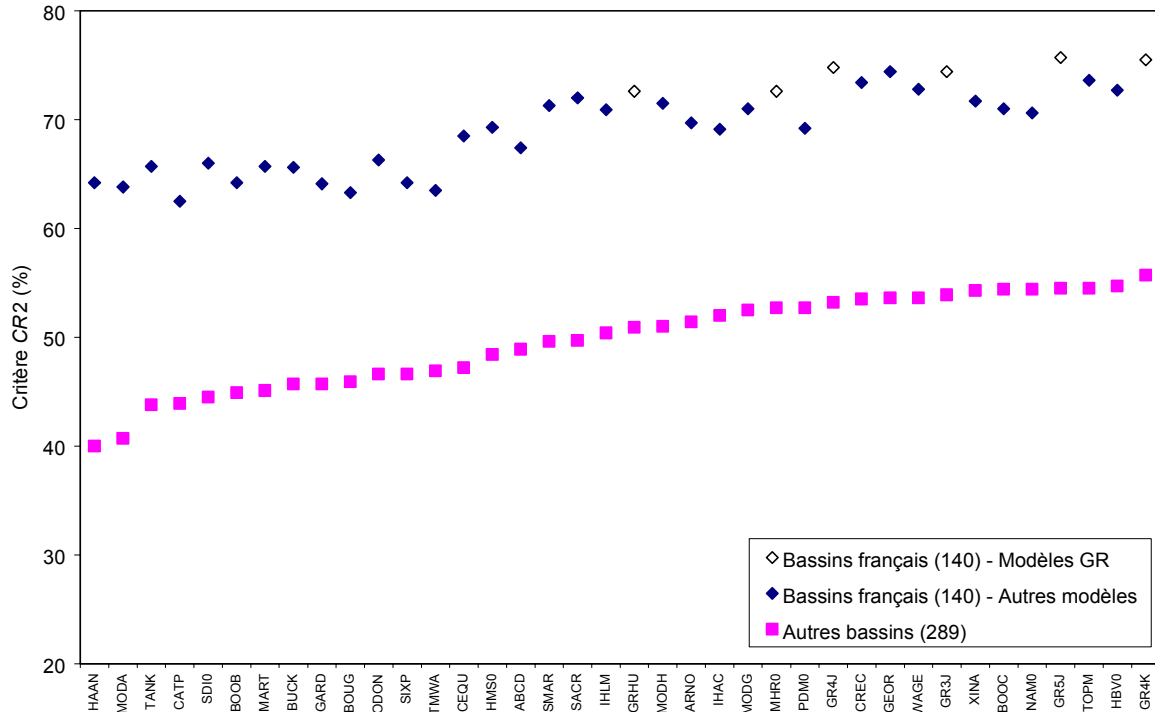


Figure 5.8: Performances moyennes des structures suivant le critère $CR2$ ($Nash(\sqrt{Q})$) sur les sous-échantillons de données (a) françaises et (b) étrangères

Pour illustrer cet effet d'échantillonnage, nous avons scindé notre échantillon en deux, une partie constituée par les 140 bassins français utilisés par Edijatno *et al.* (1999) pour la mise au point du modèle GR3J, le restant des bassins (289) constituant le deuxième sous-échantillon. La Figure 5.8 présente les performances moyennes obtenues par les structures sur ces deux sous-échantillons suivant le critère $CR2$ ($Nash(\sqrt{Q})$). Elles ont été ordonnées suivant les performances obtenues sur les 289 bassins.

Plusieurs points sont à noter. Il apparaît tout d'abord que les niveaux de performance et les classements des structures sont clairement différents d'un groupe de bassins à l'autre. En revanche, on remarque une tendance générale similaire sur les deux sous-groupes: un modèle très satisfaisant dans un cas ne sera pas complètement décevant dans l'autre. Cette relative cohérence des résultats est assez intéressante. Elle tend à indiquer que lorsqu'un grand nombre de bassins versants sont utilisés pour tester des modèles, on réussit à dégager quelques tendances sur la valeur des modèles, alors que les conclusions de précédentes études comparatives ayant porté sur un petit nombre de bassins pouvaient sembler contradictoires sur le jugement de certains modèles. Le fait que les classements obtenus dans les deux cas ne soient pas exactement les mêmes prouve par ailleurs que la richesse introduite par l'ajout de bassins est toujours bénéfique pour éprouver les modèles.

Enfin, on peut noter que, sur le sous-échantillon de bassins français, les structures de la famille GR (GR3J, GR4J, GR4K, GR5J, GRHU et MHR0), identifiées avec des symboles différents sur la Figure 5.8, obtiennent des classements meilleurs que sur l'échantillon complémentaire, pour tous se positionner dans les toutes meilleures structures. Ceci tend à

prouver la bonne adéquation de ces modèles à l'échantillon de bassins ayant servi de support à leur mise au point.

5.2.4. Peut-on parler d'équifinalité entre structures de modèles ?

Le concept d'équifinalité en modélisation (voir par exemple Beven, 1993) décrit le fait que, pour un modèle donné, des jeux de paramètres différents peuvent donner des résultats équivalents en terme de simulation des débits et de performance. Comme nous en avons déjà discuté aux chapitres 1 et 3, ce phénomène est dû principalement à la sur-paramétrisation ou à la formulation maladroite des structures. Nous voulons explorer ici la possibilité d'étendre ce concept aux structures de modèles, en regardant si des structures différentes peuvent donner des résultats similaires sur un bassin ou un échantillon de bassins. La similarité sera ici quantifiée par la proximité des critères de performance. Nous fonderons notre analyse sur les performances moyennes par bassin suivant le critère $CR1$ (Nash(Q)) en contrôle:

- pour 61 bassins, il n'y a aucune différence dans les résultats des deux meilleures structures;
- pour 80 % des bassins, la différence entre le premier et le deuxième modèle est inférieure à 2 % et elle est supérieure à 5 % pour seulement 31 des 429 bassins (7 % de l'échantillon). Sur seulement 7 de ces 31 bassins, la performance de la meilleure structure excède 60 %, ce qui indique que les plus grandes différences entre les deux meilleures structures se produisent majoritairement lorsque aucune d'entre elles ne réussit à apporter de réponse satisfaisante à la simulation des débits;
- enfin, la différence entre le premier et le troisième modèle est inférieure à 2 % pour 63 % des bassins, et la différence entre le premier et le 19^{ème} modèle (moitié des structures) est inférieure à 10 % pour 252 bassins (59 % de l'échantillon),

Ces remarques tendent à aller dans le sens d'une sorte de principe d'équifinalité entre structures de modèles: pour une importante proportion de bassins, des modèles différents peuvent donner des résultats très similaires. Les résultats présentés par Franchini et Pacciani, (1991) et complétés par Franchini *et al.* (1996) montrent également dans le cas du bassin de la Sieve en Italie, que huit modèles conceptuels avec des structures différentes peuvent donner des résultats très proches en simulation.

Bien entendu, l'utilisation de plusieurs critères pour mener l'analyse précédente aurait probablement restreint l'étendue des similarités des performances entre modèles. Cependant, il semble certain qu'elle n'aurait pas annulé cette tendance à la similarité. Nous pouvons par ailleurs rapprocher ces remarques de l'analyse faite au premier chapitre sur les comparaisons existantes de modèles à réservoirs: nous avons émis l'hypothèse que le manque de tendance nette dans les comparaisons de modèles quant à la valeur relative des différents modèles pourrait être due à une grossière équivalence dans les capacités de ces structures. Nos résultats tendent à aller dans le sens de cette hypothèse. Bien sûr, cette équivalence est essentiellement relative aux performances beaucoup plus modestes du modèle TSYK.

5.2.5. Est-ce que le nombre de paramètres accroît le niveau de performance ?

Au cours des 40 dernières années, de nombreux concepteurs de modèles ont recouru à l'accroissement de la complexité des modèles pour essayer de trouver une réponse aux échecs des modèles pluie-débit et tenter d'améliorer les performances des modèles.

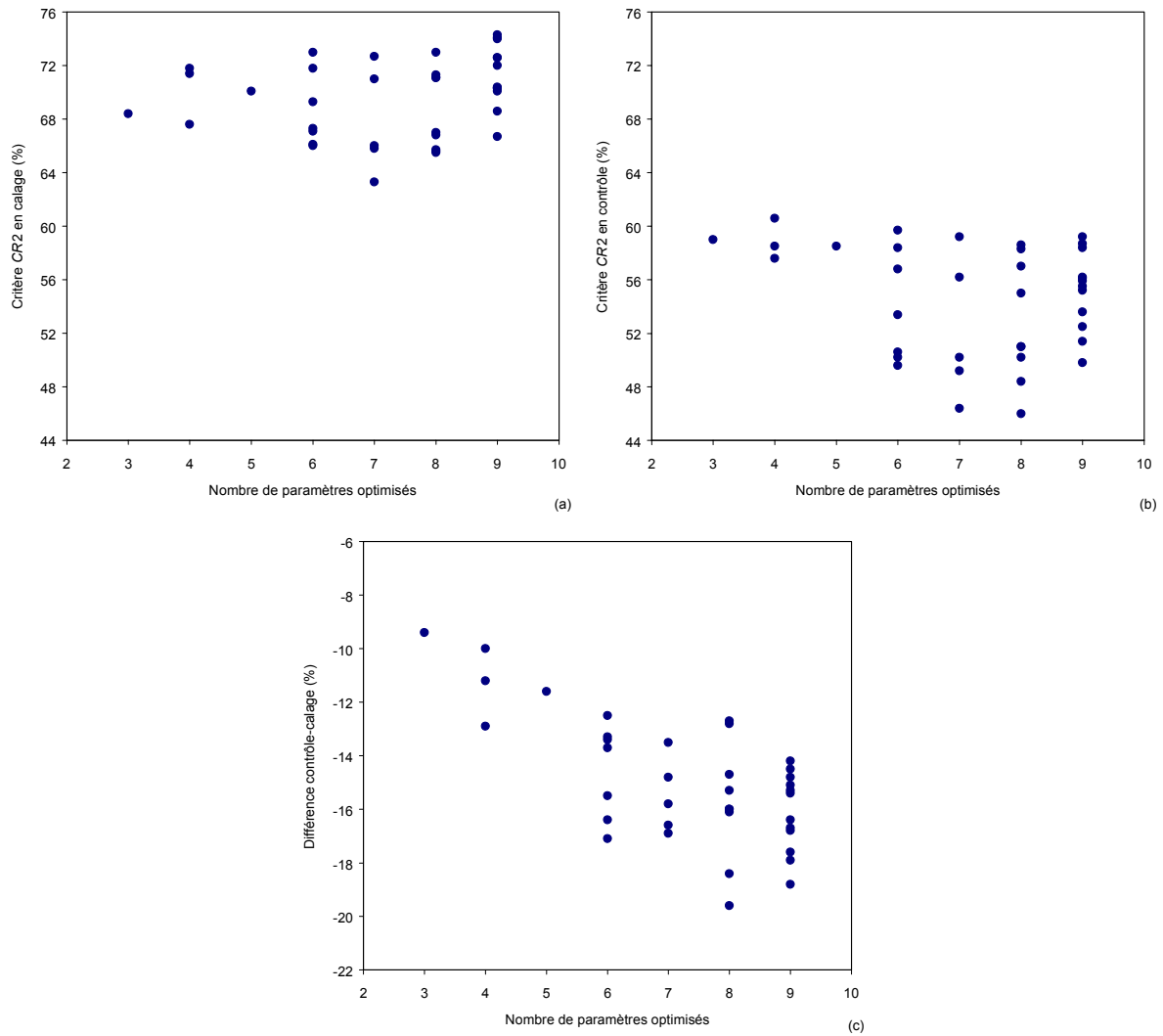


Figure 5.9: Performances moyennes des modèles en fonction du nombre de paramètres optimisés, suivant le critère $CR2$ (Nash(\sqrt{Q})) (a) au calage et (b) au contrôle; (c):diminution des performances entre calage et contrôle

Si la complexité peut être entendue de diverses manières (niveau de description des processus, discrétisation temporelle ou spatiale...), nous ne l'entendons ici que dans le sens du nombre de paramètres laissés libres dans le modèle, c'est-à-dire optimisés en phase de calage. La Figure 5.9 montre les performances des modèles suivant le critère $CR2$ au calage et au contrôle en fonction du nombre de paramètres optimisés. En phase de calage, les modèles avec un nombre élevé de paramètres bénéficient généralement de ce supplément de degrés de liberté pour obtenir de meilleures performances en réussissant à mieux reproduire les débits observés. En revanche, cette tendance n'apparaît plus au contrôle, où les modèles avec un nombre limité de paramètres réussissent à obtenir des performances équivalentes aux autres modèles.

Gan *et al.* (1997) et Ye *et al.* (1997) ont également signalé cette absence de supériorité des modèles les plus complexes sur les modèles les plus simples. Ceci peut être partiellement expliqué par la stabilité des performances du calage au contrôle, comme le montre la Figure 5.9(c). En lien avec l'analyse précédente sur la robustesse des modèles, ces résultats confirment que les modèles les plus complexes ont tendance à être moins robustes, c'est-à-dire qu'ils ont des performances moins stables que les modèles plus simples lors du passage du calage au contrôle.

Par ailleurs, que ce soit en calage ou en simulation, des modèles avec un même nombre de paramètres optimisés peuvent donner des niveaux de performances assez différents. Ceci renforce l'idée avancée par Gan *et al.* (1997) que la structure du modèle, c'est-à-dire la nature des réservoirs, la façon dont les éléments de la structure sont reliés, la paramétrisation des fonctions de production ou de routage, sont des points essentiels pour le succès d'un modèle. La complexité, liée à une plus grande distinction de processus élémentaires, ne peut garantir des performances fiables et de niveau élevé.

Nous pouvons dégager de ce qui précède une conclusion importante: la formulation des structures des modèles globaux à réservoirs ne semble actuellement pas suffisamment pertinente pour pouvoir supporter un degré de complexité élevé.

La Figure 5.9 montre enfin que des structures avec des niveaux de complexité variés peuvent donner des résultats équivalents. Une conclusion possible, déjà avancée par Nash et Sutcliffe (1970) serait alors de privilégier l'utilisation des plus simples de tous ces modèles équivalents, minimisant ainsi les problèmes de sur-paramétrisation et de détermination des paramètres. Cette idée est discutée plus en détail dans le paragraphe suivant.

5.2.6. Complémentarité de structures de modèles: est-il possible de se hisser jusqu'au modèle 'idéal' ?

Nous avons mentionné précédemment que, parmi les 38 structures de modèles testées, toutes n'obtiennent pas des performances très élevées sur les mêmes bassins et qu'il est presque toujours possible de trouver un bassin sur lequel une structure donnée sera meilleure que toutes les autres. Nous pouvons alors nous poser la question de savoir s'il est judicieux de vouloir exploiter cet avantage de certains modèles sur des bassins donnés, c'est-à-dire une possible complémentarité entre structures.

Pour mettre en évidence cette complémentarité, nous adoptons dans un premier temps une démarche uniquement fondée sur les résultats obtenus par les structures au contrôle. Dans un second temps, nous essaierons d'exploiter une correspondance entre bassins et modèles. Ceci est présenté dans la partie suivante de ce chapitre. Nous présentons maintenant les résultats de la première approche.

Plaçons nous dans la position d'un modélisateur qui aurait le choix entre les 38 structures de modèles pour modéliser la relation pluie-débit sur chaque bassin. Dans chaque cas, le test de toutes les structures permettrait de déterminer la plus performante et le modélisateur pourrait ainsi bénéficier de la structure la mieux adaptée à son cas d'étude. Nous avons procédé ainsi sur notre échantillon de 429 bassins versants, en retenant à chaque fois la meilleure performance fournie par les 38 structures. Nous considérons maintenant ce jeu de meilleures performances comme celles dérivant d'un même modèle 'idéal' hypothétique, sorte de modèle enveloppe, que nous nommerons dans la suite structure IDEA.

Précisons que ce modèle 'idéal' a été construit à partir des performances obtenues par les modèles en contrôle et non sur celles du calage. Il correspond en fait, a posteriori, au meilleur modèle composite, une association 'idéale' des 38 structures. Bien entendu, il n'est 'idéal' que dans la limite de ces 38 structures et il reste probablement différent du modèle parfait qui se caractériserait par une distribution des Nash liée aux seules erreurs de mesures.

La Figure 5.10 montre les distributions des valeurs du critère $CR1$ (Nash(Q)) obtenues en contrôle par le modèle de référence TSYK, le modèle 'idéal' IDEA et les 38 structures de modèles. La Figure 5.13 précise la composition de ce modèle 'idéal': parmi les plus fortes contributions, on retrouve les modèles jugés les plus fiables au paragraphe 5.2.3. Sur notre échantillon de bassins, le modèle IDEA atteint une moyenne du critère $CR1$ de 64,4 % alors

que la meilleure des structures atteint seulement une moyenne de 54,9 % (valeurs obtenues à partir des moyennes des performances par bassin). Cela indique qu'il existe bien une complémentarité entre structures et que son exploitation peut permettre des progrès significatifs.

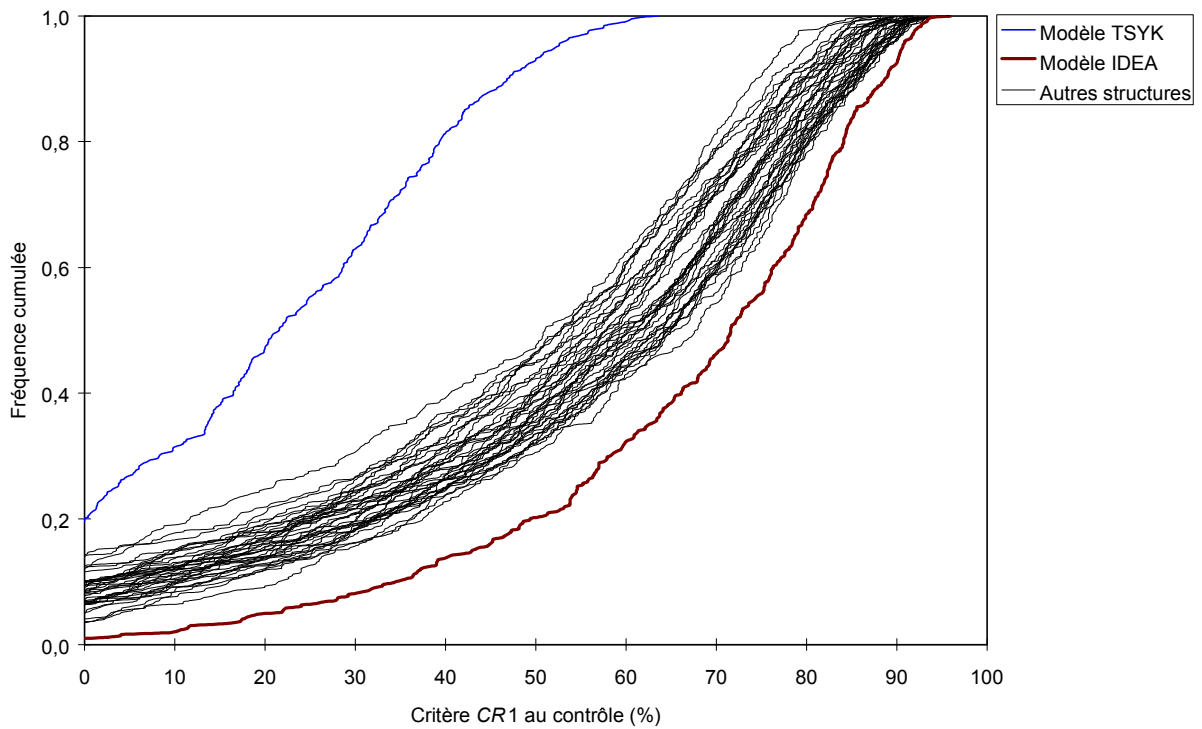
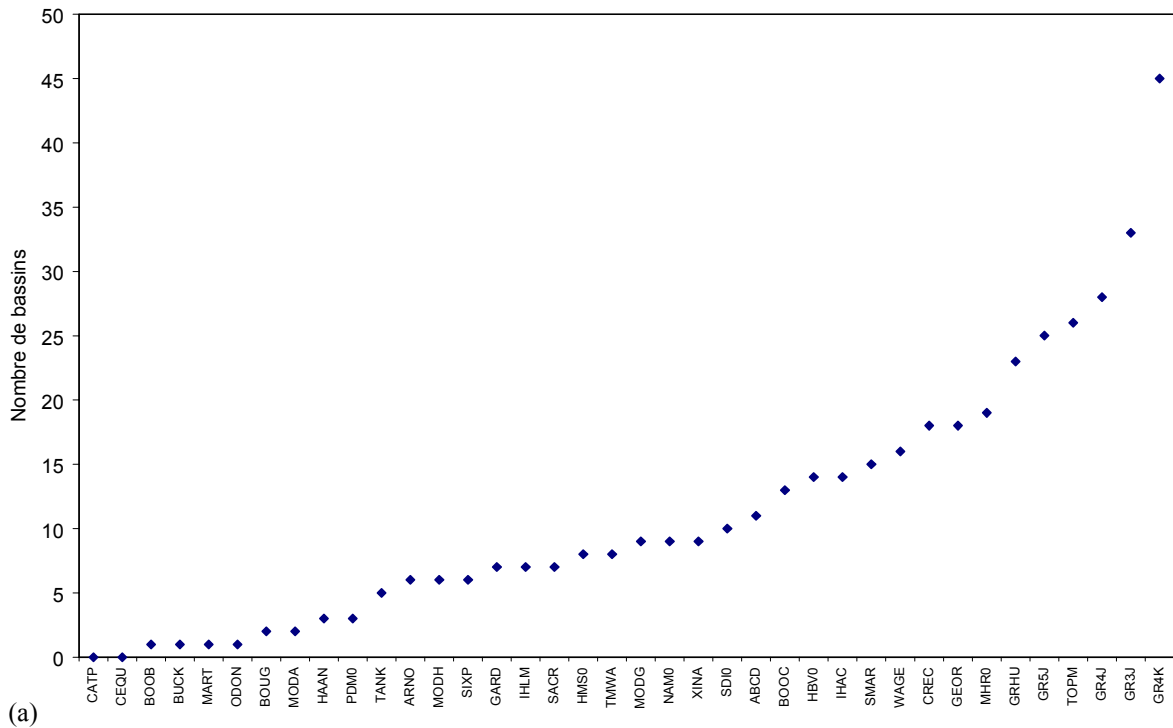


Figure 5.10: Distribution des performances des modèles TSYK et IDEA et des 38 structures sur l'échantillon des 429 bassins en contrôle suivant le critère CR1



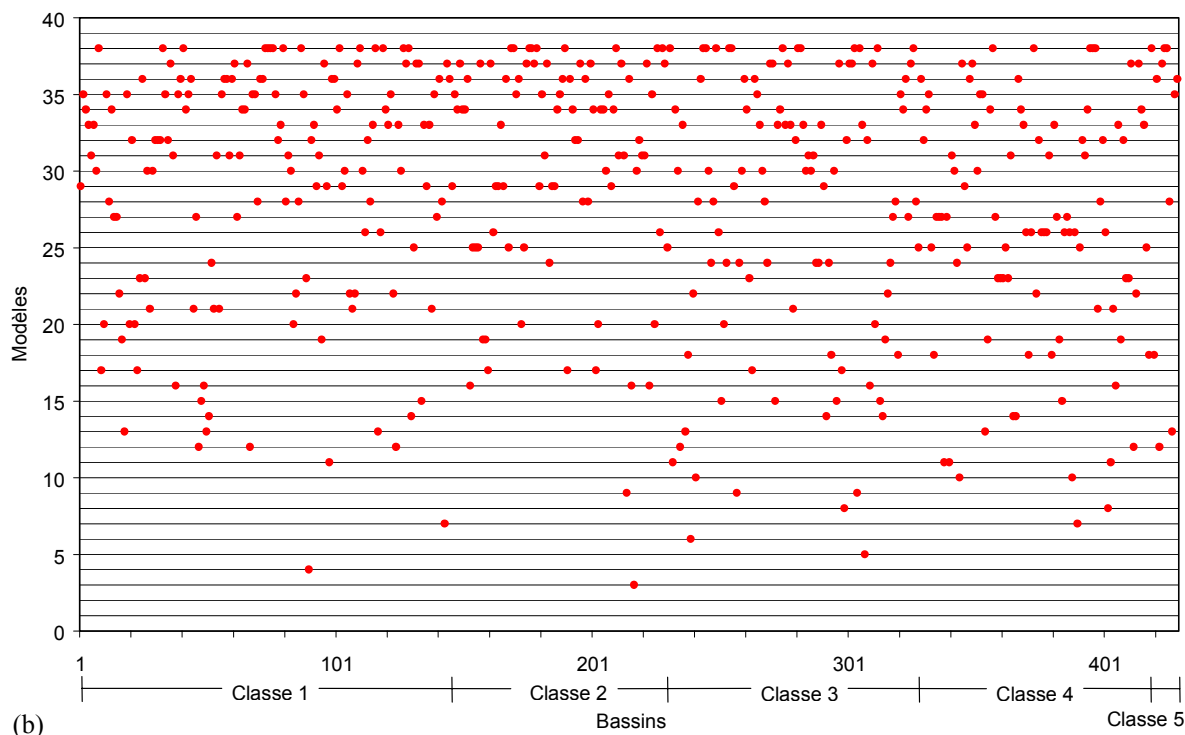


Figure 5.11: Composition du modèle idéal suivant le critère *CR1*: (a) nombre de bassins où les modèles ont été classés premiers ; (b) graphe de correspondance bassins-modèles (modèles dans le même ordre que la figure (a) et bassins ordonnés en classes (Classe 1: 1 à 144; Classe 2: 145 à 227; Classe 3: 228 à 327; Classe 4: 328 à 417; Classe 5: 418 à 429))

Remarquons cependant que ce modèle 'idéal' qui regroupe potentiellement toutes les qualités des 38 structures, n'obtient en contrôle des critères de Nash supérieurs à 80 % que dans trois cas sur dix (voir les distributions de la Figure 5.10). Cette proportion tombe à presque deux bassins sur dix dans le cas des meilleures des 38 structures. Autrement dit, il demeure encore des grosses difficultés à modéliser la transformation pluie-débit à l'échelle du bassin versant, même en associant plusieurs modèles: on réussit à obtenir des résultats satisfaisants (modèle exploitable avec une relativement bonne fiabilité) au mieux pour 30 % des bassins de l'échantillon test.

Pour tester davantage cette idée de complémentarité, nous avons cherché les structures dont l'association par paire donnait les meilleurs résultats. Nous utilisons pour cela les résultats en contrôle suivant le critère *CR2* (critère de Nash sur les racines carrées des débits). Le Tableau 5.3 fournit, pour les vingt meilleures structures, la matrice symétrique donnant le nombre de bassins pour lesquels au moins une des deux structures (en ligne et colonne) est classée parmi les quatre meilleures structures de l'échantillon. On se réfère donc à la notion de fiabilité exposée précédemment. Sur la diagonale, on trouve le nombre de bassins pour lesquels un modèle seul est classé parmi les quatre meilleures structures.

L'association de deux modèles améliore de façon claire la fiabilité de la paire ainsi construite. Prenons l'exemple d'une des meilleures associations (en gras dans le Tableau 5.3) entre les structures GR4K et GEOR. Associées, l'une des deux au moins réussit à se classer parmi les quatre structures les plus performantes pour 292 bassins, alors qu'individuellement, elles n'obtiennent ce classement que pour 204 bassins au maximum. Une amélioration de presque 90 bassins (20 % de notre échantillon) peut être obtenue lorsque deux modèles sont associés par rapport au cas où les modèles sont utilisés séparément.

	ARNO	BOOC	CREC	GEOR	GR3J	GR4J	GR4K	GR5J	GRHU	HBV0	IHAC	IHLM	MHR0	MODG	MODH	NAM0	SMAR	TOPM	WAGE	XINA
ARNO	54																			
BOOC	115	72																		
CREC	150	169	107																	
GEOR	181	198	213	146																
GR3J	214	216	245	273	166															
GR4J	224	244	261	281	250	185														
GR4K	244	243	286	292	229	267	204													
GR5J	203	211	242	257	262	278	164													
GRHU	168	174	196	228	229	255	257	247	125											
HBV0	176	191	192	228	259	264	292	262	213	131										
IHAC	112	117	159	199	214	239	253	217	179	183	67									
IHLM	117	134	165	189	219	223	243	206	183	174	135	70								
MHR0	176	197	207	230	241	244	266	252	220	207	196	176	132							
MODG	107	122	148	185	210	232	248	208	168	165	109	118	174	56						
MODH	106	117	152	185	202	215	231	200	166	174	118	111	167	101	54					
NAM0	111	133	160	190	227	241	259	215	185	182	127	138	189	119	121	72				
SMAR	114	137	158	200	209	228	246	222	167	169	126	131	185	106	116	129	67			
TOPM	188	223	230	247	270	261	290	243	256	246	216	199	245	200	191	198	202	157		
WAGE	129	148	177	214	227	242	253	210	199	205	147	149	208	138	133	143	149	199	90	
XINA	97	127	159	191	218	235	250	206	186	180	120	124	189	115	116	120	120	193	141	65

Tableau 5.3: Partie de la matrice de complémentarité montrant les meilleures associations entre structures (construite à partir des performances moyennes par bassin suivant le critère CR2 (Nash(\sqrt{Q})))

Ceci indique qu'il y a dans chaque modèle des composants spécifiques dont l'efficacité peut être complétée de façon efficace par d'autres. De ce fait, l'association de modèles peut être un des remèdes partiels aux limites de chacun d'eux. Chercher le modèle qui donne les meilleures performances lorsqu'il est associé à un modèle donné pourrait ainsi indiquer quels composants du premier pourraient être introduits dans le second avec le plus de succès. Idéalement, en associant les 38 structures, on obtiendrait le modèle IDEA, avec des gains de performance très substantiels. Dans ce sens, Shamseldin *et al.* (1997), ou plus récemment Shamseldin et O'Connor (1999) ont montré, dans un contexte de prévision, que l'on peut tirer parti de l'association de plusieurs modèles a posteriori en combinant de manière judicieuse leurs sorties. L'utilisation de plusieurs modèles peut cependant rendre certaines applications moins faciles à mettre en œuvre, comme par exemple la régionalisation. Nous pensons que des efforts doivent être consentis en amont, en essayant de combiner plusieurs structures en une seule plus performante. Le modèle résultant de ces combinaisons serait plus fiable et plus performant, même si la distribution des résultats du modèle IDEA à la Figure 5.10 fixe probablement une limite supérieure que l'on ne pourra pas dépasser à partir des concepts et outils mathématiques contenus dans les structures testées.

Nous verrons donc dans le chapitre 7 comment nous pouvons envisager d'exploiter ces complémentarités entre modèles qui se dessinent ici, en essayant de combiner divers composants de ces structures. Dans la partie suivante, nous présentons une deuxième approche des complémentarités.

5.2.7. Recherche d'une typologie bassins-modèles

La classification des bassins effectuée au chapitre 2 nous a permis, à partir de quelques variables descriptives simples du bassin, de scinder notre échantillon en cinq classes. Si nous avons vu les limites de cette classification, qui reste assez dépendante des caractéristiques utilisées, nous voulons essayer de voir dans quelle mesure il est possible d'exploiter cette information pour l'application préférentielle des modèles. Il s'agit de voir si l'on peut mettre en évidence une correspondance entre types de bassins et types de modèles.

Une telle typologie bassins-modèles permettrait à l'utilisateur, connaissant quelques caractéristiques du bassin, de pouvoir choisir le modèle qui semble a priori le plus fiable sur le bassin considéré. Par rapport à la méthode précédente de recherche de complémentarités,

l'échantillon de bassins n'est plus appréhendé dans sa globalité mais par classes de bassins du même type.

A partir de la classification ayant identifié cinq classes de bassins, nous avons calculé les performances moyennes obtenues par les modèles sur chacune des classes. Ces moyennes ont été calculées à partir des résultats moyens sur chaque bassin (et non à partir des résultats par période, ceci permettant de donner un poids équivalent à chaque bassin). La Figure 5.12 donne les résultats en contrôle dans le cas du critère $CR2$ ($Nash(\sqrt{Q})$). On constate que plus le niveau de performance moyen est élevé, moins il y a d'écart entre les performances des différentes structures. Les performances obtenues sur la classe 3 sont par exemple beaucoup plus stables d'une structure à l'autre que celles obtenues sur la classe 4.

Les classes 4 et 5 présentent toutes deux des niveaux de performance moins élevés. Ces deux classes correspondent à des bassins soumis à une ETP relativement élevée et faiblement arrosés, avec des comportements hydrologiques bien contrastés entre une saison humide et une saison sèche assez longue. Ces bassins ont par ailleurs des rendements faibles. Rappelons que l'apparent moins bon niveau de performance provient en partie des problèmes d'applicabilité du critère de Nash sur les bassins intermittents, dont nous avons discuté au chapitre 4.

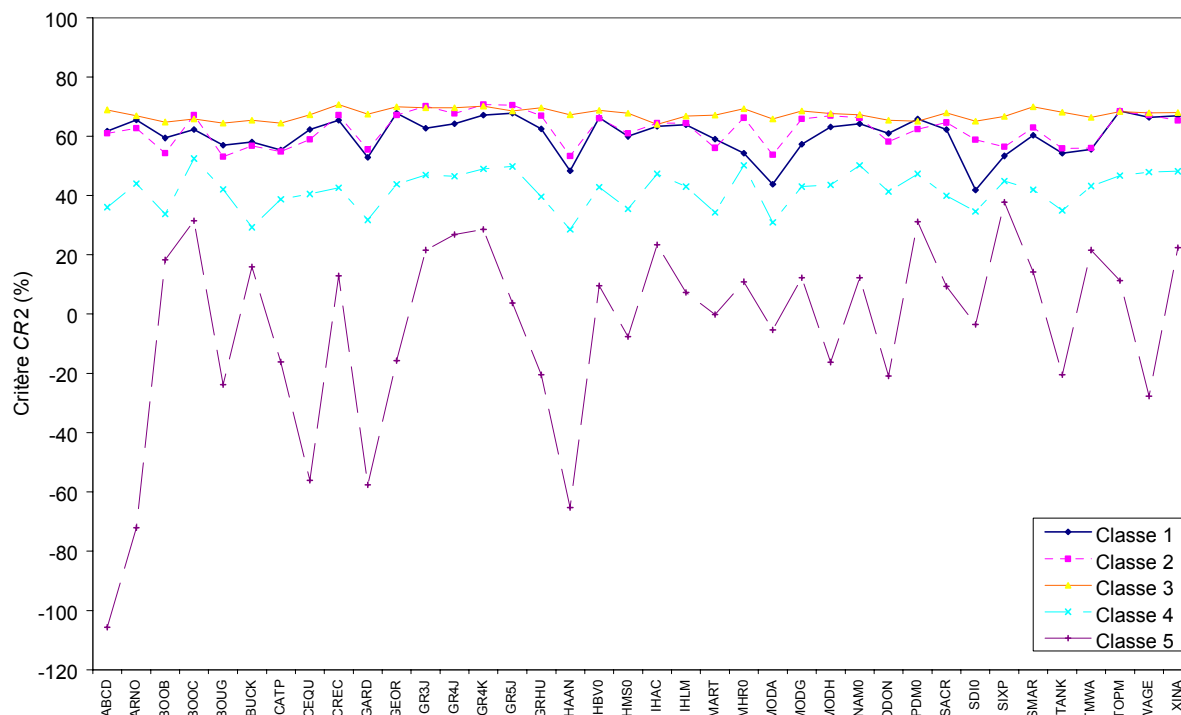


Figure 5.12: Performances moyennes des modèles sur les cinq classes de bassins pour le critère $CR2$

Essayons maintenant de voir si des modèles obtiennent de meilleurs résultats (au niveau du classement des modèles) sur certaines classes de bassins que sur d'autres. Pour cela, il est tout d'abord nécessaire de choisir un critère parmi les six dont nous disposons.

La Figure 5.13 montre, dans le cas du modèle WAGE, qu'il y a généralement une évolution des classements du modèle assez similaire d'un critère de performance à l'autre en fonction des classes de bassin. Ceci se retrouve pour les autres modèles dans une plus ou moins large mesure. De tels graphiques, réalisés pour tous les modèles permettent d'avoir une appréciation quantitative sur le comportement des modèles suivant les classes. Il est difficile

d'identifier des similarités de comportement d'un modèle à l'autre. Cependant, on peut remarquer que:

- les évolutions de classement entre les classes 1 et 2 sont généralement faibles, ce qui est en partie dû au fait que les caractéristiques des bassins sur ces deux classes sont assez proches;
- la classe 3 entraîne parfois de très nettes différences dans les performances relatives des modèles: les modèles BOOC, IHAC, NAM0, PDM0 obtiennent des résultats relatifs bien moins bons sur cette classe que sur les autres classes. Inversement, les modèles ABCD, GARD, GRHU, HAAN et TANK fournissent de bien meilleurs résultats sur cette classe;
- les évolutions de classement sur les classes 4 et 5 sont très variables, ce qui est à rapprocher de la grande variabilité des performances obtenues sur ces classes.

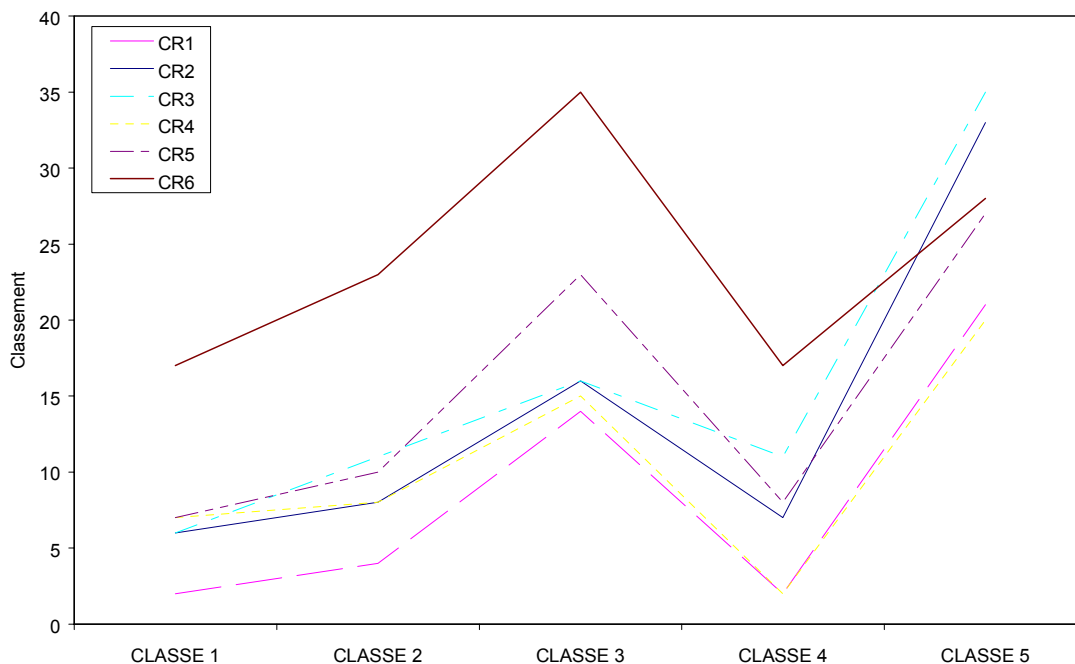


Figure 5.13: Evolution des classements des performances du modèle WAGE sur chacune des classes pour les six critères de performance

Est-il possible de déterminer si certains modèles sont plus adaptés à certaines conditions que d'autres ? Pour tenter de répondre à cette question et en se basant sur les remarques précédentes sur les évolutions comparables des performances relatives des modèles suivant les critères, nous avons utilisé dans un premier temps les six critères de qualité, ceci ayant pour but de ne pas biaiser la réponse en fonction du choix d'un critère spécifique.

Pour cibler l'analyse, nous nous sommes restreints aux modèles apparaissant dans les six premières places des classements des modèles suivant ces six critères, et nous avons compté le nombre de fois (indiqué par N1 dans le Tableau 5.4) où les modèles étaient présents. Par exemple, pour la classe 1, le modèle GEOR est classé parmi les six meilleurs modèles pour 5 fonctions critères. Le Tableau 5.4 regroupe les modèles apparaissant dans les six premières places pour au moins deux critères. Ce sont donc ceux qui paraissent les plus adaptés pour simuler les conditions hydrologiques des différentes classes. Suivant les classes, ce sont des modèles différents qui sont les plus satisfaisants. Néanmoins, certains modèles comme TOPM, GR3J (GR4K), GEOR ou BOOC sont parmi les meilleurs pour plusieurs classes de bassins.

Classe 1		Classe 2		Classe 3		Classe 4		Classe 5	
Modèle	N1	Modèle	N1	Modèle	N1	Modèle	N1	Modèle	N1
TOPM	6	GR3J	6	GR4K	6	BOOC	6	SIXP	6
GEOR	5	GR4K	6	CREC	5	MHR0	6	BOOC	5
GR4K	5	TOPM	5	GR3J	5	GR4K	5	GR4K	5
GR4J	3	GR5J	4	GR4J	5	GR5J	3	GR4J	4
GR5J	3	BOOC	3	GRHU	4	NAM0	3	PDM0	4
WAGE	3	GEOR	3	SMAR	4	GR3J	2	TOPM	3
XINA	3	GR4J	3	GEOR	2	WAGE	2	GR3J	2
CREC	2	GRHU	2			XINA	2	SMAR	2
HBV0	2								

Tableau 5.4 : Liste des modèles apparaissant pour N1 critères de performances parmi les six meilleurs modèles (par classes de bassins)

Modèles ou associations	Moyenne CR2 (%)
GR4K	63,7
GR4K-GEOR	66,4
TOPM-GR4K-CREC-BOOC-SIXP	65,2
IDEA	71,4

Tableau 5.5: Performances moyennes (critère CR2) des différentes associations de modèles sur les 429 bassins (calculées à partir des performances moyennes par bassin)

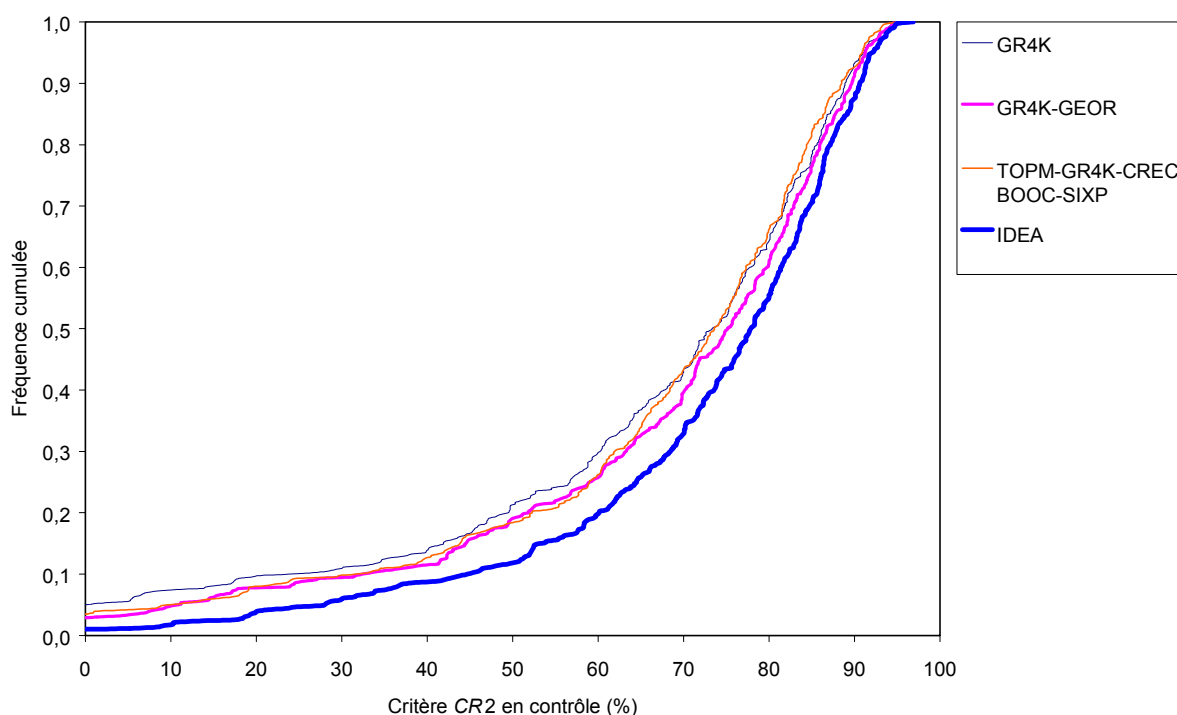


Figure 5.14: Distributions des performances suivant le critère CR2 de diverses associations de modèles sur l'échantillon des 429 bassins

Pour comparer les résultats de cette approche de complémentarité entre structures avec ceux exposés dans la partie précédente, nous avons calculé les performances obtenues dans chaque cas, suivant le critère CR2. Pour la première approche de complémentarités, nous avons retenu la paire de modèles qui paraissait la plus satisfaisante, à savoir GR4K-GEOR. Dans la deuxième approche, pour chacune des cinq classes de bassins, nous avons retenu le modèle du

Tableau 5.4 qui donne les meilleurs résultats suivant *CR2*. Les cinq modèles retenus sont respectivement TOPM, GR4K, CREC, BOOC et SIXP. Le Tableau 5.5 montre que l'association de deux structures (ici GR4K et GEOR) permet d'obtenir de meilleures performances que l'association de cinq structures en fonction du type de bassins. Le gain par rapport au meilleur modèle pris individuellement (GR4K est la structure la plus satisfaisante suivant le critère *CR2*) est plus important dans le premier cas. Ceci est également illustré à la Figure 5.14, où apparaissent les distributions des résultats, avec une légère supériorité de la première approche de complémentarité sur la deuxième. On constate que l'écart avec le modèle 'idéal', quelle que soit l'approche de complémentarité adoptée, reste important.

Les moins bons résultats obtenus en tenant compte des correspondances entre modèles et types de bassins peuvent s'expliquer par la difficulté que nous avons mentionnée au chapitre 2 d'obtenir une classification de bassins réellement satisfaisante, cette dernière étant relativement instable suivant les variables explicatives utilisées. De ce fait, il devient difficile d'assigner avec certitude tel ou tel bassin à l'une des classes.

Cette démarche de recherche de typologie bassins-modèles aurait également pu être abordée au travers d'une analyse en composantes principales (ACP), en mettant en relation bassins et modèles par le biais des performances. Une telle approche a été adoptée en prenant les résultats moyens par bassin au contrôle suivant le critère *CR2*. L'analyse a été effectuée à l'aide du logiciel STATlab, en considérant les bassins comme individus et les modèles comme variables explicatives (le logiciel limitant le nombre de variables explicatives à 50, il n'était pas possible de choisir les bassins comme variables explicatives). Pour éviter de biaiser l'ACP avec des valeurs fortement négatives du critère *CR2*, nous lui avons appliqué une transformation préalable ($X = CR2 / (2 - CR2)$) qui permet de limiter l'intervalle de variation de la variable transformée *X* à]-1; 1]. Sans cette transformation préalable, on obtient quand même des résultats assez similaires.

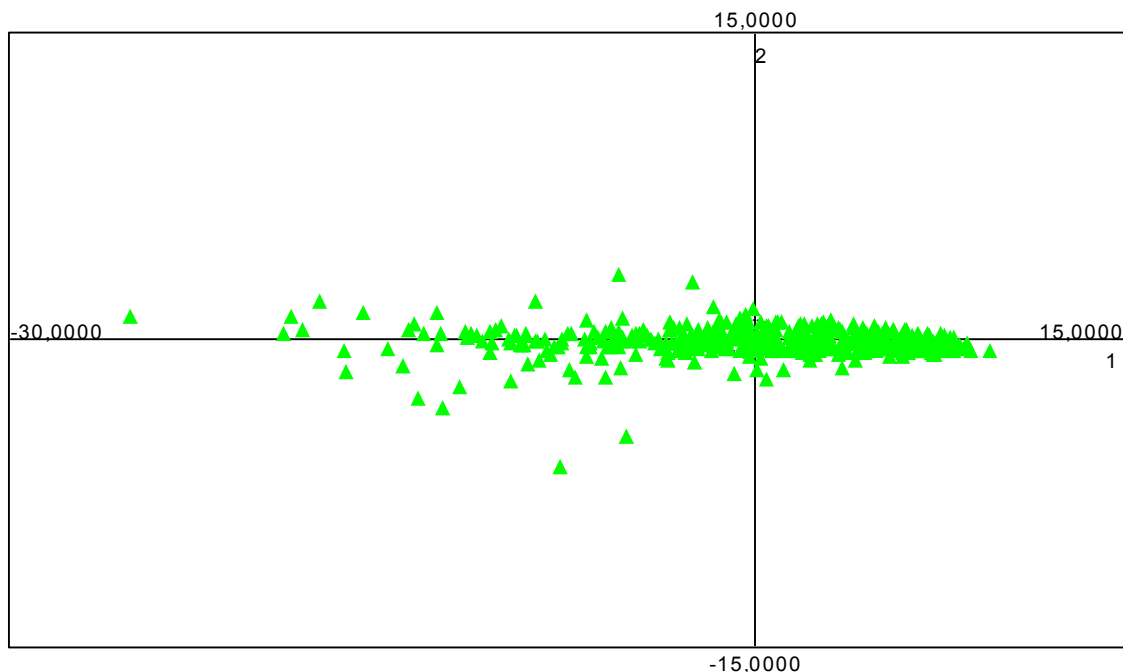


Figure 5.15: Projection des 429 bassins sur les premiers axes principaux de l'ACP bassins-modèles

Les résultats montrent que la première composante principale explique 90 % de la variance de l'échantillon de bassin, les autres composantes ayant des taux d'explication non significatifs

(inférieurs à 1,5 %) et étant donc peu interprétables. Cette première composante peut être interprétée comme la facilité de modélisation du bassin. En effet, les bassins s'ordonnent suivant cet axe à peu près par performance moyenne (sur l'ensemble des modèles) croissante. La projection sur le plan formé par les deux premiers axes principaux confirme l'allure très allongée du nuage des bassins suivant la direction du premier axe (voir Figure 5.15). Cette ACP indique que les performances des modèles sur ces bassins sont peu discriminantes pour discerner des correspondances bassins-modèles, autrement dit qu'il n'est pas possible par ce biais d'identifier des groupes de modèles attachés plus spécifiquement à certains sous-échantillons de bassins. Si tel avait été le cas, il aurait été possible d'identifier une seconde composante principale soulignant l'existence d'une telle correspondance.

Les tests effectués ici ne semblent pas permettre pas de mettre en évidence une spécificité des structures vis-à-vis des types bassins testés (climat semi-aride ou humide par exemple).

La deuxième approche de complémentarité (à partir de la typologie bassins-modèles) présente un intérêt potentiel certain, notamment par le lien qu'elle fait avec les caractéristiques des bassins. Cependant, d'après les résultats de notre étude, son efficacité se trouve limitée d'une part par la difficulté d'obtenir une typologie de bassins et une répartition en classes qui puissent être exploitées avec fiabilité, et d'autre part par le fait que notre étude met en doute l'existence d'une correspondance entre structures de modèles et types de bassins. L'utilisation de cinq modèles différents ne paraît pas justifiée au regard des résultats, puisque l'association de deux modèles seulement conduit à une meilleure amélioration des performances.

5.3. Conclusion

Le cadre de comparaison que nous avons élaboré nous a permis d'évaluer les performances de structures de modèles sur un échantillon de 429 bassins. Ce cadre comparatif minimise certaines faiblesses que nous avons pu signaler dans des études comparatives antérieures, à savoir le trop petit nombre de bassins ou la non homogénéité des procédures de test. Même si certaines sources d'incertitude subsistent, les précautions qui ont été prises ont permis, selon nous, d'atteindre un bon niveau d'objectivité dans l'évaluation des capacités des structures à simuler la transformation pluie-débit sur un bassin versant.

Nous avons pu montrer que les 38 structures testées obtiennent des résultats grossièrement équivalents, en comparaison à un modèle plus simple ne comportant pas de procédures de suivi d'humidité et de routage propres. Cependant, parmi ces 38 structures, certaines réussissent en moyenne à obtenir des performances plus élevées et paraissent plus fiables et plus robustes que d'autres.

En mettant en relation cette qualité de robustesse avec le nombre de paramètres optimisés, il a été démontré que des modèles simples (à peu de paramètres) sont capables d'atteindre des niveaux de performance moyens aussi satisfaisants que des modèles plus complexes, c'est-à-dire ayant un nombre plus élevé de paramètres. Ces modèles plus complexes sont sujets à des problèmes de sur-paramétrisation qui les empêchent d'atteindre leur valeur potentielle. Le nombre de paramètres optimisables d'un modèle pourrait être compris entre 3 et 5 seulement, ce qui rejoint les observations faites par Beven (1989) qui préconise cette même fourchette de paramètres libres. Jakeman et Hornberger (1993) ont également déduit de leur travaux que six paramètres sont suffisants pour reproduire la transformation pluie-débit sur un grand nombre de bassins. Parmi ces modèles simples, les modèles GR ont paru capables d'obtenir des performances d'un niveau aussi satisfaisant que des structures plus complexes.

Steeffel et Van Cappellen (1998) commentent l'article d'Oreskes *et al.* (1994) en disant que, pour une grande part, le test ultime d'un modèle particulier est sa simplicité relativement à ses performances sur un nombre donné d'observations. En accord avec ce point de vue, nous pouvons affirmer que, à performances équivalentes, les modèles les plus simples devraient être préférés. En effet, ils induisent moins de problèmes d'incertitude de paramètres, comme nous allons le voir dans le chapitre 6. Ceci est de la plus grande importance, par exemple en régionalisation où la fiabilité des paramètres est un point essentiel (Wheater *et al.*, 1993). Nous discutons cet aspect plus en détail dans le dernier chapitre.

Steeffel et Van Cappellen (1998) ajoutent que cependant, il y a une limite à la simplicité des modèles, cette limite étant atteinte lorsqu'un modèle échoue à donner des explications satisfaisantes aux observations (c'est-à-dire que la simplicité seule ne peut être utilisée comme un critère valide de qualité d'un modèle). Nous partageons cette opinion dans la mesure où la parcimonie ne peut être recherchée que si elle ne diminue pas les capacités du modèle à simuler les débits observés.

Cette étude apporte aussi un éclairage sur la question de la complémentarité entre modèles. Dans les tests réalisés, aucun des modèles n'a montré sa supériorité dans tous les cas. Au contraire, l'association de structures complémentaires a permis d'améliorer de façon significative les résultats des modèles pris indépendamment. Cette complémentarité a été ici envisagée en terme de performances sur les bassins. Une autre approche pour établir des complémentarités, consistant à construire des correspondances entre structures de modèles et types de bassins (tels qu'ils ont été établis au chapitre 2), n'a pas permis de mettre en évidence des spécificités nettes des modèles (l'efficacité de cette seconde approche ayant peut-être été limitée par l'absence de descripteurs physiographiques pour la classification des bassins).

Tout en conservant un faible nombre de paramètres, l'existence d'une complémentarité de performances ouvre la perspective de grouper les composants efficaces provenant de différents modèles dans un seul, lequel aurait une structure et une paramétrisation bien choisie. C'est la voie que nous nous proposons d'explorer au chapitre 7. Nous partirons de la structure la plus simple, celle du modèle GR3J. Elle a en effet prouvé, comme les autres structures de la famille GR, de bonnes qualités de performance et de robustesse, et présente l'avantage d'être suffisamment parcimonieuse pour que l'on puisse essayer de trouver la meilleure façon d'intégrer un ou deux degrés de complexité supplémentaires. La structure GR4K, identique à celle de GR3J dans laquelle on optimise un quatrième paramètre initialement fixé, représente une possibilité de complexification. D'autres sont peut-être plus judicieuses.

Nous adopterons donc l'approche conseillée par Nash et Sutcliffe (1970), en accroissant la complexité graduellement et le plus judicieusement possible pour améliorer les performances du modèle. Cette approche rejoint également le point de vue de Bergström (1991) qui pense que "l'on évite ainsi la frustration d'abandonner des concepts ou théories élégants en apparence lorsque l'on doit passer de structures de modèles complexes à de plus simples".

Chapitre 6

Chapitre 6

Exploration du modèle GR3J au travers d'une comparaison de deux méthodes d'analyse d'incertitude des paramètres

6.1. Introduction

Les paramètres de modèles hydrologiques globaux (conceptuels ou empiriques) n'ont généralement pas de correspondance directe sur le terrain et leurs valeurs sont donc déterminées par optimisation mathématique. Du fait des nombreuses erreurs qui entachent toute la démarche de modélisation (de la conception du modèle jusqu'à son application), les paramètres des modèles ainsi calculés sont eux-mêmes sujets à de nombreuses incertitudes. Or l'incertitude sur la sortie d'un modèle (et sur les applications qui les utilisent) est en partie fonction des incertitudes sur les paramètres du modèle. Ainsi, l'incertitude ΔY sur la sortie Y d'un modèle dépendant d'un paramètre A peut être donnée par:

$$\Delta Y = \Delta A \cdot \frac{\partial Y}{\partial A} \quad \text{Eq. (6.1)}$$

où ΔA est l'incertitude sur le paramètre A et $\frac{\partial Y}{\partial A}$ la sensibilité de Y par rapport à A .

Il paraît donc naturel, pour étudier l'incertitude sur les sorties, de s'intéresser à l'incertitude et à la sensibilité sur les paramètres.

La sensibilité peut être définie comme le taux de changement d'un facteur en réponse au changement d'un autre facteur, c'est-à-dire la réponse d'un système à une perturbation. L'analyse de sensibilité permet d'indiquer dans quelle mesure les résultats du modèle sont sensibles au changement des valeurs des paramètres: si un changement de la valeur d'un paramètre conduit à un faible changement dans la valeur des sorties, le modèle est alors peu sensible aux variations du paramètre. Inversement un grand changement dans les sorties du modèle consécutivement au changement de la valeur du paramètre indique une bonne sensibilité des sorties du modèle. Si une bonne sensibilité est souhaitable, elle est cependant très dépendante d'éventuelles anamorphoses (transformations mathématiques) sur les paramètres: des transformations bien choisies permettent d'accroître la sensibilité des sorties du modèle aux variations du paramètre considéré. La sensibilité d'un paramètre n'est donc pas absolue, mais doit être considérée relativement à d'autres mesures de sensibilité.

Nous allons nous intéresser à la détermination de l'incertitude sur les paramètres. L'objectif de ce chapitre est double. Nous voulons tout d'abord comparer les résultats d'une méthode classique d'estimation des incertitudes sur les paramètres, l'analyse d'incertitudes par approximation linéaire présentée par Mein et Brown (1978) ou Troutman (1985a, 1985b),

avec une méthode non conventionnelle appelée ici méthode ‘multi-calage’. Cette dernière est basée sur l’analyse statistique d’un échantillon de valeurs de paramètres obtenues par calages multiples sur différentes périodes d’enregistrement. Ceci suppose que l’on dispose sur le(s) bassin(s) étudié(s) de chroniques de données suffisamment longues pour permettre de distinguer un nombre suffisant de périodes. Ces analyses d’incertitudes permettent également d’étudier les liens entre les paramètres du modèle.

Le second objectif est d’acquérir des compléments d’information sur l’articulation structurelle du modèle GR, par le truchement des liens entre les paramètres, et de voir si cette analyse peut mettre en évidence d’éventuelles faiblesses de l’architecture du modèle. Dans la perspective de nos travaux de modifications de structures, nous avons utilisé ici le modèle GR3J d’Edijatno *et al.* (1999), avec quatre versions comprenant de trois à cinq paramètres, certains paramètres fixes de la version à trois paramètres étant optimisés.

Après une rapide synthèse sur les notions d’erreur et d’incertitude dans le contexte de la modélisation hydrologique, nous présentons les deux méthodes d’analyse d’incertitudes que nous souhaitons appliquer (méthode par approximation linéaire et méthode ‘multi-calage’). La méthodologie adoptée dans cette étude est décrite et les performances des quatre structures du modèle sont commentées. Ensuite, au travers d’une présentation comparative des résultats des deux méthodes testées, nous analysons les écarts-types des paramètres et les corrélations entre eux. Nous nous interrogeons également sur l’influence de la complexité du modèle sur l’incertitude sur les paramètres et sur les possibles effets de la longueur de la série de calage sur les résultats des analyses d’incertitudes.

6.2. Erreurs et incertitudes en modélisation pluie-débit

Les bassins versants sont des systèmes naturels complexes, difficiles à étudier et à comprendre. La modélisation de la transformation pluie-débit est sujette à diverses sources d’erreurs, causes ou conséquences de cette difficulté d’appréhension. Plusieurs auteurs, tels que Kitanidis et Bras (1980) ou O’Donnell et Canedo (1980) donnent une liste des origines multiples des erreurs, que l’on retrouve principalement dans:

- les données d’entrée (collecte des données, traitement des données, établissement de moyennes spatiales ou temporelles, données non directement observables telles que l’ETP),
- la structure du modèle: quelle que soit l’approche utilisée, le modèle reste toujours une représentation grossière d’un système naturel complexe,
- les paramètres du modèle, qui dépendent du choix de la fonction objectif, des performances de la méthode d’optimisation, de l’échantillon de données de calage et de l’identifiabilité d’un optimum.

Ces sources d’erreurs génèrent des incertitudes à chaque étape du processus de modélisation, qui se propagent tout au long de la procédure d’utilisation du modèle (du calage jusqu’aux applications). Leurs effets sont largement commentés dans la littérature (voir par exemple Sorooshian, 1985; Troutman, 1985a; Yapo *et al.*, 1996). En phase de calage, les degrés de liberté laissés au modèle par l’intermédiaire des paramètres optimisables permettent implicitement de compenser (de façon partielle) les erreurs dans les données ou celles dues à l’imperfection du modèle.

En théorie, plus le nombre de paramètres optimisés augmente, plus il sera aisé pour le modèle d’arriver à fournir des simulations proches des débits observés. Cependant cet avantage

théorique de l'ajout de paramètres optimisables se trouve limité par la troisième source d'erreurs, due à l'incertitude sur la valeur des paramètres. En effet, la phase de détermination des paramètres consiste à extraire une information sur le comportement du bassin et à la transférer au modèle par l'intermédiaire des paramètres, dont les valeurs optimisées constituent alors en quelque sorte l'identificateur hydrologique, l'empreinte digitale du bassin considéré.

Pour être fiable, ce transfert d'information suppose tout d'abord que la structure du modèle soit suffisamment en adéquation avec la réalité du bassin. En effet, plus la structure est 'juste' (c'est-à-dire capable de représenter la transformation pluie-débit), moins les paramètres doivent compenser les erreurs structurelles du modèle. Ensuite, il faut que le degré de complexité autorisé dans la structure soit en adéquation avec le degré de 'justesse' du modèle. Mieux la structure du modèle est déterminée, plus il semble possible de s'accorder des degrés de complexité supplémentaires sans accroître l'incertitude sur les paramètres. Enfin, il faut que le niveau d'information disponible sur le bassin (typiquement des chroniques hydro-météorologiques) soit en adéquation avec le niveau d'information requis pour bien déterminer les paramètres. Il apparaît donc que l'inadéquation globale 'réalité du bassin – structure du modèle – niveau de complexité – niveau d'information disponible' puisse être un facteur limitant la détermination des paramètres avec une faible incertitude.

Par ailleurs, la procédure de détermination des paramètres optimaux est elle-même source d'incertitudes, pour des raisons mathématiques (interactions entre paramètres, non-existence d'un unique optimum) ou du fait du choix de la fonction objectif et des performances de l'algorithme de calage. Ces aspects contribuent également à limiter la fiabilité de l'optimum identifié des paramètres. Ceux-ci apparaissent donc comme de bons intégrateurs des erreurs de la chaîne de modélisation. A ce titre, l'étude de l'incertitude sur leurs valeurs semble être un bon moyen d'évaluer la fiabilité du modèle. Elle peut également mettre en lumière des problèmes structureaux du modèle.

Puisque certaines erreurs affectant les simulations du modèle peuvent être considérées comme des variables aléatoires, il est possible de donner une formulation générale du modèle sous un cadre statistique:

$$Y_i = P_i(X, \beta) + e_i \quad 1 \leq i \leq n \quad \text{Eq. (6.2)}$$

où, au pas de temps i , Y_i est la variable observée (débit), $P_i(X, \beta)$ est la variable estimée dépendant d'une(de) variable(s) d'entrée X et du vecteur des paramètres β , et e_i est l'erreur associée. L'Eq. (6.2) est la plus fréquemment adoptée dans la littérature. Cependant, le découplage additif de P_i et e_i ne va pas de soi. De manière plus générale, on peut écrire:

$$Y_i = f(P_i(X, \beta), e_i) \quad 1 \leq i \leq n \quad \text{Eq. (6.3)}$$

Souvent des hypothèses sont faites sur les erreurs: elles sont supposées suivre une distribution normale, avec une moyenne nulle et une variance finie σ^2 , être non corrélées et homoscédastiques. Cependant, dans le cas de la modélisation pluie-débit, ces hypothèses sont rarement vérifiées. Les méthodes d'analyse d'incertitude en hydrologie peuvent être classées, grosso modo, en deux grandes catégories: déterministes ou probabilistes. Melching (1995) présente quelques-unes de ces méthodes. La première catégorie s'intéresse généralement à déterminer l'incertitude à partir d'un point de l'espace des paramètres (l'optimum par exemple). Elle repose sur une approximation au premier ordre, c'est-à-dire une linéarisation locale du système et ses résultats sont donc conditionnés par la validité de cette approximation. La démarche de Mein et Brown (1978) et différentes versions de la méthode des moments (Melching, 1992) entrent dans cette catégorie d'approches.

En réponse aux faiblesses parfois avérées des approximations des méthodes précédentes, une autre catégorie d'approches a été développée. Elles reposent essentiellement sur des procédures de type Monte-Carlo, qui utilisent l'idée d'associer aux variables ou paramètres incertains des distributions de probabilité. On génère alors des jeux de paramètres en fonction de ces distributions, et à chaque fois, on fait tourner le modèle. En fonction d'un critère de satisfaction donné (sur les sorties par exemple), il est alors possible de sélectionner les jeux de paramètres satisfaisants et de construire les distributions de probabilité expérimentales correspondantes, avec parallèlement l'estimation d'intervalles de confiance sur les sorties. Dans cette catégorie, on trouve notamment la méthode Generalized Likelihood Uncertainty Estimation (GLUE) de Beven et Binley (1992), celle proposée par Van Straten et Keesman (1991) ou l'algorithme Metropolis mis en œuvre par Kuczera et Parent (1998).

Dans les études comparatives de certaines de ces approches (voir par exemple Scavia *et al.*, 1981; Yeh et Tung, 1993), les résultats des analyses par des procédures de type Monte-Carlo sont généralement prises comme référence. Si les résultats des premières méthodes remettent parfois en cause les hypothèses faites, elles paraissent donner des résultats satisfaisants dans un certain nombre de cas. Elles sont par ailleurs nettement moins exigeantes au niveau calcul que les méthodes de type Monte-Carlo.

Dans la suite de notre travail, nous nous limiterons à l'utilisation d'une méthode fondée sur l'approximation linéaire et d'une méthode qui pourrait être rapprochée d'une approche Monte-Carlo mais portant sur un nombre (très) limité de tirages.

6.3. Méthodes pour l'analyse d'incertitudes sur les paramètres

Dans les paragraphes suivants, nous donnons une brève description des deux approches d'évaluation des incertitudes qui vont être comparées.

6.3.1. Analyse d'incertitudes par approximation linéaire

L'analyse d'incertitudes est considérée par de nombreux auteurs comme faisant partie intégrante de toute étude hydrologique et elle est potentiellement utile dans les phases de formulation, d'optimisation ou de vérification du modèle (Vemuri *et al.*, 1969; McCuen, 1973). Sorooshian et Arfi (1982) mentionnent que de telles procédures peuvent donner des informations utiles sur la nature des interactions entre les paramètres, pour établir des mesures de confiance sur les paramètres ou pour détecter une faible identifiabilité des paramètres.

Le cadre mathématique de l'analyse d'incertitudes par approximation linéaire (que nous désignerons dans la suite AIAL par commodité), parfois abusivement désignée par 'analyse de sensibilité', a été détaillé par Mein et Brown (1978) puis repris par Troutman (1985a), et est utilisé ici pour calculer les écarts-types des paramètres et la matrice des corrélations entre paramètres. Les détails des calculs sont présentés dans les références précédentes et par Nascimento (1995) qui a appliqué cette approche au modèle GR4J. Nous n'en présentons que les grandes lignes en Annexe 4. La méthode repose sur le principe d'approximation des erreurs du modèle par une formulation linéaire par rapport aux paramètres au voisinage de l'optimum, en utilisant une approximation du modèle au premier ordre du développement en série de Taylor de l'Eq. (6.2). Pour évaluer ce développement, on calcule l'effet d'une petite variation de la valeur optimale de chaque paramètre sur la chronique des débits. A chaque fois, on peut ainsi évaluer un résidu, différence entre débits calculés avec l'optimum du paramètre et débits calculés avec la valeur modifiée de l'optimum. Enfin on relie ces résidus de façon linéaire aux erreurs du modèle, par une équation du type:

$$Q - Q_0 = \sum_{i=1}^p \alpha_i \frac{Q_i - Q_0}{\varepsilon} + \mu \quad \text{Eq. (6.4)}$$

où Q est le vecteur des débits observés, Q_0 le vecteur des débits calculés par le modèle avec le jeu optimum des paramètres, Q_i le vecteur des débits calculés par le modèle avec le jeu optimum des paramètres dont la composante i a été modifiée d'une petite quantité ε et $\alpha_1, \dots, \alpha_p$ sont p coefficients à déterminer, avec p nombre de paramètres du modèle. On détermine les valeurs des coefficients $\alpha_1, \dots, \alpha_p$ qui minimisent l'erreur μ .

La matrice des variances-covariances de $\alpha_1, \dots, \alpha_p$ correspond approximativement à celle des paramètres du modèle. On peut alors obtenir une évaluation des écarts-types et de la matrice de corrélation des paramètres du modèle (Kabouya et Michel, 1991).

Cette méthode, qui s'appuie sur le modèle linéaire général, avance des hypothèses de normalité et d'indépendance des résidus du modèle qui sont rarement vérifiées. Nascimento (1995) mentionne que des transformations appropriées permettent de rendre les résidus plus conformes à ces hypothèses. Les problèmes d'autocorrélation peuvent également être limités en ne considérant pas tous les résidus de la série temporelle mais seulement un tous les n pas de temps ($n = 5$ par exemple). Cette solution, employée par Nascimento (1995) pose cependant le problème de la représentativité des résidus considérés: on peut manquer par ce procédé de nombreux épisodes où les résidus sont importants. Tous les pas de temps ont donc été pris en compte ici.

Nous pouvons mentionner que des approches alternatives ont été proposées par d'autres auteurs pour évaluer les corrélations entre paramètres, avec des évaluations graphiques de la surface de réponse aux environs de l'optimum (Gupta et Sorooshian, 1983) ou le calcul de deux autres mesures de sensibilité, la concentricité χ et l'interaction τ (Sorooshian et Arfi, 1982). L'approche du maximum de vraisemblance, elle, introduit l'utilisation des dérivées seconde de la fonction objectif par rapport aux paramètres. Nous nous en tiendrons cependant ici à l'approche décrite précédemment.

6.3.2. La procédure 'multi-calages'

Puisque les paramètres sont calés en utilisant des données, il existe un lien intrinsèque entre les caractéristiques des données de calage et les valeurs optimisées des paramètres. Ces valeurs dépendent de l'information contenue dans ces jeux de données (et également des erreurs qui les affectent). Le calage d'un modèle sur différents jeux de données d'un bassin est susceptible, par conséquent, de conduire à des optima différents des paramètres, ceci dépendant par exemple de l'aridité ou de l'humidité de la période, de l'hétérogénéité de la pluie ou de la variabilité des événements qui la composent.

Etant donné cette dépendance des paramètres par rapport aux données, une mesure de l'incertitude sur les paramètres pourrait être fournie par calage du modèle sur différentes périodes d'enregistrement, en étudiant les propriétés statistiques de l'échantillon de paramètres obtenus. C'est l'approche que nous proposons dans la méthode 'multi-calage'. Elle se décompose en plusieurs phases:

- 1- fractionnement de la période totale d'enregistrement en n sous-périodes indépendantes;
- 2- calage du modèle sur chacune des sous-périodes, fournissant au total un échantillon de n vecteurs de paramètres;
- 3- calcul de la moyenne et de l'écart-type de chaque paramètre et détermination des corrélations entre paramètres sur l'échantillon des n valeurs.

L'échantillon des n vecteurs de paramètres est considéré comme une population statistique. Le fait que la période d'enregistrement disponible soit scindée en sous-périodes sans recouvrement garantit une meilleure indépendance des paramètres en passant d'une période à une autre, bien que cela réduise en contrepartie le nombre de sous-périodes identifiées et donc le cardinal de l'échantillon des paramètres optimisés. L'approche 'multi-calage' nécessite donc d'avoir de longues chroniques de données sur les bassins étudiés.

L'incertitude sur les paramètres est quantifiée de deux manières différentes, par l'intermédiaire d'une analyse mathématique des résidus du modèle dans la première approche et par la quantification de la variabilité des paramètres dans l'approche 'multi-calage'. La première méthode accorde plus d'importance à l'incertitude due aux caractéristiques structurelles du modèle, alors que la seconde accorde plus d'importance au rôle des données dans l'incertitude sur les paramètres.

6.4. Le modèle GR3J

Pour tester et comparer les deux approches d'analyse d'incertitudes, nous avons choisi un modèle pluie-débit simple. Comme l'ont souligné Dawdy et O'Donnell (1965), le nombre croissant de paramètres dans un modèle diminue la sensibilité de chacun d'eux et peut introduire lors de la phase d'optimisation des problèmes de convergence. Le modèle GR3J à trois paramètres (Edijatno *et al.*, 1999) a été utilisé. Une représentation schématique en est donnée à la Figure 6.1. Les notations adoptées ici pour les trois paramètres sont:

- X1: le coefficient d'échange, pour la simulation des apports ou pertes en eau
- X2: la capacité à un jour du réservoir de routage,
- X3: le temps de base de l'hydrogramme unitaire HU1.

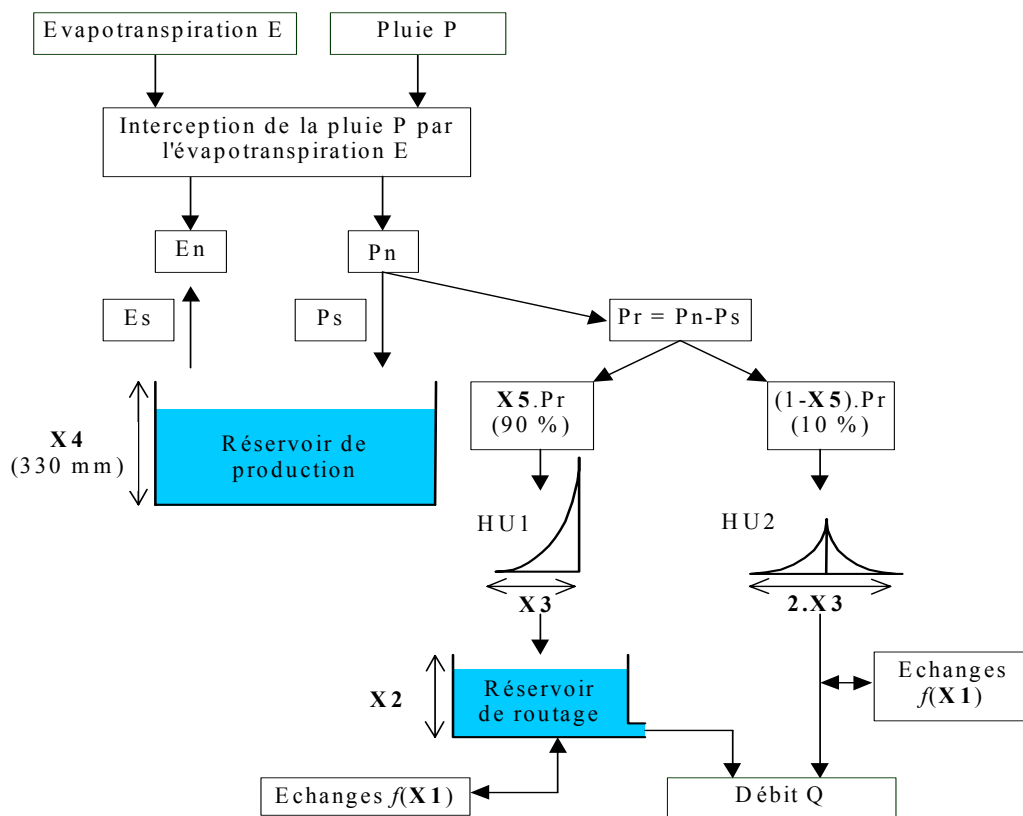


Figure 6.1: Structure du modèle GR utilisé pour tester les méthodes d'analyse d'incertitudes

Remarquons que cette structure empirique n'a pas de vocation de représentation conceptuelle du bassin versant. Par exemple, il n'y a pas de correspondance a priori entre la nappe et le réservoir de routage (souvent identifié dans les modèles conceptuels comme un réservoir 'eau souterraine'). Du fait du terme d'échange, le bassin est considéré comme un système ouvert dans le modèle GR, pouvant avoir des échanges avec des nappes 'profondes'.

Il y a par ailleurs deux paramètres qui sont fixés à des valeurs constantes dans la structure de GR3J:

- X4: la capacité du réservoir de production, fixée à 330 mm,
- X5: le coefficient de partage des deux branches d'écoulement, fixé à 90 %.

Ces valeurs ont été choisies empiriquement comme permettant d'obtenir des résultats convenables sur de nombreux bassins versants. Pour les besoins de notre étude, nous avons donné la possibilité à ces paramètres d'être optimisés. Nous aurons donc ici des structures du modèle GR de trois à cinq paramètres. Quatre structures ont été testées:

- GR31: structure avec X1, X2 et X3 optimisés et X4 et X5 fixés respectivement à 300 mm et 90 % (ici la valeur de 300 mm semblait mieux convenir aux données sélectionnées que les 330 mm originaux);
- GR41: structure avec X1, X2, X3 et X4 optimisés et X5 fixé à 90 %;
- GR42: structure avec X1, X2, X3 et X5 optimisés et X4 fixé à 300 mm;
- GR51: structure avec X1, X2, X3, X4 et X5 optimisés.

Ainsi l'évolution de la variance des paramètres et de leurs corrélations en fonction de la complexité du modèle (nombre de paramètres libres) pourra être étudiée.

Les valeurs transformées des paramètres telles qu'elles sont prises en compte par l'algorithme de calage seront utilisées dans ce chapitre pour illustrer les résultats. En désignant par X_i et x_i respectivement les valeurs transformées et réelles des paramètres, on a les correspondances suivantes:

$X1 = \operatorname{argsinh}(x1)$	$-9,99 < X1 < 9,99$
$X2 = \ln(x2)$	$-1,99 < X2 < 9,99$
$X3 = \ln(x3-0.5)$	$-4,99 < X3 < 9,99$
$X4 = \ln(x4)$	$-1,99 < X4 < 9,99$
$X5 = 19,98.x5-9,99$	$-9,99 < X5 < 9,99$

6.5. Bassins, données et périodes

6.5.1. Choix d'un échantillon de bassins

Nous disposons dans notre échantillon de données d'une série de bassins avec des chroniques suffisamment longues pour pouvoir appliquer la méthode 'multi-calage'. Nous avons retenu 17 bassins versants de la base de données MOPEX (MOdel Parameter Estimation EXperiment), sur lesquels le modèle GR donne des résultats acceptables. Le choix de prendre plusieurs bassins plutôt qu'un seul permet d'obtenir des résultats moins dépendants des caractéristiques d'un seul bassin. Ces bassins sont détaillés dans le Tableau 6.1. Ils sont situés dans le centre des Etats-Unis, dans les Etats de l'Arkansas, l'Oklahoma, du Kansas et du Missouri. Les surfaces des bassins vont de 966 à 6895 km², avec une pluie annuelle moyenne entre 650 et 1300 mm et une ETP entre 950 et 1560 mm. Le coefficient d'écoulement moyen varie entre 8 et 40 %, comme l'illustre la Figure 6.2. Cette figure indique par ailleurs qu'il

existe une relation linéaire assez nette entre pluie et ETP annuelles moyennes pour les bassins étudiés. Les données d'ETP sont des moyennes interannuelles de chaque jour calendaire. Sur chaque bassin, le modèle utilisera donc une même chronique de données d'ETP pour toutes les années.

N°	Code	Nom de la station	Surface (km²)
1	US144200	Arkansas River à Valley Center, KS	3437
2	US152000	Chikaskia River près de Blackwell, OK	4815
3	US153000	Black Beer Creek à Pawnee, OK	1492
4	US177500	Bird Creek près de Sperry, OK	2344
5	US186000	Spring River près de Waco, MO	3015
6	US187000	Shoal Creek à l'amont de Joplin, MO	1106
7	US189000	Elk River près de Tiff City, MO	2259
8	US191000	Big Cabin Creek près de Big Cabin, OK	1166
9	US196500	Illinois River près de Tahlequah, OK	2484
10	US243500	Deep Fork près de Beggs, OK	5227
11	US250000	Lee Creek près de Van Buren, ARK	1103
12	US252000	Mulberry River près de Mulberry, ARK	966
13	US261500	Fourche Lafave près de Gravelly, ARK	1062
14	US311500	Deep Red Run près de Randlett, OK	1598
15	US332500	Blue River près de Blue, OK	1233
16	US335000	Clear Boggy près de Caney, OK	1865
17	US340000	Little River près de Horatio, ARK	6895

Tableau 6.1: Liste des 17 bassins américains étudiés (ARK: Arkansas, KS: Kansas, MO: Missouri, OK: Oklahoma)

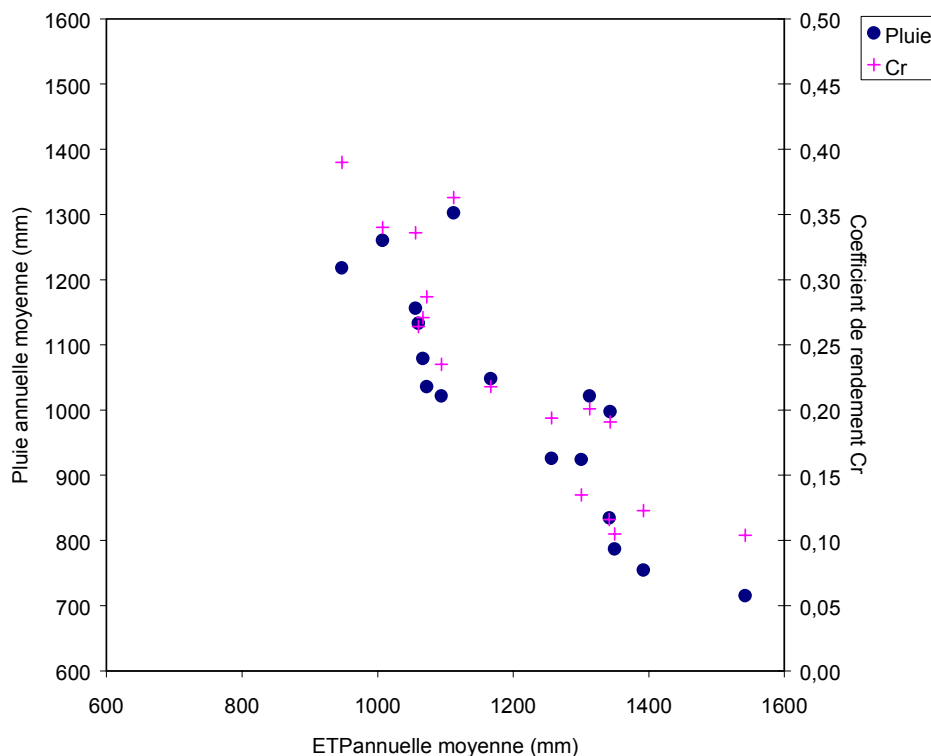


Figure 6.2: ETP moyenne annuelle et coefficient d'écoulement moyen en fonction de la pluie moyenne annuelle sur les 17 bassins versants étudiés

L'avantage de choisir ces bassins réside d'une part dans l'uniformité du type et de l'origine des données, et d'autre part dans le fait qu'une même période d'enregistrement allant de 1951

à 1987 était disponible pour tous ces bassins, rendant le traitement systématique plus aisé. D'après les informations dont nous disposons sur ces bassins, ils ne semblent pas avoir connu de changement majeur au cours de cette période. Bien que le bassin US340000 ait été jugé, au chapitre 2, peu propice à la modélisation d'après les informations disponibles (voir Annexe 3), la régularisation des débits signalée semble en fait avoir un faible effet, permettant au modèle d'obtenir de bons résultats. Pour étudier l'influence de la longueur de la série de calage sur les résultats des deux méthodes d'analyses d'incertitudes, nous avons scindé successivement la période totale en 36, 18, 12, 9, 6 et 4 sous-périodes de 1, 2, 3, 4, 6 et 9 ans respectivement. A chaque fois, une année précédant la période d'étude a été réservée pour l'initialisation du système, suivant ainsi la procédure de découpage en sous-périodes expliquée au chapitre 2. Un type de découpage similaire a été employé par Yapo *et al.* (1996), mais ces auteurs ont utilisé des années 'hydrologiques', avec chevauchement des périodes et seulement six mois de mise en route. Il leur a donc été possible d'obtenir des nombres très similaires de périodes longues et courtes. Considérant le lien de dépendance entre périodes induit par ce découpage, nous avons préféré opter pour des périodes sans recouvrement.

La situation idéale serait d'avoir un enregistrement suffisamment long pour pouvoir choisir un grand nombre de périodes, que celles-ci soient longues ou courtes. C'est ce qu'ont fait O'Donnell et Canedo (1980). Cependant ces auteurs ont utilisé des chroniques de pluie mises bout à bout et des données synthétiques de débit pour obtenir une séquence de 328 années, ce qui limite l'intérêt de leur étude.

Un découpage systématique a également été préféré à un découpage en fonction des caractéristiques des années, comme l'ont fait Sorooshian *et al.* (1983) sur années sèches et humides. Nous allons voir dans le paragraphe suivant que le découpage sans choix a priori donne également une assez bonne variabilité des caractéristiques des périodes.

6.5.2. Brève analyse sur les périodes sélectionnées

Le découpage en sous-périodes fournit des jeux de données différents pour les calages successifs du modèle. Les caractéristiques de ces jeux de données semblent pouvoir être un facteur d'influence prépondérant sur les résultats de la méthode 'multi-calage' que nous nous proposons de tester. Nous faisons donc ici quelques commentaires sur les caractéristiques des différentes sous-périodes étudiées.

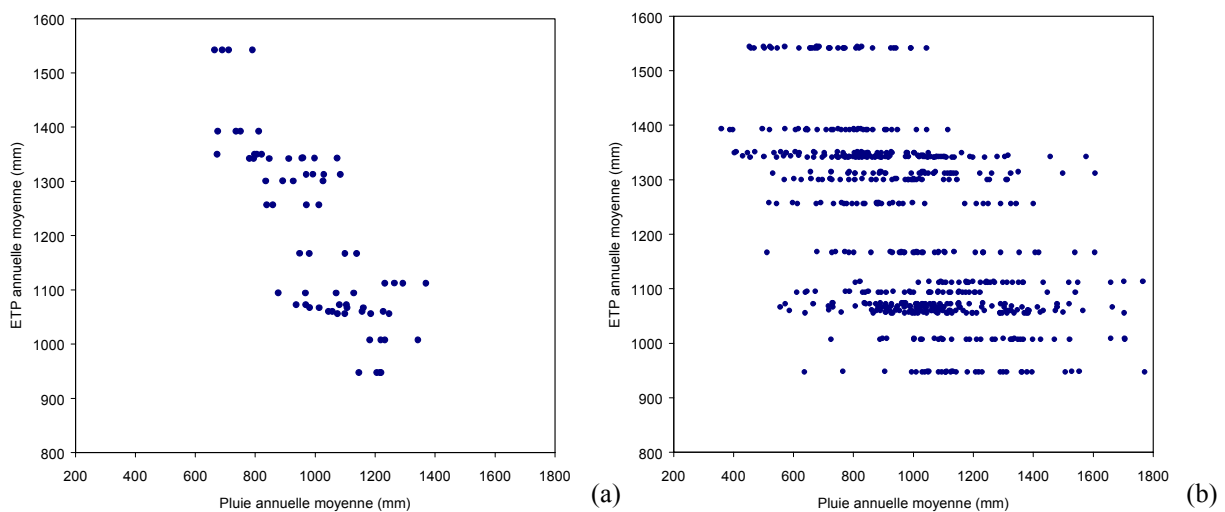


Figure 6.3: Variations des pluies et ETP moyennes pour des périodes de (a) 9 ans et (b) 1 an sur les 17 bassins versants (avec une seule valeur d'ETP interannuelle par bassin)

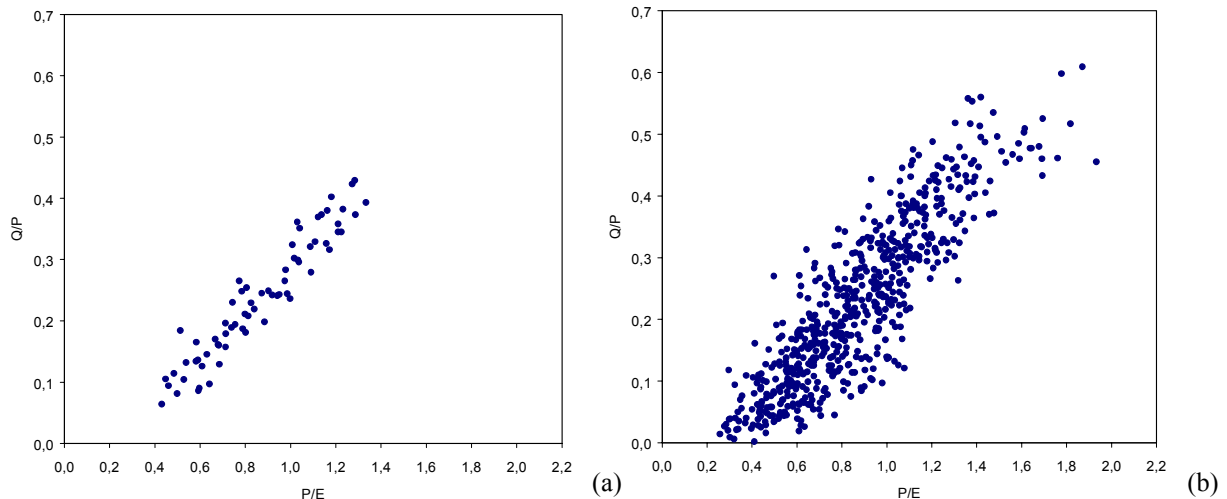


Figure 6.4: Rapport du débit à la pluie (Q/P) en fonction du rapport de la pluie à l'ETP (P/E) pour des périodes de (a) 9 ans et (b) 1 an sur les 17 bassins versants

Comme il fallait s'y attendre, il apparaît tout d'abord que la diminution de la longueur des périodes a pour effet d'augmenter la variabilité des caractéristiques moyennes de ces périodes. La Figure 6.3 montre par exemple que la variabilité des pluies annuelles moyennes est beaucoup plus importante dans le cas de courtes périodes que dans le cas de périodes longues. Par ailleurs il est important de noter (voir Figure 6.3 b) que les bassins étudiés sont soumis à des régimes pluviométriques très variables d'une année sur l'autre, avec des pluies annuelles pouvant varier (sur les 36 années disponibles) du simple au double ou du simple au triple. La Figure 6.4 montre que la gamme de variation des coefficients de rendement moyens des bassins évolue également significativement lorsque la longueur de la période change.

La spécificité des périodes (écart par rapport aux caractéristiques moyennes) augmente lorsque la longueur de la période diminue (résultat attendu). Le calage des paramètres du modèle sera ainsi effectué sur d'autant plus de cas s'écartant du comportement moyen du bassin que la série de calage sera courte. Par ailleurs, les données d'ETP étant des moyennes interannuelles, leur variabilité d'une année sur l'autre n'est pas prise en compte (voir Figure 6.3). Or il semble, d'après la Figure 6.4 (a et b), que dans le cas de ces 17 bassins, le rendement soit relativement bien corrélé au rapport de la pluie à l'ETP (P/E) lorsque la période est longue, cette corrélation diminuant lorsque la période diminue. Le rapport P/E est donc un bon explicateur du rendement du bassin sur des pas de temps de plusieurs années (c'est-à-dire lorsque l'ETP moyenne se rapproche de l'ETP moyenne interannuelle), mais son intérêt diminue lorsque la période se raccourcit. Ceci pourrait indiquer, même si ce n'est ici qu'une hypothèse, que l'absence de variabilité de l'ETP pourrait avoir une influence sur la description du comportement du bassin par le modèle, en particulier sur la valeur des paramètres de rendement.

6.6. Méthodologie

La méthodologie de comparaison appliquée dans cette étude est illustrée à la Figure 6.5. Le calage des paramètres des structures a été réalisé à l'aide de la méthode 'pas-à-pas' décrite au chapitre 3 en utilisant comme fonction objectif le critère de Nash-Sutcliffe (1970) calculé sur les débits (critère $CR1$).

Pour un bassin donné, après chaque calage sur une période donnée, des tests en contrôle sont effectués sur toutes les périodes restantes. Ainsi, avec une division de l'enregistrement en n sous-périodes, on obtient sur chaque bassin $n(n-1)$ validations, ce qui représente sur chaque bassin 12, 30, 72, 132, 306 et 1260 tests pour des périodes respectivement de 9, 6, 4, 3, 2 et 1 ans. Il existe donc un rapport de un à cent du nombre de tests entre les périodes les plus courtes et les plus longues. Le calage effectué sur la période complète de 36 ans fournit un jeu de paramètres de 'référence', auquel nous comparerons les jeux de paramètres obtenus sur les sous-périodes. La comparaison des résultats des deux approches d'analyse d'incertitude a été menée au niveau du bassin, mais également au niveau de l'échantillon des bassins (à une échelle régionale en quelque sorte) en faisant une moyenne des résultats sur cet échantillon.

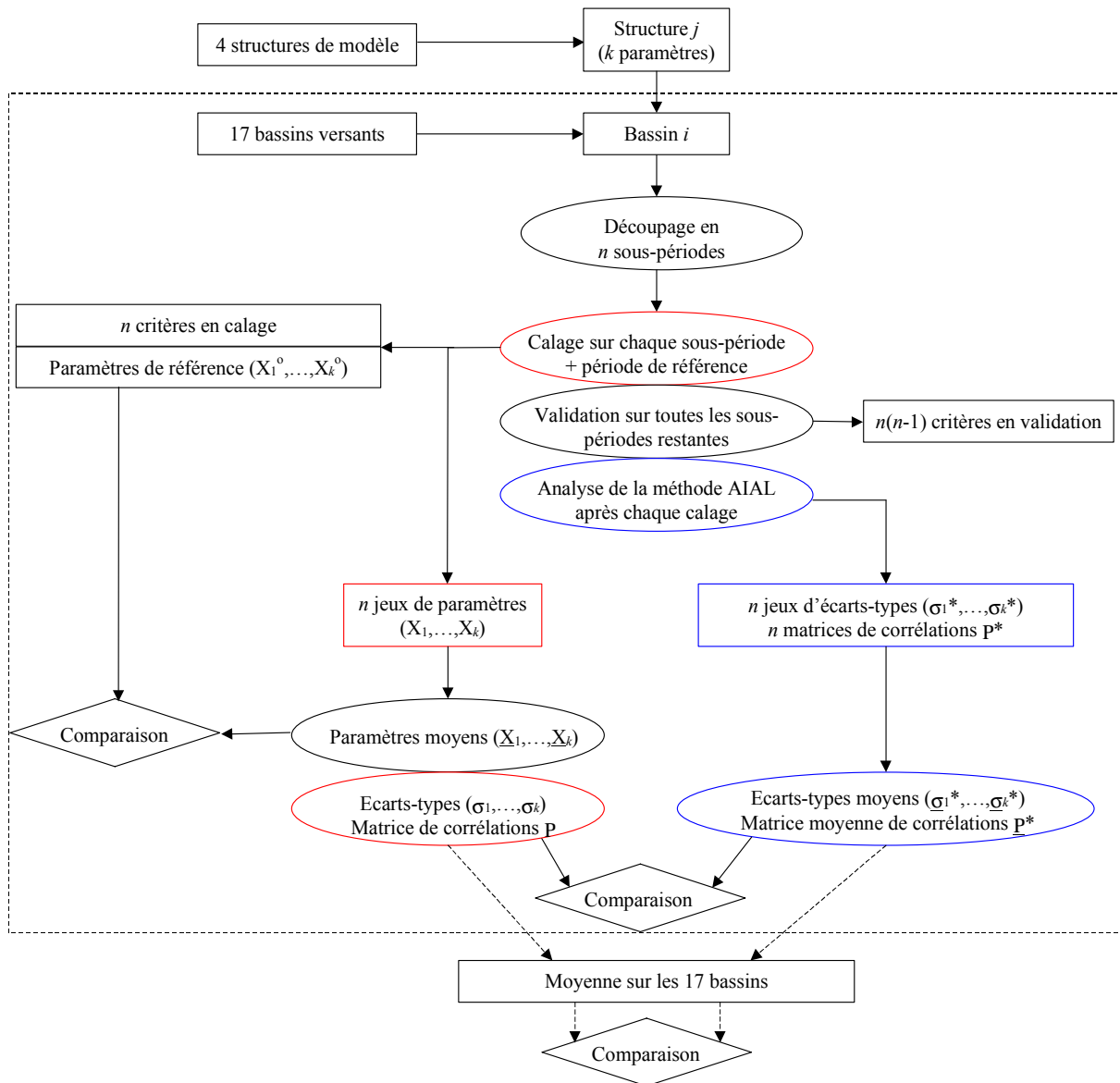


Figure 6.5: Schéma de la méthodologie de comparaison des deux méthodes d'analyse d'incertitudes (partie gauche du graphique: procédure 'multi-calage'; partie droite: procédure AIAL)

6.7. Résultats et discussion

6.7.1. Performances des différentes structures du modèle

Nous allons présenter ici une analyse des performances du modèle. Elles sont l'image de l'adéquation de la structure du modèle et du jeu de paramètres calés au bassin étudié. La bonne performance du modèle sur un bassin est une condition nécessaire pour assurer une bonne fiabilité des paramètres, même si elle ne représente pas une condition suffisante.

La Figure 6.6 montre les distributions des résultats du modèle GR41 obtenus sur les 17 bassins versants au calage et en contrôle pour les périodes de 1 à 9 ans. Les courbes obtenues pour les autres structures, montrant des comportements très similaires, ne sont pas présentées ici.

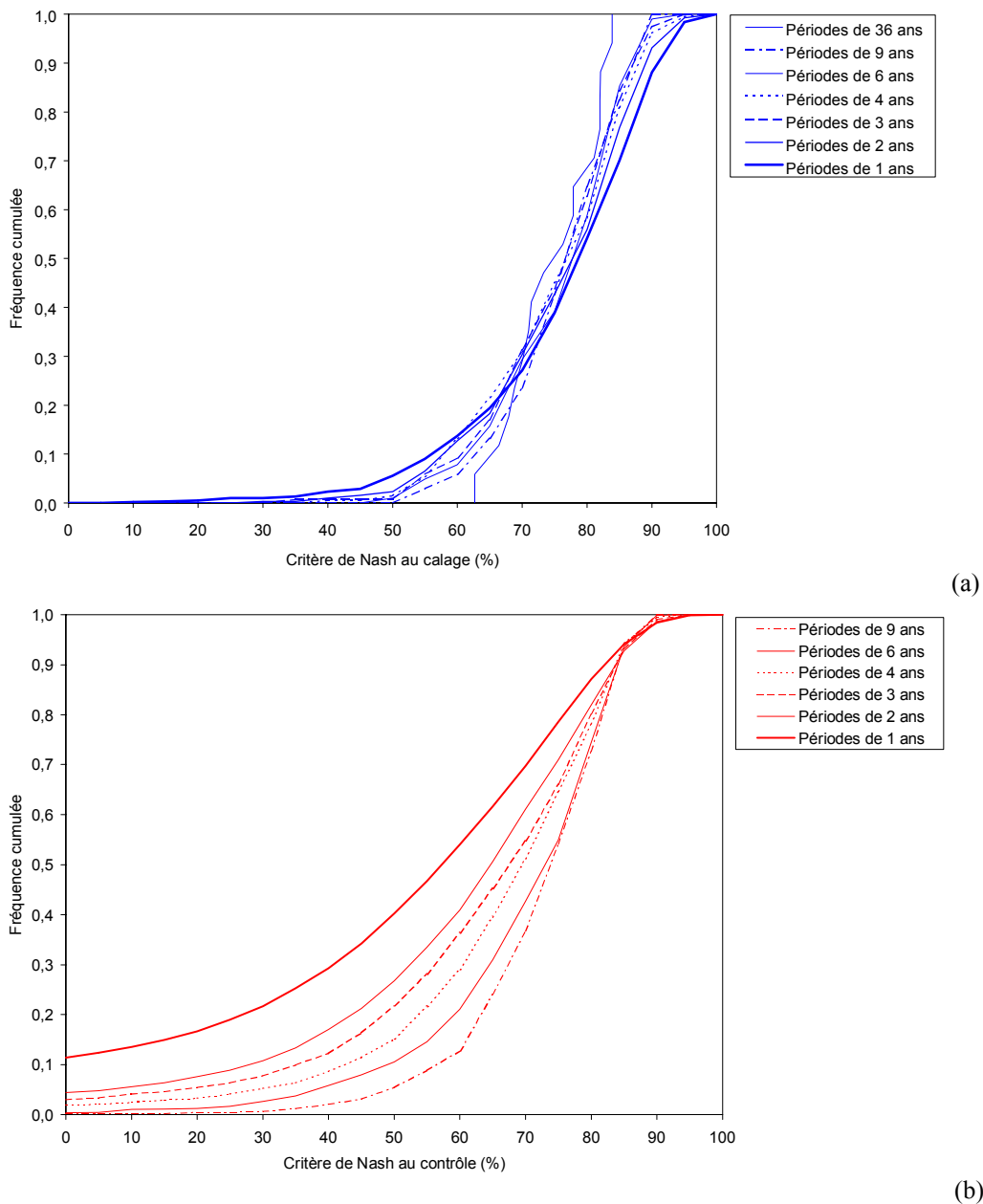


Figure 6.6: Distributions des résultats obtenus par le modèle GR41 (a) en calage et (b) en contrôle

En calage, plus la longueur de la période augmente, plus la pente des distributions augmente. Ceci indique que la performance du modèle devient moins sensible au choix de la période lorsque celle-ci augmente. Nous retrouvons ici une des conclusions de Yapo *et al.* (1996) avec la même fonction objectif dans le cas du modèle NWSRFS-SMA à 13 paramètres appliqué sur un bassin américain.

Par ailleurs, la Figure 6.6 (a) montre que le modèle réussit à atteindre des valeurs de critère plus élevées lorsque la période est plus courte et que, lorsque la longueur de la période augmente, le modèle est moins souvent en échec, avec moins de faibles valeurs du critère. Ceci se traduit par un entrecroisement des courbes de distribution correspondant aux différentes longueurs de période, et une évolution non significative de la performance moyenne avec la longueur de la période. Ces résultats sont en désaccord avec ceux de Yapo *et al.* (1996), qui montrent une diminution croissante des performances du modèle lorsque la longueur de la période diminue. Paradoxalement, nous retrouvons ici des résultats beaucoup plus similaires à ceux qu'ils exposent en utilisant le critère HMLE (Heteroscedastic Maximum Likelihood Error) pour fonction objectif.

Cette différence pourrait provenir du choix de la méthode de découpage en sous-périodes. Ici, nous disposons pour chaque bassin de 36 périodes de calage de 1 an mais seulement 4 périodes de 9 ans. Un nombre plus important de périodes longues conduirait probablement à un nombre plus important d'échecs du modèle. En contrôle, il existe en fait plus de chances de trouver parmi 36 périodes une période très similaire à la période de calage que lorsque l'on n'a que quatre périodes de neuf ans (seulement trois cas à comparer).

Pour vérifier cette hypothèse, nous avons appliqué la méthode de découpage de Yapo *et al.* (1996) en périodes non indépendantes, dans le cas des périodes de 9 ans. 28 sous-périodes ont ainsi pu être distinguées pour caler le modèle. Les résultats, donnés à la Figure 6.7, montrent une différence faible entre les distributions en calage d'une part et les distributions en contrôle d'autre part lorsque l'on compare les deux découpages, avec seulement une faible diminution des performances en utilisant des périodes non indépendantes. Il est probable que cette légère différence entre les deux techniques de découpage se retrouve également lorsque la longueur de la période diminue, conservant ainsi un entrecroisement des distributions à la Figure 6.6 (a). Le choix de la méthode de découpage ne semble donc pas influencer de manière drastique les performances du modèle.

Deux autres hypothèses peuvent être avancées pour expliquer la différence entre les deux études. Elle pourrait provenir de l'échantillon de données, Yapo *et al.* (1996) n'ayant utilisé dans leurs travaux qu'un seul bassin versant. La seconde hypothèse est que le modèle à 13 paramètres de Yapo *et al.* (1996) aurait besoin de séries de données relativement longues (supérieures à huit ans) pour obtenir des résultats stables, comme l'affirment ces auteurs, et que la structure GR41 à quatre paramètres serait donc moins exigeante de ce point de vue. Au regard des tests effectués au chapitre précédent, ce dernier point ne semble pas être la raison principale. Néanmoins, tester ces deux hypothèses nous demanderait de reprendre entièrement l'approche de Yapo *et al.* (1996) sur les mêmes données, ce qui nous entraînerait trop loin du cadre de notre étude. Nous nous en tiendrons donc ici à ce constat.

En validation, plus la période est courte, plus les résultats moyens diminuent, avec une proportion plus importante de faibles valeurs du critère de performance (voir Figure 6.6 (b)). Par ailleurs, plus la longueur de la période augmente, moins l'écart entre plus faibles et plus fortes valeurs du critère est important, ce qui tend à montrer que la sensibilité des performances du modèle aux données diminue lorsque la longueur de la période augmente.

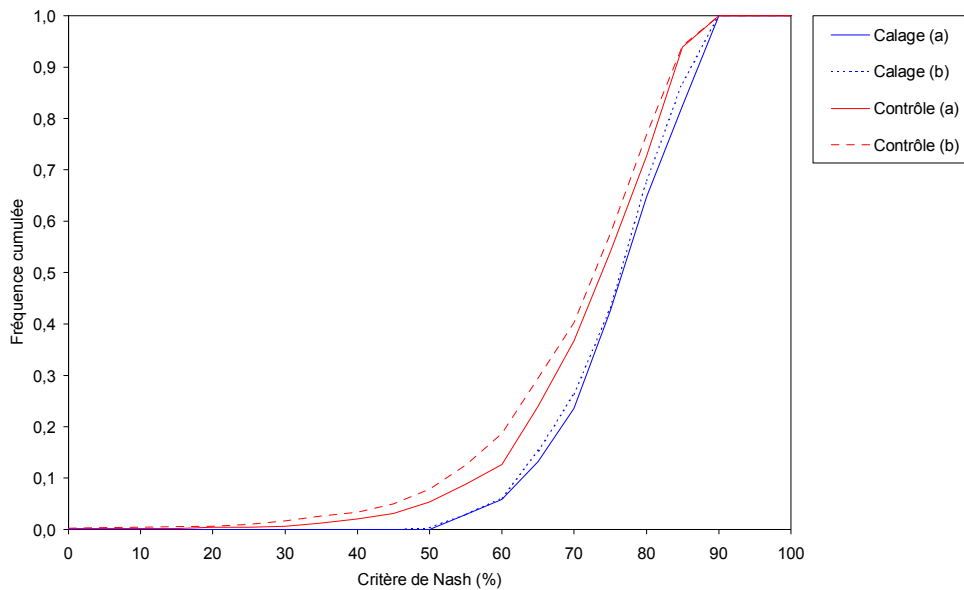


Figure 6.7: Distributions des résultats en calage et en contrôle pour le modèle GR41 en utilisant des périodes de 9 ans (a) indépendantes et (b) non indépendantes

Cette tendance claire pourrait venir du fait que les courtes périodes ne contiennent pas une variabilité suffisante d'événements hydrologiques pour pouvoir fournir assez d'information sur le comportement du bassin pour une identification fiable des paramètres. Il est souvent difficile pour un modèle calé sur une année sèche de pouvoir donner de bons résultats sur une période très humide. Ceci se traduirait par une plus grande variabilité des jeux de paramètres lorsqu'ils sont calés sur des périodes plus courtes, ce que nous étudierons dans la suite. Comme nous l'avons vu dans l'analyse des périodes de calage, l'allongement de la période réduit l'écart des caractéristiques de la période à un comportement moyen du bassin à long terme. En allongeant la longueur de la période, celle-ci inclue donc une plus grande variété de conditions climatiques. Les calages sur plus longues périodes rendent ainsi le jeu de paramètres moins spécifique à des conditions climatiques particulières. Par ailleurs, comme dans le cas du calage, la Figure 6.7 montre que le choix de la procédure de découpage influence peu les résultats en contrôle.

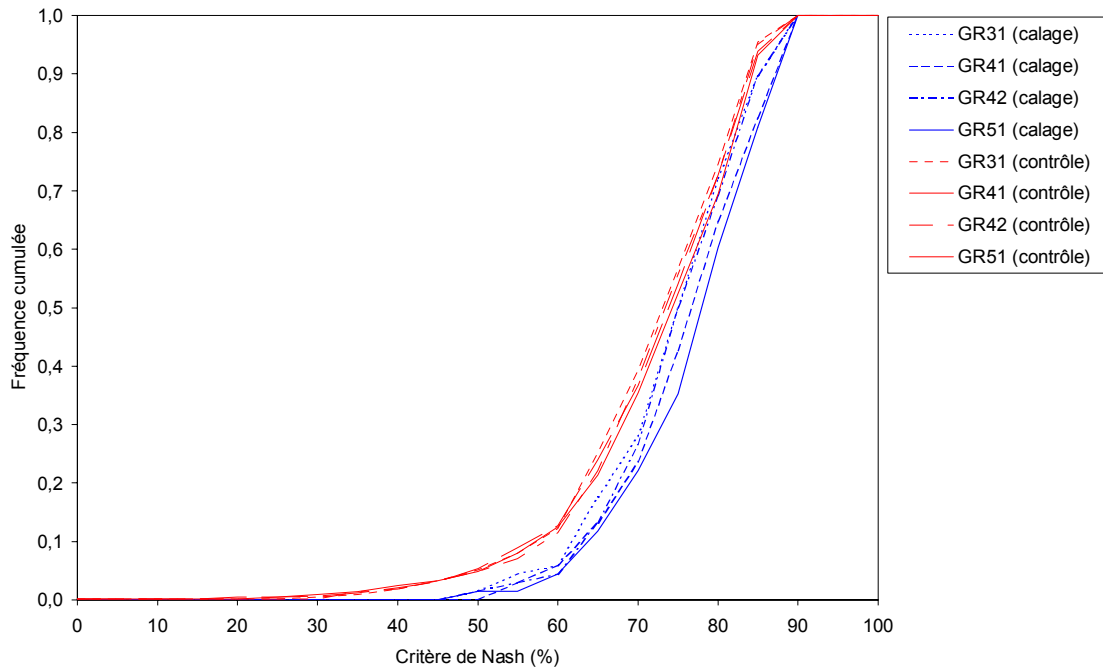
L'allongement de la période de calage conduirait donc à des performances plus stables du modèle, grâce à une moins grande dépendance à l'échantillon de données. Notons tout de même que nous aurions dû, pour être complet et évaluer plus finement l'effet de la longueur de la période de calage sur les performances en contrôle, faire des combinaisons de calage/contrôle en croisant les longueurs des périodes (par exemple, calage sur 9 ans et contrôle sur 4 ans). Dans les résultats précédents, nous nous sommes placés dans le cas particulier de calage/contrôle sur des périodes de même longueur. Différentes longueurs permettraient probablement de définir une éventuelle longueur optimale pour la période de calage.

Nous allons maintenant voir si l'augmentation de la complexité de la structure induit de meilleures performances du modèle. La Figure 6.8 fournit une illustration comparative des résultats des quatre structures pour trois longueurs de période (1, 4, et 9 ans).

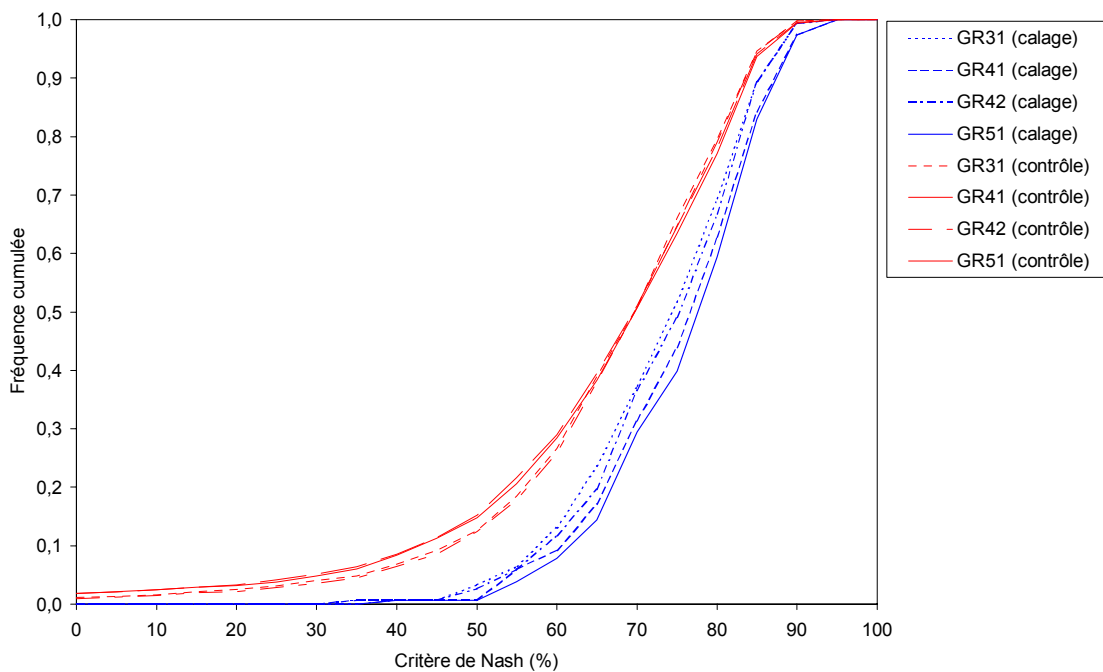
En calage, les différences entre modèles sont faibles, avec par ordre de performance croissante, GR31, GR42, GR41 et GR51, quelle que soit la longueur de période considérée. La complexité croissante avec davantage de paramètres libres donne une plus grande flexibilité au modèle pour s'adapter aux spécificités de la période de calage. Les résultats des

structures GR31 et GR42 d'une part, et GR41 et GR51 d'autre part, sont très similaires, indiquant que le gain de performance en utilisant le paramètre supplémentaire X5 (coefficient de partage de la pluie efficace) est négligeable. La différence cette fois entre ces deux couples de structures est assez significative avec de meilleures performances pour les structures GR41 et GR51 qui laissent libre le paramètre X4 (capacité du réservoir de production).

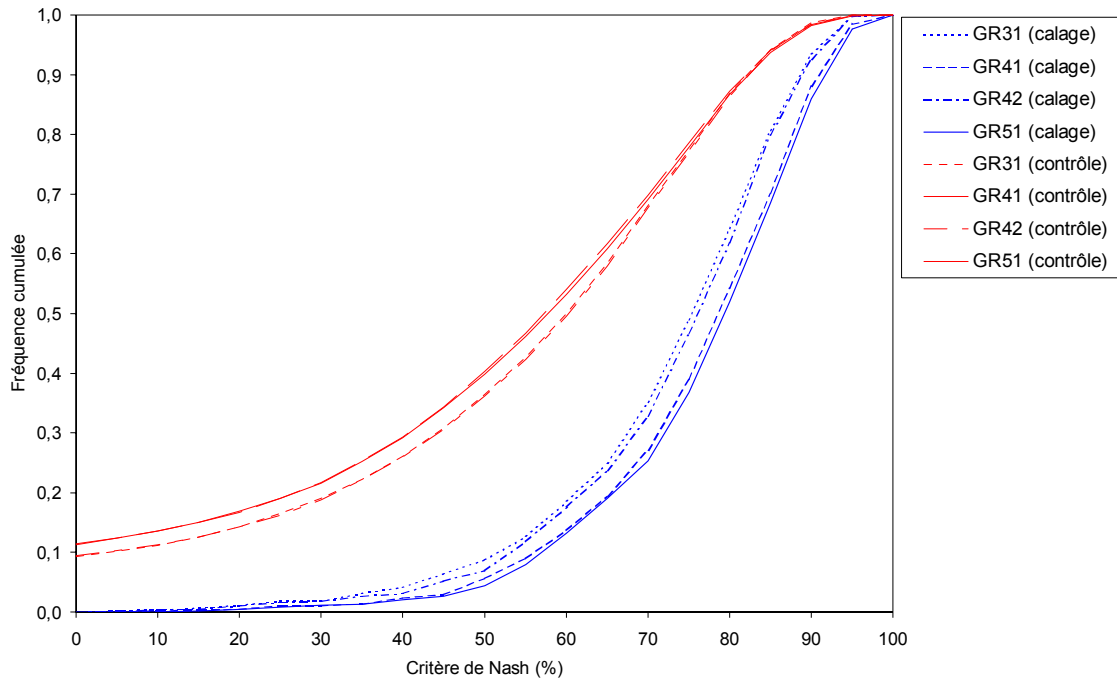
En contrôle, cette similarité entre structures est conservée. L'amélioration apportée par l'introduction du paramètre X5 est quasiment nulle. L'avantage des deux structures GR41 et GR51 sur les structures GR31 et GR42 se maintient en contrôle pour les longues périodes, mais s'inverse pour les courtes périodes, où l'optimisation du réservoir de production n'est pas bénéfique.



(a)



(b)



(c)

Figure 6.8: Distributions des résultats en calage et en contrôle pour les structures GR31, GR41, GR42, GR51 dans les cas de périodes de (a) 9 ans, (b) 4 ans et (c) 1 an.

Ce comportement est probablement à mettre en lien avec les commentaires précédents soulignant l'apparente difficulté d'obtenir des paramètres robustes lorsque la période devient trop courte. Dans le cas de la capacité du réservoir de production, une capacité moyenne permet d'obtenir des résultats plus stables lorsque l'on ne dispose que de courtes périodes de calage. La liberté laissée au deuxième paramètre de rendement (le réservoir de production) aurait pour effet de 'surcaler' le module de production du modèle sur la période de calage, le rendant ainsi trop dépendant des caractéristiques de cette période.

Quelques conclusions peuvent être tirées de l'analyse des performances des structures:

1. La longueur de la période utilisée pour le calage du modèle n'a pas montré d'influence claire sur les performances moyennes en calage. L'augmentation de la période a cependant pour effet de rendre les performances moins sensibles aux caractéristiques des données de calage. En phase de contrôle, des distributions de résultats plus satisfaisantes sont obtenues lorsque la longueur de la période augmente. Les périodes plus longues, en mettant à disposition en phase de calage une plus grande variabilité d'événements hydrologiques (c'est-à-dire une quantité d'information plus grande sur le comportement du bassin) semblent permettre l'identification de paramètres plus robustes. Ils sont en effet moins dépendants des caractéristiques d'une ou deux années particulières. Ce dernier résultat doit être pris prudemment, étant donné que nous nous sommes placés dans le cas particulier de tests en contrôle sur des périodes de même longueur que celles utilisées en calage. Le choix d'un découpage en sous-périodes indépendantes, qui réduit le nombre de longues périodes, ne semble pas être à l'origine de ce comportement.
2. L'ajout de degrés de liberté supplémentaires dans la structure donne au modèle la possibilité d'obtenir de meilleures performances au calage (résultat attendu). L'optimisation du coefficient de partage X5 n'apporte qu'un gain marginal en terme d'efficacité, alors que celle de la capacité du réservoir de production X4 permet des améliorations significatives. En contrôle, l'apport de X5 est quasiment nul. L'avantage apporté par X4 se maintient pour les longues périodes. En revanche, il diminue

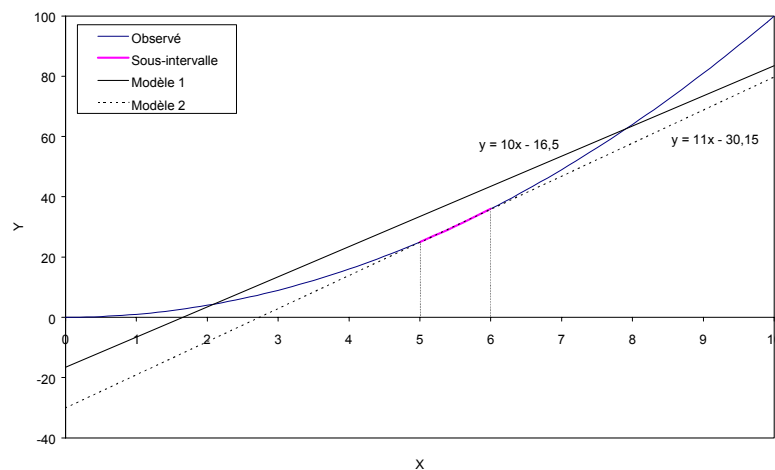
progressivement (et même s'annule) lorsque la longueur de la période diminue. Ceci pourrait provenir d'une situation de surcalage⁵ du module de production pour les courtes périodes. Ces résultats tendent à indiquer qu'ici l'augmentation de la complexité du modèle ne garantit ni de meilleures performances en contrôle, ni une meilleure robustesse. L'ajout de X5 ne semble donc pas être une solution bénéfique de complexification de la structure du modèle GR, alors que conceptuellement, l'idée d'une proportion variable de deux composantes d'écoulement d'un bassin à l'autre paraît tout à fait réaliste. Optimiser X4, qui paraissait être une solution envisageable au vu des résultats du modèle GR4K dans la comparaison, pourrait ne pas être non plus la solution la plus favorable.

6.7.2. Valeurs des paramètres calés

Dans cette partie, nous allons nous intéresser aux valeurs des paramètres obtenus par calage, pour compléter les observations faites sur les performances des structures. Nous allons notamment comparer les paramètres moyens obtenus sur chaque bassin (moyenne sur toutes les sous-périodes) avec les paramètres de 'référence' obtenus par calage sur la période entière d'enregistrement. Tous les résultats suivants utilisent des valeurs transformées des paramètres qui sont adimensionnelles.

Il y a généralement un bon accord entre les paramètres de bassin et les paramètres de 'référence', en particulier pour les paramètres X3 (temps de base de l'hydrogramme) et X4 (capacité du réservoir de production). Un exemple est donné à la Figure 6.9 dans le cas du paramètre X4 de la structure GR51. Il n'y a pas d'évolution significative de ces correspondances pour les paramètres X3 et X4 lorsque la longueur de la période varie.

⁵ Nous entendons par surcalage la situation où les paramètres du modèle deviennent trop spécifiques à l'échantillon de calage, le gain d'ajustement sur cet échantillon ayant pour contrepartie une moins bonne transposabilité du modèle calé à d'autres observations. Ceci est illustré à la figure ci-dessous dans le cas de l'exemple simple suivant: on veut modéliser la relation liant une variable Y à une variable X par une relation linéaire (ici, les observations sont supposées correspondre à la relation $Y = X^2$ inconnue a priori du modélisateur). Si l'on considère l'intervalle total d'observation ($X \in [0;10]$) pour caler le modèle (modèle 1), on obtient la meilleure relation linéaire décrivant toutes les observations (erreur quadratique moyenne: 57,8). Si en revanche, on ne prend en compte qu'un sous-intervalle ($X \in [5;6]$) pour le calage, le modèle ajusté (modèle 2) est très satisfaisant sur cette partie des observations (erreur quadratique moyenne: 0,01), mais il est nettement moins satisfaisant que le modèle 1 si l'on considère l'ensemble de la période (erreur quadratique moyenne 141,1). Ici, le sous-intervalle [5;6] est l'analogue d'une courte période d'observation. On a un phénomène de surcalage dans le sens où l'on a obtenu de bons résultats qui ne sont pas confirmés quand on étend la période d'observation.



Cependant cet accord, dans le cas des paramètres X1 (coefficient d'échange) et X2 (capacité du réservoir de routage), est meilleur lorsque la longueur de la période de calage augmente, spécialement dans le cas des modèles GR31 et GR41 et dans une moindre mesure pour les structures GR42 et GR51. Ceci est confirmé par la Figure 6.10 montrant l'évolution des paramètres régionaux (moyenne sur tous les bassins) avec la longueur de la période pour les quatre structures. Lorsque la longueur de la période diminue, le paramètre X1 augmente et le paramètre X2 diminue. X1 et X2 semblent atteindre un plateau lorsque la période augmente, prenant des valeurs plus proches de la valeur de référence.

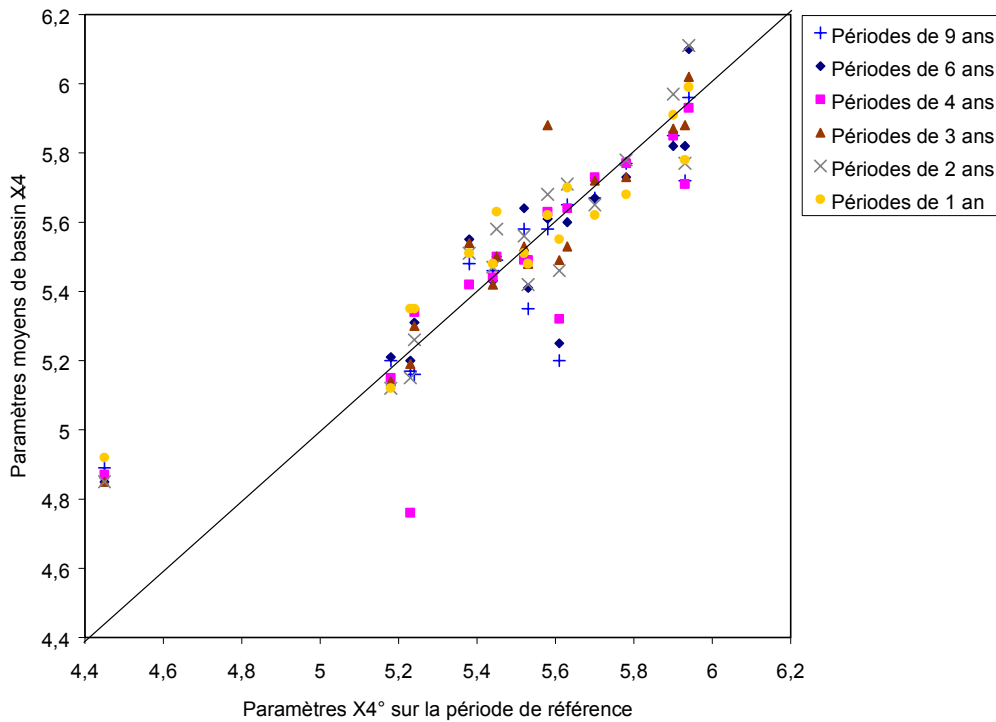
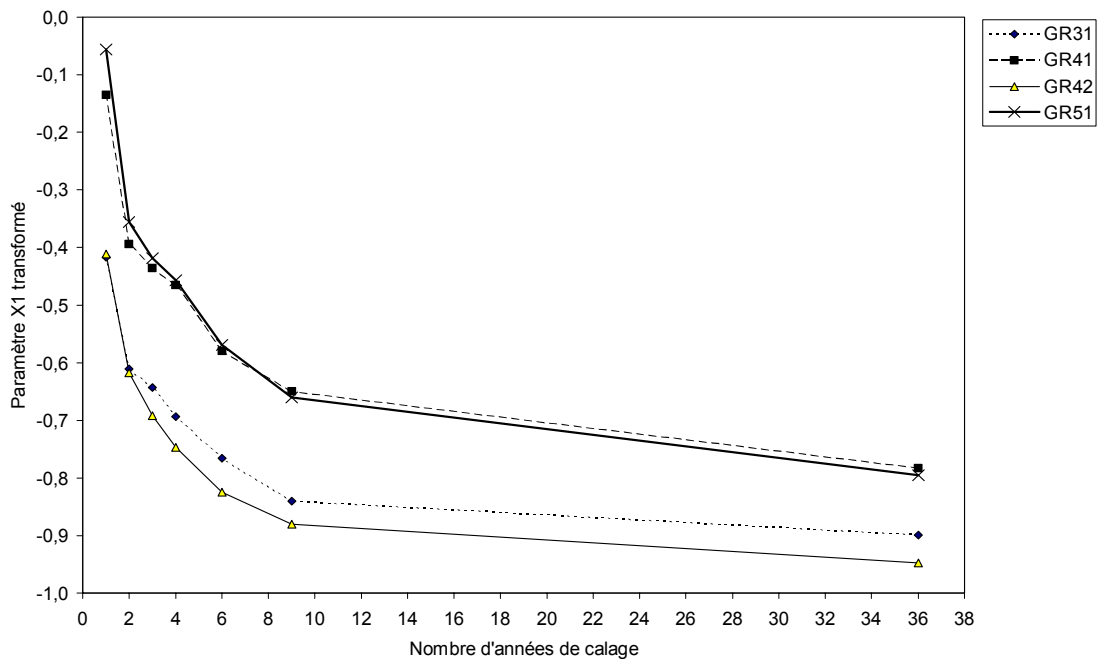
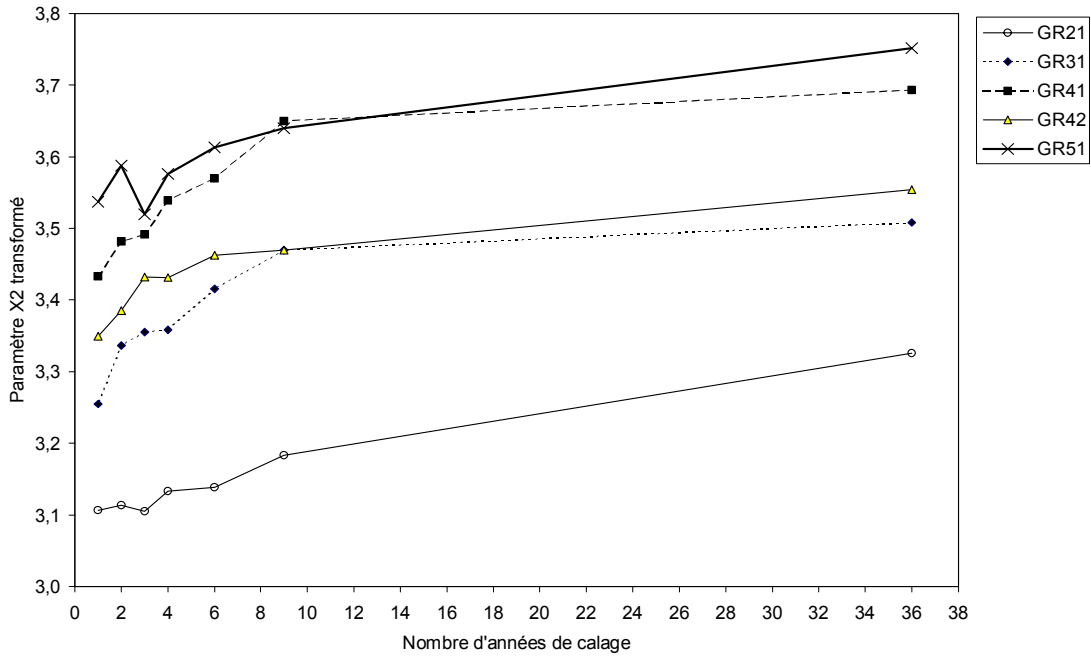


Figure 6.9: Comparaison des paramètres moyens $\underline{X}4$ et des paramètres de référence $X4^0$



(a)



(b)

Figure 6.10: Evolution des paramètres 'régionaux' X1 (a) et X2 (b) avec la longueur de la période

Dans le cas du paramètre X1, la Figure 6.11 montre que l'augmentation progressive des valeurs 'régionales' est due à une part croissante de valeurs positives du paramètre lorsque la période diminue: pour les périodes de 9 ans, presque 95 % des paramètres sont négatifs alors que seulement 70 % des paramètres le sont pour les périodes de 1 an. Par ailleurs, la Figure 6.12 illustre le fait qu'il semble exister un lien entre le paramètre X1 et le rendement observé du bassin, X1 augmentant (avec davantage de pertes en eau) lorsque le rendement diminue. Cette tendance n'est pas observable sur l'autre paramètre de rendement X4 du modèle (capacité du réservoir de production).

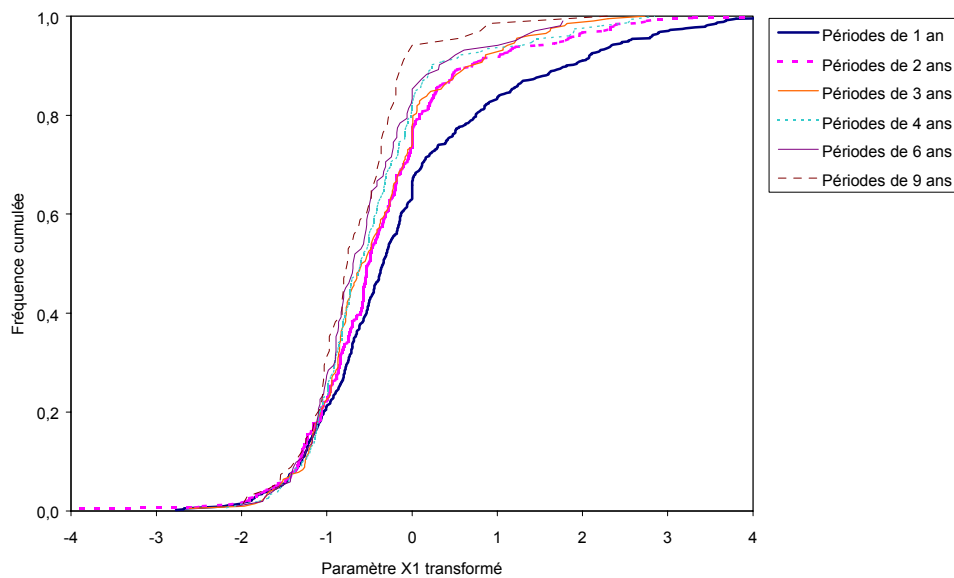


Figure 6.11: Distributions des valeurs de X1 pour la structure GR51 pour différentes longueurs de périodes

Ceci indiquerait que la dépendance de la valeur de X1 avec la longueur de la période vient d'une part du fait que les conditions de rendement sur le bassin sont plus variables d'une période à l'autre lorsque la période est courte et d'autre part du fait que X1 semble plus sensible que l'autre paramètre de rendement du modèle (X4) pour permettre l'ajustement du bilan en eau du bassin.

Cependant, nous devons rappeler ici que les données d'ETP utilisées, qui agissent sur le réservoir de production, sont ici des moyennes interannuelles, négligeant les variations entre années. Ceci pourrait favoriser la stabilité du paramètre X4.

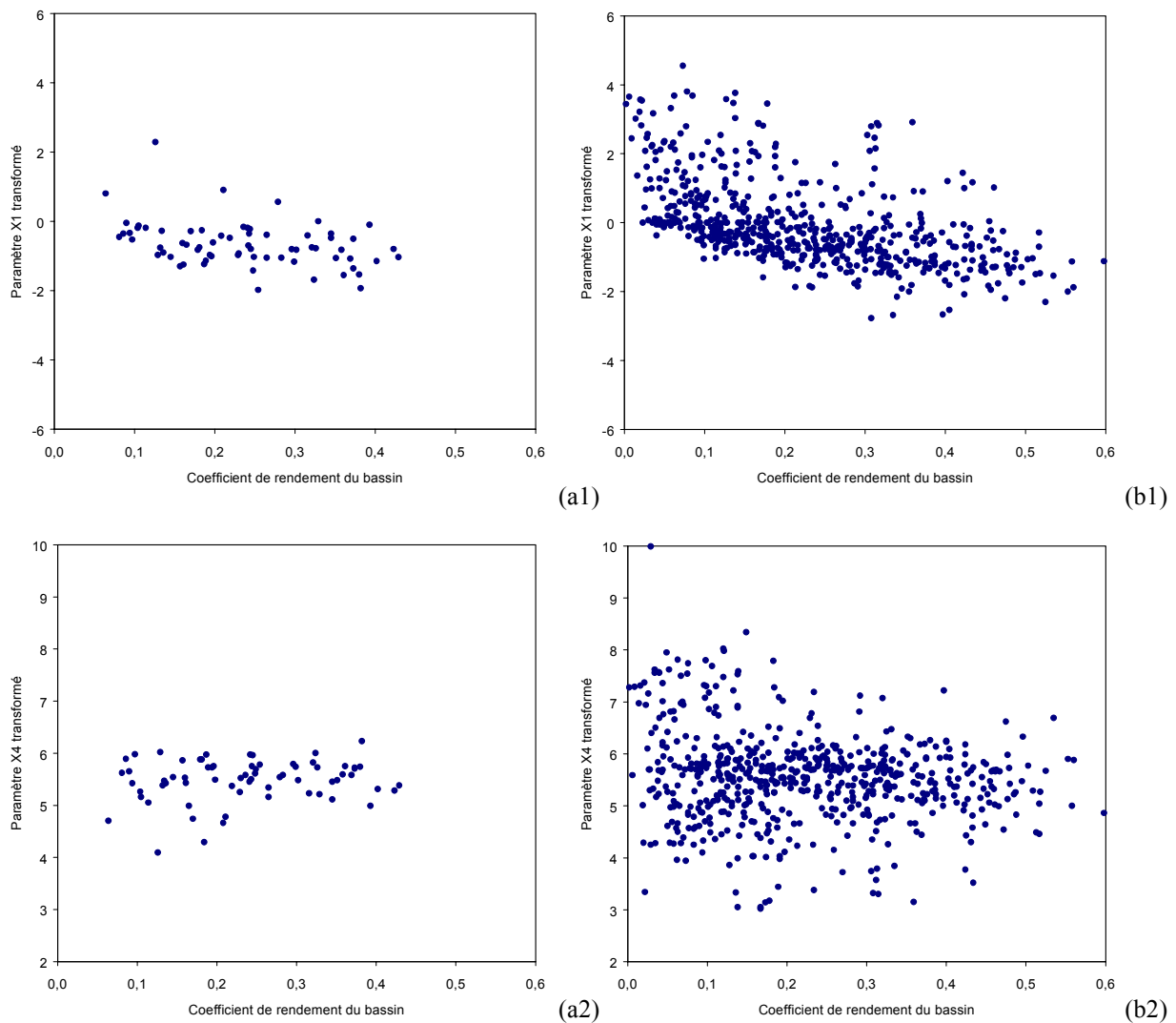


Figure 6.12: Liens entre les paramètres X1 et X4 de la structure GR51 et le rendement du bassin pour des périodes (a1,a2) de 9 ans et (b1,b2) de 1 an

Pour essayer d'identifier la nature du lien possible de X2 avec la longueur de la période et voir s'il est lié à une interaction entre X1 et X2, nous avons testé une structure GR21 avec deux paramètres optimisés (X2 et X3), pour laquelle nous avons fixé le paramètre d'échange à zéro. X2 montre toujours une tendance (moins importante) à l'augmentation avec la longueur de la période (voir Figure 6.10). La raison majeure de la variation de X2 en fonction de la longueur de la période ne semble donc pas provenir de sa liaison avec X1, mais davantage de l'influence des caractéristiques de la période. Ici encore, cette dépendance croissante lorsque

la période raccourcit semble pouvoir être traduite en terme de diminution de la variabilité des conditions hydrologiques rencontrées sur la période.

Nous avons vérifié par ailleurs (voir Annexe 5) que le changement des paramètres X1 et X2 avec la longueur de la période n'était pas dû à l'influence croissante de la période d'initialisation lorsque la longueur de la période d'étude diminue.

La Figure 6.10 montre également que les paramètres des structures GR31 et GR42 d'une part et GR41 et GR51 d'autre part, sont assez similaires, comme l'étaient les performances de ces différentes structures. L'optimisation de la capacité X4 du réservoir de production influence beaucoup les valeurs de X1 et X2. Ceci est facilement compréhensible dans le cas du paramètre X1, étant donné que c'est comme X4 un paramètre de rendement du modèle. Le changement dans les valeurs de X2, paramètre de routage, pourrait être dû à des interactions entre X1 et X2. Cette hypothèse est discutée dans la suite.

Il y a une bonne correspondance des valeurs du paramètre X3 (temps de base de l'hydrogramme) entre les structures GR31 et GR41 d'une part et GR42 et GR51 d'autre part. L'influence de l'ajout du coefficient de partage X5 sur ce paramètre X3 est donc plus importante que celle du paramètre X4. En effet, X5 gouverne la quantité d'eau qui transite par l'hydrogramme unitaire HU1 qui dépend uniquement de X3. Cependant l'optimisation de X5 diminue l'accord entre les paramètres X3 de 'référence' et moyens, ce qui pourrait indiquer une plus grande incertitude sur la détermination de X3 lorsque X5 est optimisé.

Les valeurs de X4 pour les structures GR41 et GR51 sont quasiment identiques, montrant que l'optimisation du coefficient de partage X5 n'influence pas la capacité du réservoir de production.

Nous pouvons résumer les commentaires précédents sur l'évolution des paramètres en fonction de la longueur de la série de calage ou de la structure du modèle, par les éléments suivants:

1. Une dépendance des paramètres X1 et X2 avec la longueur de la période de calage a été observée. Elle est à mettre en lien direct avec la variabilité des conditions hydro-climatiques rencontrées sur la période, dont ces deux paramètres restent dépendants. Le modèle pourrait utiliser le paramètre d'échange X1 avec plus de flexibilité que l'autre paramètre de rendement X4 pour ajuster les bilans en eau et s'adapter aux spécificités des périodes. Une autre cause pourrait venir du fait que des ETP moyennes interannuelles, agissant sur le réservoir de production, sont utilisées ici.
2. L'optimisation de la capacité X4 du réservoir de production influence logiquement la valeur de l'autre paramètre de rendement X1 (coefficient d'échange) puisque les deux paramètres agissent tous deux sur le bilan en eau. Le changement consécutif des valeurs de X2 indiquerait que des interactions existent entre X1 et X2.
3. L'optimisation du coefficient de partage X5 influence la valeur du temps de base X3 de l'hydrogramme unitaire. Lorsque X5 est libre, les paramètres X3 moyens et de référence sont moins bien corrélés, ceci pouvant être le signe d'une plus grande incertitude sur X3. Cependant, l'optimisation de X5 réduit légèrement la dépendance de X1 et X2 à la longueur de la période de calage, ceci étant probablement le simple effet de l'augmentation du nombre de degrés de liberté.

6.7.3. Corrélations entre paramètres

Nous avons mentionné dans les paragraphes précédents de possibles interactions entre les paramètres de la structure. Nous allons essayer d'examiner ici cette hypothèse plus en détail.

Pour cela, nous utilisons les corrélations entre paramètres fournis par les deux méthodes d'analyse d'incertitude, l'analyse par approximation linéaire et l'approche 'multi-calage'.

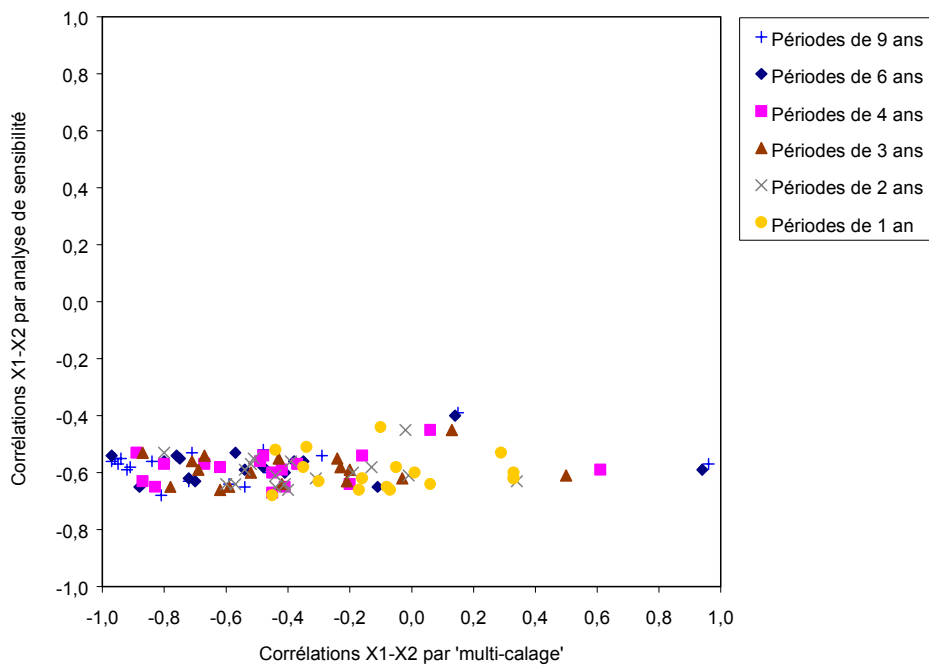


Figure 6.13: Corrélations entre les paramètres X1 et X2 de la structure GR31 fournies par approche 'multi-calage' et méthode AIAL (valeurs moyennées par bassin) pour les cinq longueurs de périodes

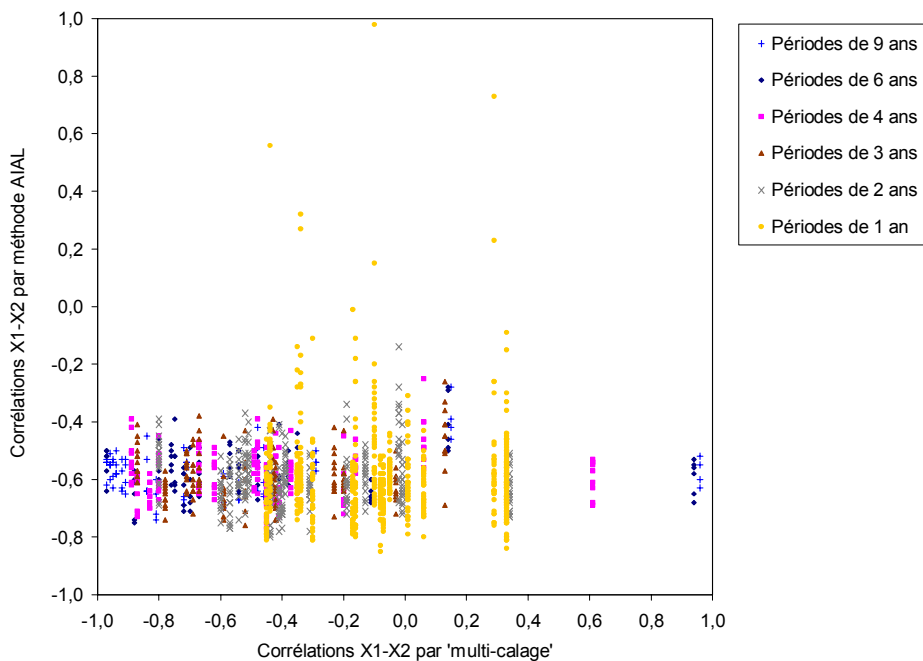


Figure 6.14: Corrélations entre les paramètres X1 et X2 de la structure GR31 fournies par approche 'multi-calage' et méthode AIAL (valeurs non moyennées par bassin) pour les cinq longueurs de périodes

Un fait général observé est qu'il existe une très faible correspondance (s'il y en a une) entre les corrélations fournies par les deux méthodes, quelle que soit la structure du modèle ou la longueur de la période de calage. La Figure 6.13 illustre cela pour les paramètres X1 et X2 de

la structure GR31. Les corrélations données par la méthode ‘multi-calage’ varient toujours sur des intervalles plus grands que pour la méthode d’AIAL. Comme le montre la Figure 6.14, cette indépendance n’est pas due au fait que l’on utilise ici des valeurs moyennes sur plusieurs périodes dans le cas de l’AIAL.

La plage de variation des corrélations données par l’AIAL ne varie pas beaucoup lorsque la longueur de la période diminue, alors qu’elle tend à décroître avec l’autre méthode. Ceci proviendrait du fait qu’avec les courtes périodes, il y a plus de jeux de paramètres pour calculer les coefficients de corrélation, laissant ceux-ci moins susceptibles de varier à cause de *horsains* (points de l’échantillon situés très en dehors du nuage regroupant la majorité des points). L’AIAL accorde bien plus d’importance aux relations structurelles entre paramètres que ne le fait l’approche multi-calage, fournissant ainsi plus d’information sur les problèmes rencontrés lors de l’optimisation.

Bien qu’il y ait un important désaccord entre les deux approches, les valeurs moyennes ‘régionales’ indiquent généralement des tendances similaires comme le montre la Figure 6.15 pour les corrélations entre paramètres X1 et X2 des quatre structures. En représentant ces valeurs en fonction de la longueur de la série de calage, on s’aperçoit que les corrélations données par méthode AIAL sont assez stables, alors que l’autre approche donne des résultats plus variables, ce qui peut être relié aux commentaires précédents.

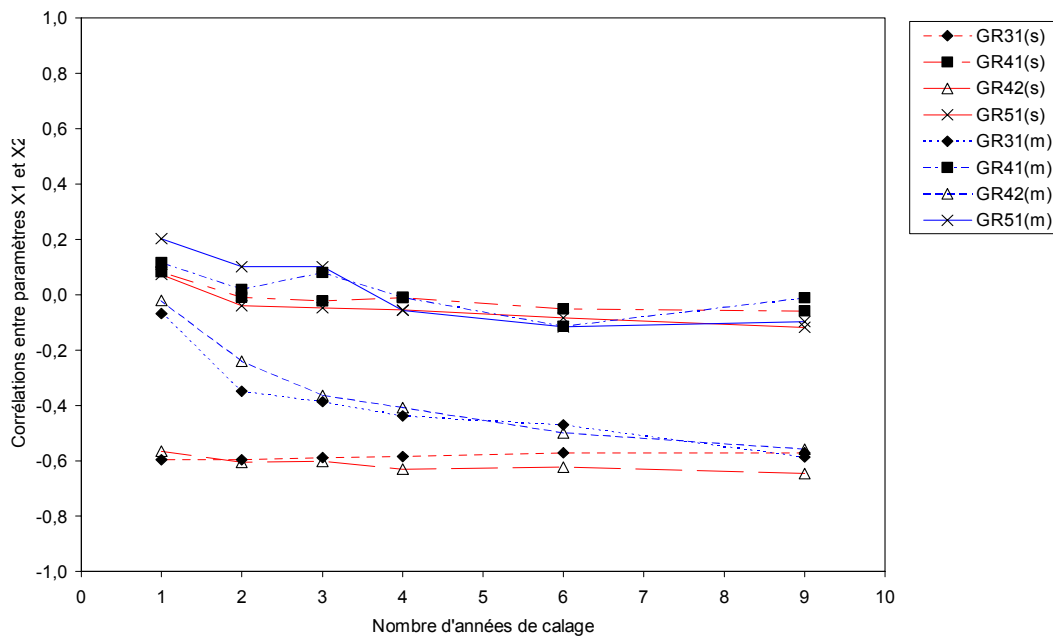


Figure 6.15: Corrélations moyennes entre paramètres X1 et X2 obtenues par AIAL (s) et par ‘multi-calage’ (m)

Les corrélations observées, que ce soit par l’une ou l’autre des méthodes, sont faibles (inférieures à 0,4) dans beaucoup de cas. C’est le cas par exemple entre X4 (capacité du réservoir de production) et X5 (coefficient de partage), ou entre X1 (coefficient d’échange) et X5. Parmi les corrélations les plus significatives, on trouve la corrélation entre X1 et X2 pour les structures GR31 et GR42. Cette corrélation devient faible lorsque le paramètre X4 est libre, c’est-à-dire lorsque l’on passe aux structures GR41 et GR51 respectivement (voir Figure 6.15). Par ailleurs, l’optimisation de X4 change la corrélation entre X2 et X3 (faible diminution). Dans les structures GR41 et GR51, une corrélation assez élevée (environ -0,6) existe aussi entre les paramètres X2 et X4, et dans une moindre mesure, entre X1 et X4.

Les principaux résultats de l’analyse de corrélation sont résumés par les points suivants:

1. Il y a peu de correspondance entre les résultats des deux méthodes d'analyses d'incertitude. L'analyse d'incertitude par approximation linéaire est plus influencée par les caractéristiques structurelles de la surface de réponse qui ne semblent pas varier beaucoup avec le changement des données présentes dans la période de calage. L'approche 'multi-calage', elle, reflète davantage la variabilité des caractéristiques des données de calage.
2. Les corrélations entre les paramètres X1 et X2, et entre ces deux paramètres et le paramètre X4, confirment les liens supposés précédemment. La corrélation entre X1 et X2 est atténuée lorsque le second paramètre de rendement (X4) est optimisé. Cependant les corrélations entre X1 et X4 ou X2 et X4 sont assez significatives. L'introduction du coefficient de partage X5 ne modifie pas beaucoup les corrélations entre paramètres. Les deux paramètres de routage interagissent faiblement, mais leur lien est légèrement modifié lorsque X4 est optimisé, en raison du lien entre X4 et X2.

6.7.4. Incertitude sur les paramètres

Comme dernier point d'analyse sur les structures, nous nous intéressons ici à l'incertitude sur les paramètres quantifiée par les écarts-types sur les paramètres fournis par les deux méthodes d'analyse.

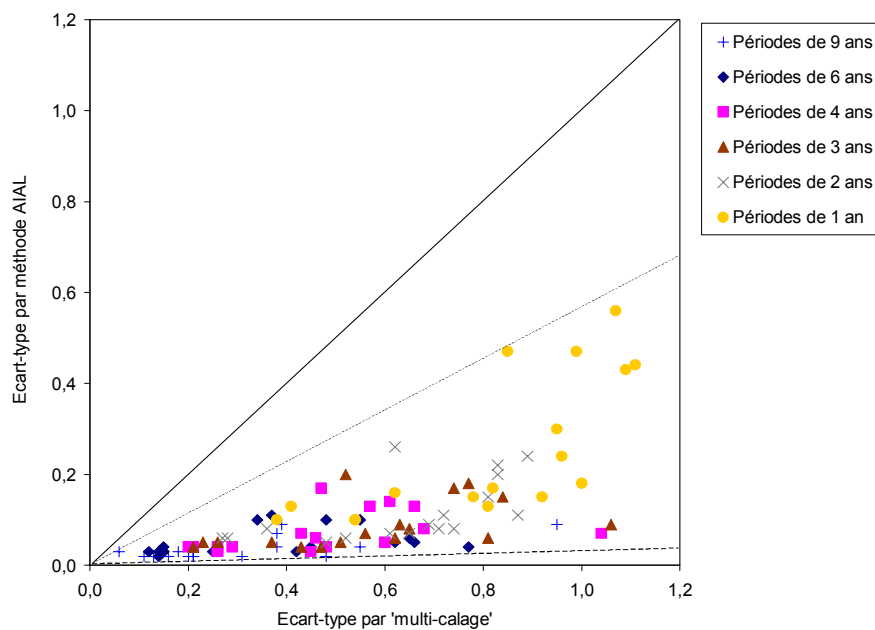


Figure 6.16: Comparaison des écarts-types du paramètre X4 de la structure GR41 fournis par les méthodes AIAL et 'multi-calage'

De manière générale, il y a un fort désaccord entre les résultats des deux méthodes. Les écarts-types fournis par la méthode AIAL sont plus faibles que ceux fournis par la méthode 'multi-calage', comme le montre la Figure 6.16 dans le cas du paramètre X4 de la structure GR41.

La décroissance générale des valeurs des écarts-types lorsque la période augmente (voir par exemple le cas du paramètre X1 à la Figure 6.17) est assez aisément explicable. Le calcul de l'écart-type prend en compte (de façon inversement proportionnelle en $1/\sqrt{n}$) la taille de l'échantillon auquel il se réfère, faisant tendre mathématiquement l'écart-type vers zéro lorsque l'échantillon de données devient grand (Troutman, 1985a). Dans le cas de la méthode AIAL, le calcul prend en compte de la même façon le nombre de pas de temps de la période

considérée, faisant diminuer l'écart-type lorsque la période augmente comme nous pouvons l'observer ici.

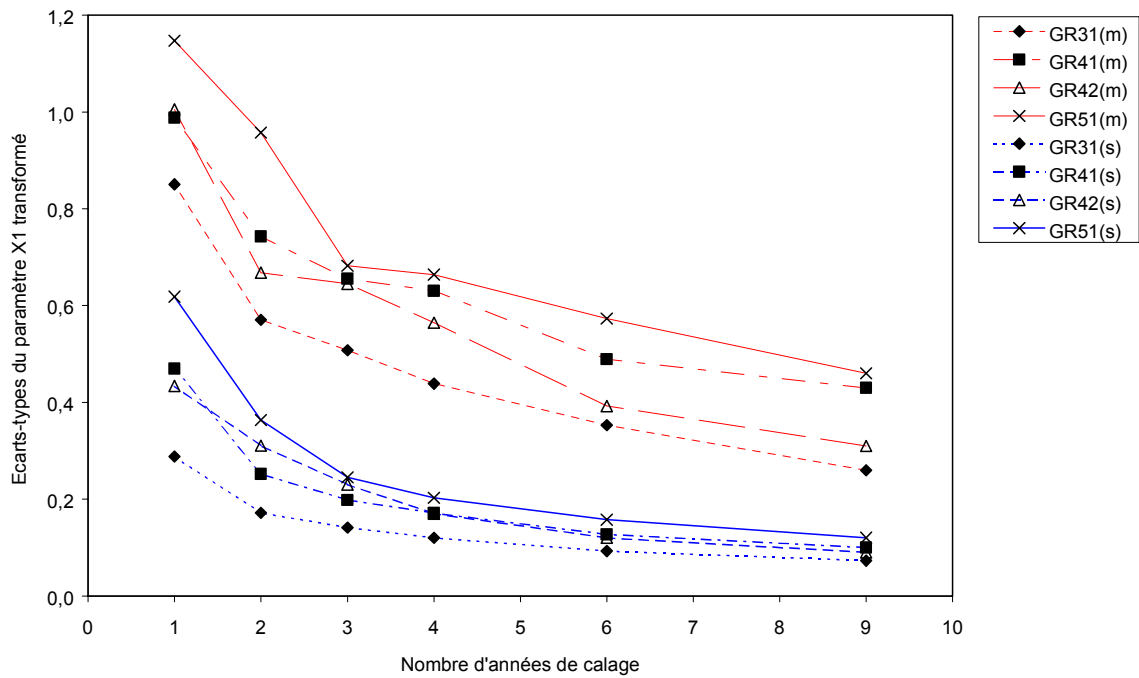


Figure 6.17: Ecarts-types du paramètre X1 fournis par la méthode AIAL (s) et par ‘multi-calage’ (m) en fonction de la longueur de la période de calage

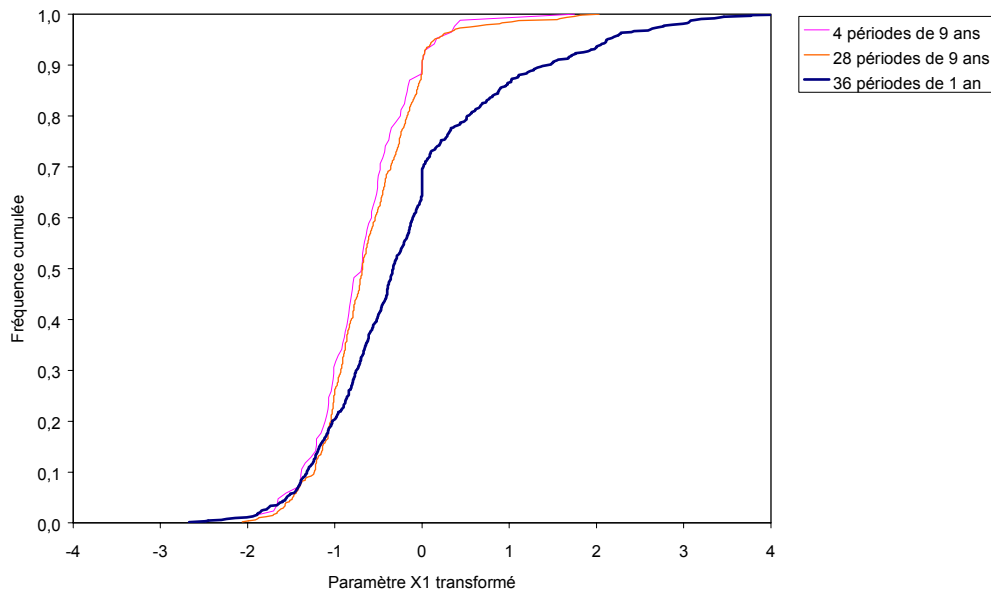


Figure 6.18: Distributions des valeurs du paramètre X1 du modèle GR41 en prenant 4 périodes de 9 ans, 28 périodes de 9 ans ou 36 périodes de 1 an.

Dans le cas de l'approche ‘multi-calage’, on prend en compte le cardinal de l'échantillon des paramètres calés sur les différentes périodes. Ce cardinal diminuant lorsque la longueur de la période augmente (il y a en effet moins de périodes indépendantes), l'écart-type devrait avoir tendance à augmenter avec la longueur de la période. Or ici, cet effet est compensé par le fait que les intervalles de variations des paramètres sont plus grands lorsque la période est courte

(comme cela a été mentionné dans l'analyse des valeurs des paramètres). Ceci est illustré par exemple à la Figure 6.12 pour les paramètres X1 et X4. Dans l'approche 'multi-calage', c'est donc bien la diminution de la variabilité des paramètres qui fait diminuer l'écart-type lorsque la période augmente. Si nous avions eu des nombres équivalents de courtes et longues périodes, l'effet de diminution aurait été amplifié. En effet, l'augmentation du nombre de périodes longues, par exemple, n'augmente pas consécutivement l'intervalle de variation des paramètres, comme le montre la Figure 6.18 pour le paramètre X1 avec 4 ou 28 périodes de 9 ans. Néanmoins le fait que les 4 périodes soient indépendantes et que les 28 périodes ne le soient pas favorise probablement cette ressemblance des distributions.

Même si les raisons paraissent différentes, les deux approches convergent vers une indication de la décroissance des écarts-types lorsque la période augmente. Ceci rejoint les commentaires précédents soulignant le moins grand nombre d'échecs du modèle en contrôle lorsque la période de calage augmente.

Dans le cas du paramètre X1 (voir Figure 6.17) et du paramètre X2, la complexité croissante de la structure du modèle augmente les écarts-types sur les paramètres, c'est-à-dire leur incertitude. Dans le cas du paramètre X3, cet écart-type est toujours très faible et n'évolue pas significativement d'une structure à l'autre.

6.8. Conclusion

Deux méthodes pour estimer les incertitudes sur les paramètres des modèles de simulation des débits ont été comparées:

- la première est une approche classique, l'analyse d'incertitude par approximation linéaire, couramment appliquée en modélisation hydrologique. Elle détermine l'incertitude sur les paramètres à partir de l'estimation de la matrice des variances-covariances des paramètres du modèle;
- la deuxième approche, appelée ici 'multi-calage', détermine l'incertitude sur les paramètres par une analyse statistique d'un échantillon de paramètres calés sur différentes périodes de données.

Ces deux méthodes ont été appliquées à quatre structures de complexité croissante du modèle GR, sur un échantillon de 17 bassins versants. Il a été montré que ces deux méthodes fournissent des résultats très différents quant aux écarts-types sur les paramètres et aux corrélations entre paramètres.

L'analyse d'incertitude par approximation linéaire fournit des écarts-types généralement inférieurs à ceux obtenus par la procédure 'multi-calage'. La raison principale réside dans le fait que l'approche 'multi-calage' prend en compte l'incertitude liée au choix de la période de calage, alors que l'approche classique ne considère qu'un échantillon de données. Cette dernière méthode donnerait donc des valeurs moins réalistes des écarts-types sur les paramètres, et sous-estime certainement l'incertitude qui leur est liée. La variabilité naturelle des données à partir desquelles les paramètres sont calés semble jouer un rôle important dans l'incertitude sur les paramètres du modèle.

Les deux approches fournissent des résultats également différents pour l'estimation des corrélations entre paramètres. L'analyse d'incertitude par approximation linéaire décrit mieux le lien structurel entre paramètres en s'intéressant aux liens entre ceux-ci dans le voisinage de l'optimum et les erreurs du modèle. Les résultats de l'approche 'multi-calage', en

caractérisant les liens existant entre paramètres dans un échantillon de valeurs optimisées, sont très influencés par les caractéristiques et la longueur des périodes de calage.

L'analyse d'incertitude classique semble donc davantage adaptée pour analyser les corrélations structurelles entre paramètres. Elle est donc bien adaptée pour identifier la nature des difficultés de calage dues à la structure du modèle. Au contraire, la méthode 'multi-calage' semble fournir des valeurs d'écart-types plus réalistes que l'analyse de sensibilité, en prenant en compte l'incertitude dans la détermination des paramètres liée au choix, à l'échantillonnage des données de calage. L'approche 'multi-calage' paraît ainsi être un complément utile à l'autre méthode d'analyse d'incertitudes pour estimer l'incertitude sur les paramètres.

L'application de la méthode 'multi-calage' présuppose cependant que de longues chroniques de données soient disponibles pour distinguer un nombre suffisant de périodes de calage.

Cette étude nous a permis par ailleurs d'avoir des éclaircissements sur la structuration interne du modèle GR et sur les liens entre ses paramètres. L'optimisation de la capacité du réservoir de production améliore la qualité des simulations comme nous l'avons vu dans les tests de comparaison, alors que celle du coefficient de partage des écoulements apparaît sans intérêt du point de vue des performances. L'accroissement de la complexité du modèle a pour effet d'augmenter les écart-types sur les paramètres, c'est-à-dire qu'il existe une plus grande incertitude dans la détermination de ces paramètres. Cependant, l'augmentation du nombre de degrés de liberté peut, dans certains cas, réduire les corrélations entre paramètres. En effet, un plus grand nombre de paramètres permet de mieux définir le rôle de chacun d'eux dans le processus de transformation pluie-débit.

Les deux paramètres de rendement n'ont pas paru également sensibles à l'ajustement des bilans en eau sur le bassin. Le paramètre d'échange semble, sur cet échantillon de bassins, être utilisé avec plus de souplesse par le modèle que la capacité du réservoir de production, ce qui se traduit par une plus grande dépendance aux caractéristiques de la période de calage. Nous n'avons pu apporter d'éclaircissement ici sur le rôle de l'utilisation d'une ETP moyenne interannuelle et ses éventuels effets sur la stabilité de la capacité du réservoir de production.

D'après cette étude, il semble que l'allongement de la période de calage permette d'obtenir des paramètres et performances du modèle plus stables. Il faut en fait entendre par cet allongement de la période davantage une prise en compte d'une plus grande variabilité de conditions hydro-climatiques qu'un accroissement du nombre de données. Le paramètre d'échange et la capacité du réservoir de routage ont paru les plus sensibles à cette longueur de la période. Il ne nous est cependant pas possible ici de conclure quant à une longueur de période qui permettrait d'avoir des résultats du modèle relativement insensibles aux données, comme l'ont fait Yapo *et al.* (1996) dans le cas du modèle de Sacramento en préconisant des périodes de calage de 8 ans.

Un des enseignements utiles pour la suite de notre travail est que l'introduction de degrés de liberté supplémentaires n'est pas aisément couronnée de succès ! Les deux paramètres que nous avons choisi d'optimiser (en plus des trois paramètres initiaux de GR3J), solutions simples de complexification du modèle, ne donnent pas entière satisfaction. Nous allons donc voir au prochain chapitre si d'autres modifications peuvent être plus staisfaisantes.

Chapitre 7

Chapitre 7

Tentative d'amélioration du modèle GR3J par une approche empirique

7.1. Introduction

La comparaison de structures de modèles pluie-débit a ouvert quelques perspectives prometteuses pour la mise au point d'un modèle plus performant, en soulignant les forces des différents modèles et la complémentarité existant entre certaines structures. Par ailleurs, les travaux d'analyse d'incertitude réalisés au chapitre précédent tendent à indiquer que si l'optimisation du réservoir de production du modèle GR3J permet un gain substantiel de performance, cette solution de complexification, de même que l'optimisation du coefficient de partage des écoulements, pourraient ne pas être les solutions simples les plus bénéfiques. Ces travaux ont par ailleurs souligné qu'il demeure certains points d'interrogation ou ambiguïtés dans le fonctionnement du modèle et le comportement des paramètres, en particulier au niveau de la fonction d'échange.

Nous présentons ici les résultats de la recherche d'une structure améliorée du modèle simple journalier GR3J, qui a pour but d'essayer d'en améliorer les performances et, nous espérons, la cohérence interne. Ce modèle présente le double avantage de donner des résultats satisfaisants par rapport aux autres modèles testés dans la comparaison et d'être suffisamment simple pour que l'on puisse déterminer le plus judicieusement où introduire des degrés de liberté supplémentaires. Les possibilités d'introduction de modifications deviennent en effet vite beaucoup trop nombreuses dans un modèle déjà complexe pour que la recherche d'une version améliorée soit efficace.

La recherche d'amélioration d'un modèle pluie-débit n'est pas rare en hydrologie. La plupart des modèles existants sont le fruit de modifications/améliorations successives qui ont progressivement changé les caractéristiques de la structure initiale. A titre d'exemple, nous pouvons citer le cas de TOPMODEL, pour lequel Beven (1986) a introduit dans l'indice topographique la prise en compte de la transmissivité des sols; le cas du modèle TANK, pour lequel Sugawara (1995) a proposé une structure améliorée du réservoir de production; le cas du modèle SMAR, pour lequel Tan et O'Connor (1996) ont introduit une fonction d'infiltration améliorant les résultats de la version originale; ou encore le cas du modèle HBV auquel Lindström *et al.* (1997) ont apporté des modifications pour tenir compte de l'altitude et de la végétation et réaliser un meilleur traitement de la neige. Récemment, Jayawardena et Zhou (2000) ont également proposé une paramétrisation différente de la fonction de production du modèle Xinanjiang. Dans tous les cas, ces évolutions de la structure du modèle ont pour but d'améliorer les performances. Cette situation où l'on ne semble pas avoir de

caractéristiques très bien définies des structures des modèles est certainement révélatrice de la difficulté de modéliser la relation pluie-débit et d'un certain état d'ignorance quant aux solutions efficaces à apporter, en l'absence d'une quelconque théorie généralisée sur le fonctionnement d'un bassin versant.

Dans les paragraphes suivants, nous détaillons les différentes phases de la recherche d'une version améliorée du modèle GR3J. La démarche adoptée est résolument empirique, c'est-à-dire que le seul guide et juge dans le choix de modifications a été la qualité des simulations obtenues sur l'échantillon test. Klemeš (1982), dans le débat qui a eu lieu au cours des années 80 sur la vocation de l'hydrologie et la valeur des différentes approches de modélisation, a présenté un point de vue très critique sur la démarche empirique. Il pense notamment que la recherche d'améliorations d'un modèle empirique en l'absence d'informations supplémentaires sur le bassin tend à être scientifiquement stérile et à avoir un faible ratio bénéfice/coût. Nous croyons, au contraire, que cette recherche peut permettre de faire des progrès substantiels sur la qualité de simulations des modèles et que cet enjeu est très important, notamment pour assurer une bonne crédibilité aux modèles dans les applications et questions auxquels ils sont amenés à répondre dans un contexte opérationnel.

Notre démarche va nous conduire à construire des modèles 'hybrides', résultats du couplage d'outils mathématiques provenant de différents modèles, et dont l'adéquation aux bassins versants est plus satisfaisante. Etant donné que nous partons d'une structure bien définie dans laquelle nous essayerons d'introduire des modifications, nous parlerons plutôt de 'versions modifiées' ou 'versions améliorées' dans la suite.

Après un rappel des résultats de la comparaison centrés sur le modèle GR3J, nous présentons brièvement la méthodologie de recherche d'une version améliorée avant d'en donner les résultats. Nous verrons notamment, en écho aux conclusions de la comparaison faisant état d'une fourchette de trois à cinq paramètres optimisables, quel est le degré de complexité le plus approprié, et où, dans la structure du modèle, une modification peut apporter le plus de bénéfices. La limitation de la complexité du modèle sera d'ailleurs un garde-fou à la multiplication des solutions possibles.

7.2. Rappels sur les résultats de la comparaison de modèles

Deux aspects majeurs de l'analyse de la comparaison peuvent être utilisés dans la recherche d'une version améliorée, respectivement les performances obtenues par les différents modèles et la complémentarité entre les structures. Dans ce qui suit, nous donnons tout d'abord un rappel des résultats de la comparaison, en centrant notre attention sur le modèle GR3J et en essayant à partir de cela de distinguer des structures potentiellement utiles dans notre démarche d'amélioration du modèle. Nous utiliserons les résultats obtenus suivant les six critères de performance *CR1* à *CR6* (les critères de Nash sur les débits, sur les racines carrées des débits et sur les logarithmes des débits, le critère Nash-bis, le critère d'erreur absolue et le critère de bilan. Nous garderons cependant à l'esprit que le critère de bilan n'est considéré que comme complément des autres et ne peut être utilisé seul pour juger de la qualité d'un modèle.

7.2.1. Performances et fiabilité des modèles

Les performances sont révélatrices des forces et des faiblesses des différentes structures. Le Tableau 7.1 donne les huit structures qui ont obtenu les meilleures performances moyennes sur l'échantillon test suivant les six critères. Le Tableau 7.2 résume les qualités de fiabilité, en

donnant les structures ayant été classées le plus grand nombre de fois parmi les six meilleures structures sur les 429 bassins.

	1	2	3	4	5	6	7	8
CR1	GR4K	WAGE	TOPM	BOOC	NAM0	GR3J	MHR0	HBV0
CR2	GR4K	GR5J	HBV0	TOPM	GR3J	GEOR	XINA	GR4J
CR3	GR5J	GEOR	TOPM	HBV0	GR4J	GR4K	CREC	XINA
CR4	GR4K	HBV0	GR4J	TOPM	GEOR	MHR0	GRHU	GR3J
CR5	GR4K	GR5J	GEOR	TOPM	GR4J	GR3J	HBV0	BOOC
CR6	MHR0	GR4K	MODG	HBV0	TANK	BOOC	SMAR	GR3J

Tableau 7.1: Structures ayant obtenu les meilleurs résultats moyens suivant les cinq critères de qualité

	1	2	3	4	5	6	7	8
CR1	GR4K	GR4J	GR3J	MHR0	TOPM	GR5J	GRHU	SMAR
CR2	GR4K	GR4J	GR3J	GR5J	TOPM	GEOR	MHR0	HBV0
CR3	GR5J	GR4J	GEOR	GR4K	TOPM	GR3J	HBV0	GRHU
CR4	GR4K	GR4J	GR3J	MHR0	GRHU	GR5J	TOPM	CREC
CR5	GR4K	GR5J	TOPM	GR4J	GR3J	GRHU	GEOR	MHR0
CR6	HAAN	SMAR	HBV0	GR4J	GARD	ARNO	MODA	TANK

Tableau 7.2: Structures classées le plus grand nombre de fois parmi les six meilleures structures

Sur les 38 structures initialement testées, seules quelques-unes apparaissent comme étant à la fois performantes et fiables, au rang desquelles nous pouvons citer principalement les structures TOPM, HBV0, GEOR et celles de la famille GR.

7.2.2. Complémentarité entre modèles

Dans la perspective d'amélioration de structures de modèles, la recherche de complémentarité indique quelles associations peuvent permettre d'obtenir de meilleures performances moyennes sur l'échantillon test. Ainsi l'association de deux structures A et B, en conservant sur chaque bassin les performances de la meilleure d'entre elles, donne un résultat globalement meilleur que pour les structures A et B prises individuellement. La marge d'amélioration ainsi obtenue illustre ce que chacune des deux structures peut apporter à l'autre.

La mise en évidence de complémentarités a été réalisée sur les critères CR1 à CR5 (celle sur le critère de bilan CR6 étant moins intéressante). A chaque fois, on peut construire une matrice de complémentarité, comme nous l'avons présenté au chapitre 5. Celle-ci est obtenue, à partir des performances moyennes sur chaque bassin, en comptant le nombre de bassins où au moins l'une des deux structures de l'association est classée parmi les six structures les plus performantes. En prenant tous les couples de modèles possibles, on peut ainsi construire la matrice complète. Le Tableau 7.3 (a-b-c-d) présente pour chacun des critères les meilleures associations. Comme nous souhaitons travailler à partir de GR3J, nous n'avons retenu que les associations incluant ce modèle.

Association	Nombres de bassins				Association	Nombres de bassins			
	Modèle 1	Modèle 2	Association	Gain		Modèle 1	Modèle 2	Association	Gain
TOPM-GR3J	128	157	244	87	GEOR-GR3J	146	166	273	107
MHR0-GR3J	140	157	241	84	TOPM-GR3J	157	166	270	104
GEOR-GR3J	106	157	240	83	GR5J-GR3J	164	166	262	96
SMAR-GR3J	108	157	234	77	HBV0-GR3J	131	166	259	93
GR4J-GR3J	159	157	233	74	GR4J-GR3J	185	166	250	65
CREC-GR3J	107	157	233	76	CREC-GR3J	107	166	245	79

Nombres de bassins					Nombres de bassins				
Association	Modèle 1	Modèle 2	Association	Gain	Association	Modèle 1	Modèle 2	Association	Gain
GR5J-GR3J	190	135	275	85	GEOR-GR3J	107	160	245	85
GEOR-GR3J	170	135	262	92	TOPM-GR3J	123	160	243	83
TOPM-GR3J	152	135	247	95	GR4J-GR3J	167	160	242	75
HBV0-GR3J	126	135	243	108	SMAR-GR3J	103	160	237	77
GR4J-GR3J	180	135	240	60	MHR0-GR3J	131	160	237	77
MHR0-GR3J	114	135	212	77	CREC-GR3J	108	160	236	76

Nombres de bassins				
Association	Modèle 1	Modèle 2	Association	Gain
TOPM-GR3J	170	152	273	103
GR5J-GR3J	175	152	258	83
GEOR-GR3J	132	152	250	98
HBV0-GR3J	108	152	235	83
CREC-GR3J	100	152	227	75
MHR0-GR3J	117	152	226	74

Tableau 7.3: Meilleures associations de modèles incluant GR3J et nombres de bassins correspondants pour les critères (a) CR1, (b) CR2, (c) CR3, (d) CR4 et (e) CR5.

Comme précédemment dans le cas des performances moyennes, seules quelques structures viennent compléter de façon substantielle les résultats obtenus par GR3J, par exemple TOPM, HBV0, GR5J, GEOR ou CREC.

7.2.3. Sélection des structures et commentaires

Sur les 38 structures testées dans la comparaison, peu apparaissent à la fois performantes et complémentaires de GR3J. En fonction des tableaux précédents, les dix structures suivantes ont été retenues pour la suite de l'analyse:

BOOC, CREC, GEOR, GR4J, GR4K, GR5J, HBV0, SMAR, TOPM et WAGE⁶

Nous retrouvons, parmi ces structures, bon nombre de celles classées parmi les plus performantes sur les différentes classes de bassins constituées lors de l'essai de typologie entre bassins et modèles (voir chapitre 5). Essayons maintenant de mettre en évidence quelles sortes d'outils sont utilisées dans ces structures et de voir si leurs ressemblances ou différences peuvent nous donner des idées de modifications à intégrer dans la structure de GR3J pour en améliorer les performances.

◆ Au niveau de la production:

Seul TOPM utilise un réservoir d'interception proprement dit. Les modèles GR effectuent une neutralisation de la pluie par l'ETP, c'est-à-dire qu'ils utilisent en fait un réservoir d'interception de capacité nulle. Quatre modèles GR4J, GR4K, CREC, HBV0 utilisent un réservoir de production en 'cul-de-sac', sans sortie d'écoulement. Pour les autres, il existe une percolation sortant du réservoir de production. La détermination de l'évapotranspiration réelle à partir de l'ETP utilise des fonctions variées. Outre l'utilisation de réservoirs de suivi d'humidité du sol, certains modèles font appel à d'autres outils intervenant sur le calcul du bilan. La version testée de BOOC utilise un coefficient correcteur de surface et SMAR une correction de l'ETP. Les modèles GEOR et GR mettent en oeuvre des fonctions de pertes ou d'échange en eau.

⁶ Remarquons que la structure MHR0 qui présente les deux qualités de bonnes performances et de complémentarité vis-à-vis de GR3J n'a cependant pas été retenue pour la suite de l'analyse: l'approche originale de discrétisation spatiale conceptuelle du bassin qu'elle adopte, bien qu'efficace, n'est pas aisément exploitable pour explorer des modifications de la structure de GR3J.

De deux à cinq paramètres interviennent au niveau de la production suivant les structures, sans qu'il y ait de lien avec le niveau de performance.

◆ Au niveau du routage:

La plupart des modèles ont deux composantes d'écoulement, sauf les modèles GEOR, HBV0, GR5J et SMAR qui en ont trois.

Lorsqu'il y a détermination d'une pluie nette, le partage de celle-ci en différentes composantes est réalisée par un coefficient fixe dans le cas des modèles GR, un paramètre optimisé dans le cas de SMAR, ou par une fonction paramétrée d'un état du système (niveau dans un réservoir d'humidité) dans le cas de TOPM ou WAGE.

Le routage est assuré par un ou des réservoirs avec des lois de vidange linéaires dans le cas de GEOR, HBV0 et WAGE, linéaires et quadratiques dans le cas de BOOC, CREC, SMAR, en puissance 5 pour les modèles GR4J et GR4K, linéaires et en puissance 5 pour GR5J et exponentielle et quadratique dans le cas de TOPM. Les modèles GR et HBV0 utilisent également des hydrogrammes unitaires qui assurent un délai progressif, les autres modèles utilisant simplement un pur décalage dans le temps.

De deux à six paramètres interviennent au niveau du routage, sans qu'il y ait de lien avec le niveau de performance.

Les modèles GEOR, HBV0, SMAR et WAGE présentent des seuils de fonctionnement (débordement de réservoir ou activation d'une vidange au-dessus d'un seuil). Cependant, il ne nous a pas paru intéressant d'introduire de tels seuils dans les essais d'amélioration: ils compliquent généralement l'optimisation en créant des discontinuités et induisent également des problèmes d'initialisation. Nous préférons donc adopter des types de fonctionnement continus sur les plages de variation des variables internes du modèle.

Suivant ces commentaires, nous allons orienter la recherche d'une version modifiée dans plusieurs directions. Elles correspondent aux différences existant entre GR3J et les dix autres modèles mentionnés précédemment, et peuvent être regroupées dans les catégories suivantes:

1. utilisation de facteurs de correction;
2. modification de la fonction de rendement;
3. modification de la fonction d'échange;
4. modification du partage de la pluie nette après production;
5. introduction d'une percolation du réservoir de production;
6. introduction d'un réservoir de routage supplémentaire;
7. modification de l'hydrogramme unitaire.

Enfin, l'ensemble des paramètres fixes du modèle ont été vérifiés. L'introduction de facteurs de correction n'est envisagée ici qu'à titre de référence. L'utilisation de tels coefficients est en effet délicate et ambiguë car elle se situe à mi-chemin entre le travail de traitement des données et la modélisation pluie-débit proprement dite. Notons par ailleurs que dans les modifications de la fonction de production, nous nous sommes limités à des solutions relativement simples. Nous n'avons pas exploré les possibilités d'introduire une sous-couche au sein du réservoir de production comme c'est le cas dans le modèle GRHU par exemple ou une deuxième variable interne de suivi d'humidité. Ces solutions rendent le fonctionnement du réservoir plus complexe mais restent potentiellement efficaces.

Le modèle GR3J compte un paramètre de rendement et deux paramètres de routage. Les paramètres supplémentaires interviendront dans l'une et/ou l'autre de ces fonctions.

7.3. Méthodologie de recherche

L'éventail de modifications envisageables est très vaste, suivant que celles-ci sont effectuées indépendamment ou combinées. La démarche empirique adoptée est longue puisqu'elle nécessite de procéder par essai-erreur et tâtonnements successifs. Elle s'inscrit dans la continuité de la démarche adoptée jusqu'alors dans le développement des modèles GR, où les évolutions de la structure sont faites en fonction des résultats obtenus sur les échantillons de bassins, c'est-à-dire en fonction de ce que suggèrent les données.

Nous nous sommes fixés, comme cela a été dit précédemment, de partir de la structure de GR3J d'Edijatno *et al.* (1999) (dont les détails sont donnés en Annexe 1), version la plus simple testée dans la comparaison et donnant des performances assez satisfaisantes par rapport à des structures plus complexes. Une attention particulière a été apportée au choix d'un niveau de complexité 'raisonnable' (nombre de paramètres optimisés) des versions modifiées. Les résultats de la comparaison des structures suggèrent que trois à cinq paramètres sont suffisants. Le niveau de complexité a finalement été déterminé en fonction des tests des nouvelles structures.

L'analyse a été réalisée à partir des cinq critères d'appréciation des performances mentionnés précédemment. Elle repose en priorité sur la performance moyenne obtenue en contrôle. Les quantiles (0,3 notamment) des distributions des résultats ont complété cette analyse sur la moyenne. Les versions modifiées ont été évaluées sur l'échantillon global des 429 bassins. Nous demandons également à ces structures d'obtenir simultanément de bonnes performances sur des sous-échantillons de bassins (par exemple bassins français et étrangers), ceci garantissant une meilleure robustesse à la structure et permettant de ne pas privilégier une structure qui obtiendrait de très bons résultats sur un petit groupe de bassins, masquant de moindres performances sur le reste de l'échantillon.

Le choix des structures améliorées est fondé sur ces critères de performance et de robustesse. Etant donné le nombre de versions testées, certaines ont obtenu des résultats très voisins. Le choix a alors été fait en fonction de la simplicité de la structure et de sa lisibilité.

7.4. Vers une nouvelle version ?

Beaucoup de versions ont été testées, comprenant de zéro à six paramètres optimisés et correspondant à de nombreuses combinaisons des différentes options présentées précédemment.

7.4.1. Les versions testées

Les sept grands types de modifications présentés précédemment ont été réalisés successivement puis combinés dans la structure de GR3J. Nous présentons ici brièvement les différentes solutions testées pour chacun d'eux. La Figure 7.1 résume les différentes possibilités de transformations envisagées (symbolisées par des pointillés sur la figure).

1. Facteurs de correction

Les premiers facteurs de corrections utilisés sont de simples coefficients multiplicateurs appliqués aux entrées (correction de pluie, correction d'ETP) ou en sortie au débit (correction de surface). Ils ont généralement la même action sur toutes les variables auxquelles ils s'appliquent, indépendamment de leur magnitude. Une fonction de correction permettant de corriger de façon privilégiée les crues a également été utilisée, ainsi que la fonction proposée par Rakem (1999) qui agit sur la pluie nette après production. Ces solutions ont généralement donné des résultats acceptables, avec un meilleur comportement sur les bassins étrangers que sur les bassins français.

2. Modification de la fonction de production/rendement

Les variantes testées permettent de modifier, à l'aide d'un paramètre, la détermination de l'évaporation réelle et/ou de la séparation de la pluie en fonction du remplissage du réservoir de production. Les différentes fonctions testées donnent des résultats assez similaires à ceux obtenus en utilisant la fonction de rendement initiale et en laissant la capacité du réservoir libre (version GR4K).

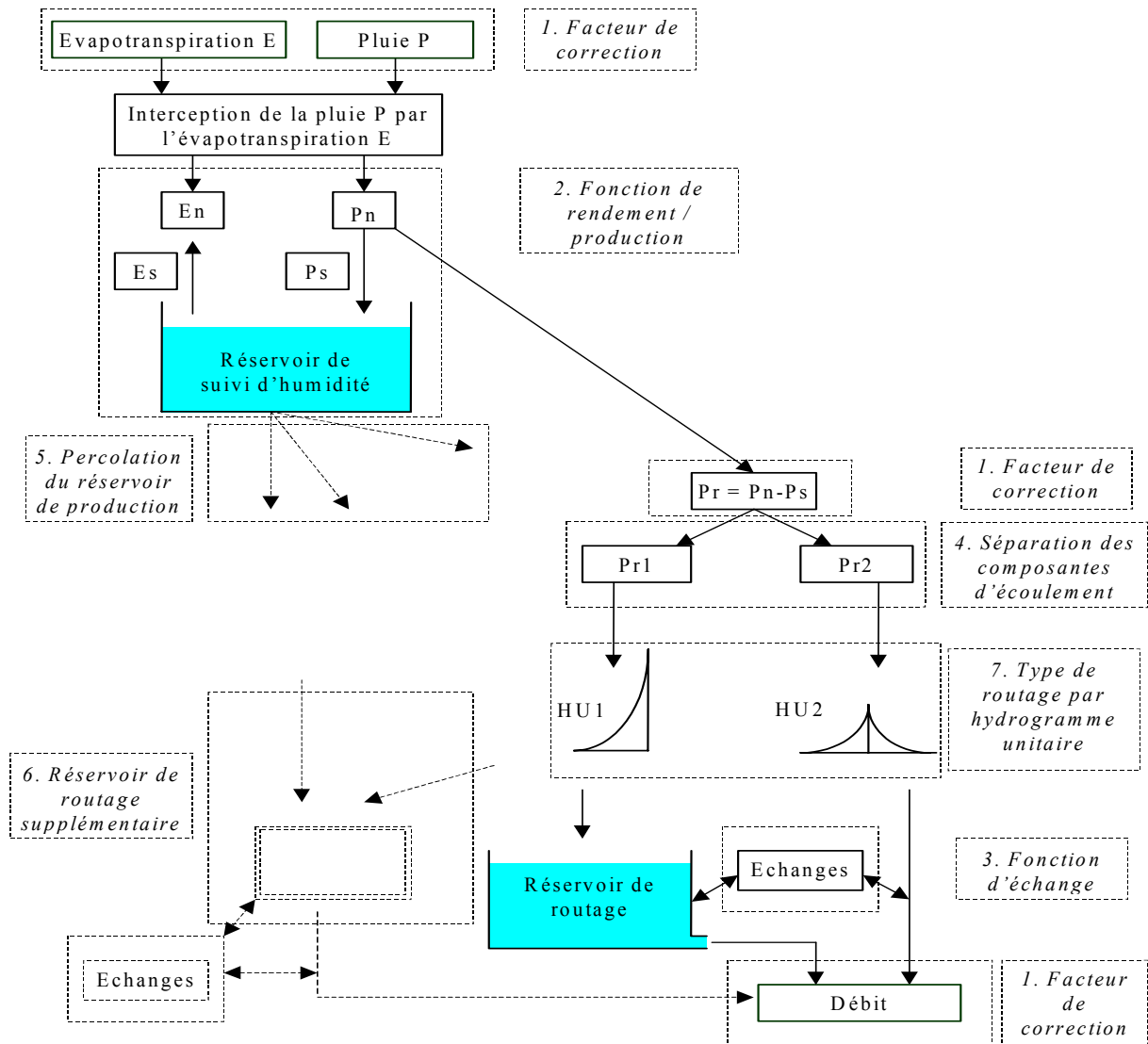


Figure 7.1: Schéma de la structure du modèle GR3J et modifications testées en pointillés

3. *Modification des échanges*

Différentes fonctions d'échange ont été testées, dépendant éventuellement d'autres états du système que le taux de remplissage du réservoir de routage, notamment dans le cas d'un réservoir de routage supplémentaire. Nous avons aussi utilisé des fonctions du type de celle proposée par Nascimento (1995). Il n'y a pas eu d'améliorations substantielles par rapport à la fonction d'échange actuelle, même si plusieurs versions sont apparues équivalentes.

4. *Partage des écoulements*

De nombreuses variantes ont été envisagées: partage paramétré (type IHAC), fonction de partage dépendant d'états du système (type TOPM, WAGE). Globalement les progrès enregistrés au regard du nombre de paramètres supplémentaires utilisés ne sont pas très encourageants, même si quelques versions incluant de telles fonctions de partage (notamment à six paramètres avec un réservoir supplémentaire) ont donné des résultats très satisfaisants.

5. *Percolation du réservoir sol*

S'inspirant notamment des solutions proposées dans GR5J ou GEOR, différentes formulations de percolations ont été essayées, en combinaison avec différentes options de routage de cet écoulement. Des résultats prometteurs ont été obtenus, en assez nette amélioration par rapport à GR3J ou GR4K, notamment dans la restitution des étiages (critère de Nash $CR3$ sur les logarithmes des débits).

6. *Réservoir de routage supplémentaire*

Pour enrichir le routage existant à deux écoulements, une troisième branche, en parallèle des deux premières, a été introduite. Différents types de réservoirs et de lois de vidanges ont été testés, en lien avec différents partages des écoulements en amont. L'introduction d'un réservoir exponentiel de type TOPM a donné de bons résultats.

7. *Utilisation des hydrogrammes*

Des fonctions différentes et des agencements différents des hydrogrammes ont été testés, notamment lors de l'introduction de percolations ou d'une troisième composante d'écoulement, sans fournir globalement de réelles améliorations.

Les autres modifications envisagées, qui relèvent ici plus de l'ajustement que de la modification de structure proprement dite, concernent les valeurs des paramètres fixes de GR. Les valeurs de tous ces paramètres (exposants, constantes...) ont été vérifiées et éventuellement modifiées. Il s'agit par exemple de la proportion du partage des écoulements, de la puissance de la loi de vidange du réservoir de routage, de la puissance de la loi d'échange ou de celle intervenant dans la fonction de l'hydrogramme unitaire.

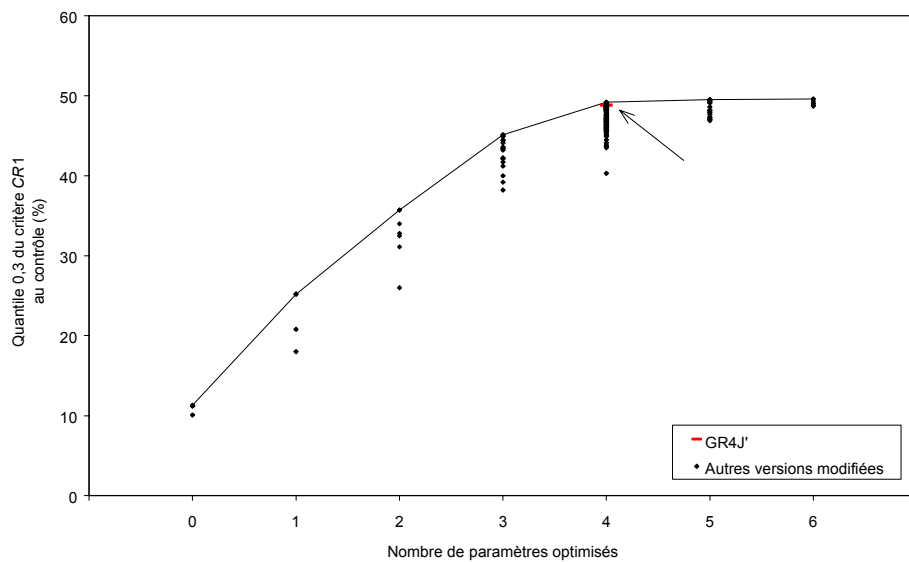
C'est parfois la combinaison de différentes options précédentes qui a permis d'apporter le plus d'amélioration. Au total, 235 structures modifiées ont été testées sur l'échantillon des 429 bassins, dont les trois quarts sont à quatre paramètres. Le temps de calcul nécessaire pour tester une version à quatre paramètres est de l'ordre de 5h30.

7.4.2. Choix d'un niveau de complexité

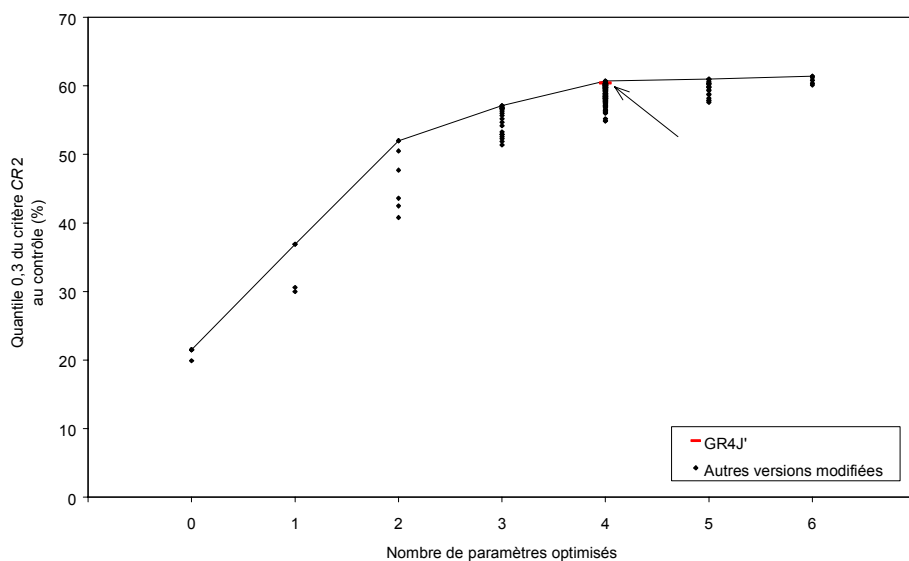
Le choix du nombre de paramètres optimisés s'est fait en fonction des résultats des tests des différentes structures. La Figure 7.2 montre les performances obtenues au contrôle sur l'échantillon test global pour quatre critères ($CR1$: Nash(Q), $CR2$: Nash(\sqrt{Q}), $CR3$: Nash(ln(Q)) et $CR6$: bilan). Nous avons reporté les résultats de toutes les structures modifiées

testées et joint les meilleures performances pour tracer une courbe enveloppe. Les quantiles 0,3 des distributions, ici plus représentatifs, ont été préférés aux moyennes des performances pour ces représentations: dans le cas des structures les plus simples, les plus nombreux contrôles négatifs affectent lourdement la moyenne. Il apparaît que plus le nombre de degrés de liberté augmente, plus le gain relatif de performance enregistré au contrôle diminue. Le gain obtenu en introduisant un degré de liberté supplémentaire reste assez sensible jusqu'à quatre paramètres, mais est faible voire négligeable au delà. Ceci est vrai quel que soit le critère utilisé. Ce palier dans les performances se fait même ressentir dès le troisième paramètre dans le cas du critère de bilan *CR6*.

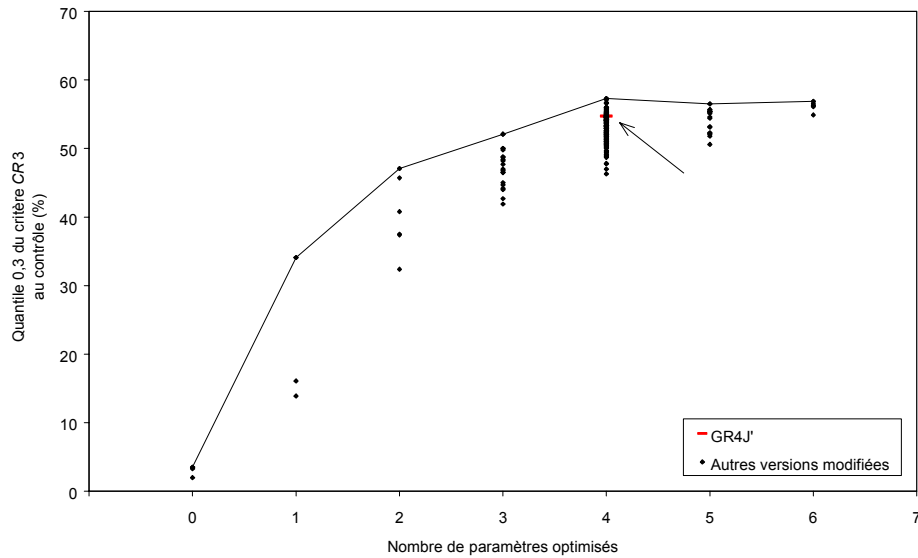
Rejoignant une des conclusions de la comparaison de modèles, il semble donc qu'au delà de quatre paramètres optimisés, l'ajout de degrés de liberté supplémentaires ne semble pas justifiable en terme d'amélioration des performances et au regard des complications induites lors de l'optimisation (compensations entre paramètres, optimum secondaires...).



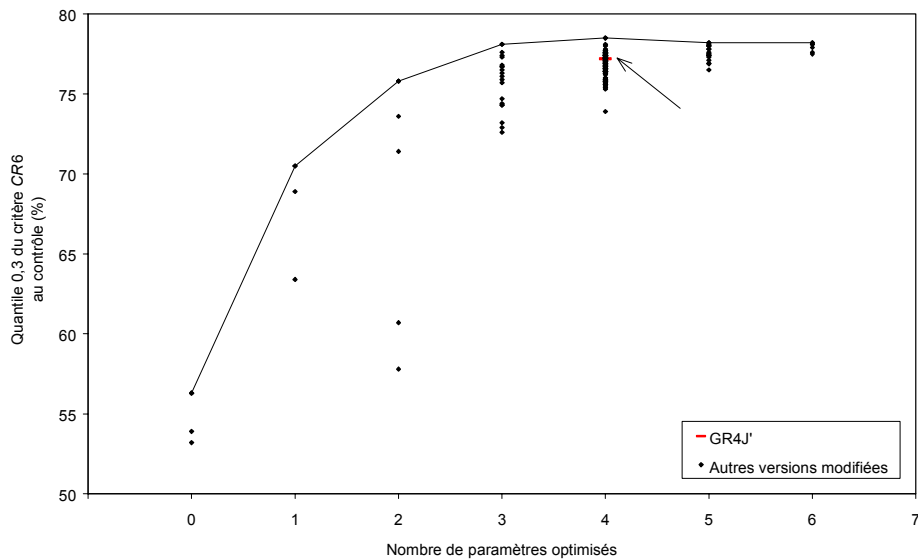
(a)



(b)



(c)



(d)

Figure 7.2: Evolution des performances des structures modifiées en fonction du nombre de paramètres optimisés pour les critères d'évaluation $CR1$, $CR2$, $CR3$ et $CR6$.

Remarquons que les meilleures performances pour chaque paramétrage n'ont pas été obtenues pour tous les critères par la même structure. Ainsi la structure GR4J' identifiée distinctement à la Figure 7.2 (et dont nous reparlons par la suite) apparaît parmi les toutes meilleures structures à quatre paramètres dans le cas des critères $CR1$ et $CR2$, mais elle est relativement moins performante pour les critères $CR3$ et $CR6$.

7.4.3. Description de la structure modifiée retenue

Le choix d'une version définitive dans notre étude a été très délicat, du fait des niveaux de performance assez semblables atteints par plusieurs structures. En effet, même si l'on a gardé un faible niveau de complexité, limitant ainsi les possibilités de modifications de la structure, de nombreuses versions ont donné des résultats très proches. Ceci rejoint les commentaires formulés lors de l'analyse des résultats de la comparaison qui indiquent que des structures de modèles différentes arrivent à donner des résultats équivalents, s'écartant ainsi d'un principe d'unicité en modélisation où à un système donné correspondrait une seule formulation du modèle. La Figure 7.3 montre les performances très rapprochées de 30 structures modifiées,

suivant quatre critères d'évaluation. La différenciation a été tout de même facilitée par l'utilisation de plusieurs critères, qui donnent des clés supplémentaires de critique entre les modèles, ainsi que par l'utilisation des quantiles des distributions.

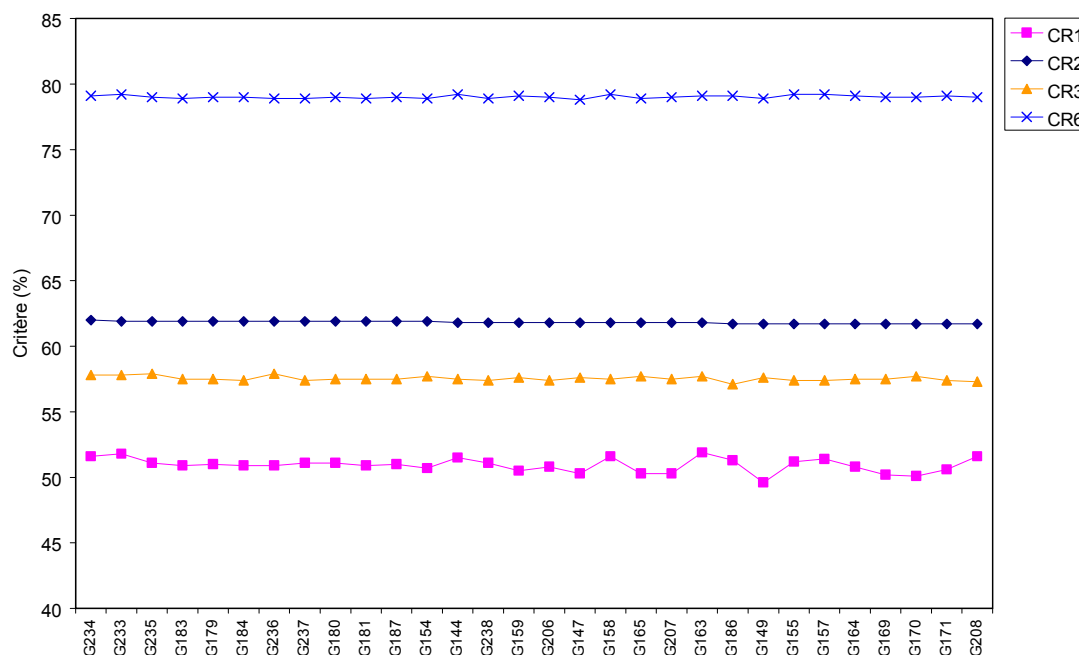


Figure 7.3: Performances moyennes suivant quatre critères de 30 structures modifiées à quatre paramètres, ordonnées par performance décroissante suivant CR2

La version retenue (désignée dans la suite par GR4J'), dont les performances sont marquées distinctement par une flèche à la Figure 7.2, est une version à quatre paramètres. On peut constater qu'elle n'est pas la meilleure pour tous les critères, même si dans chaque cas, elle est proche de la meilleure structure. Elle a été choisie après compromis entre les différents critères de qualité et les objectifs de robustesse énoncés précédemment, notamment sur des sous-échantillons de l'échantillon global.

Dans cette version, de nombreuses valeurs ont été, comme dans les précédentes versions de GR, fixées à des valeurs constantes. Toutes ces valeurs ont été testées et choisies empiriquement de manière à optimiser les résultats.

L'architecture de la version retenue est présentée à la Figure 7.4 et les détails mathématiques sont fournis en Annexe 1. Les quatre modifications introduites par rapport à la version GR3J d'Edijatno *et al.* (1999) sont les suivantes:

- La capacité A du réservoir sol est optimisée (paramètre X4). On introduit ainsi un deuxième paramètre au niveau du rendement (quatrième paramètre du modèle).
- Une percolation du réservoir sol a été introduite. Elle est issue d'une loi de vidange en puissance 5, de la même forme que celle utilisée dans le réservoir de routage. La capacité à un jour dont dépend cette vidange est fixée à $2,25.X4$ (l'origine de cette relation est expliquée dans la suite). La percolation est donc également dépendante du quatrième paramètre du modèle. La vidange ainsi produite est additionnée à la pluie nette avant séparation (90 % - 10 %) des deux composantes d'écoulement.
- La forme de l'hydrogramme unitaire a été légèrement modifiée en passant d'une puissance 3 à une puissance 2,5 pour la courbe en 'S' de l'hydrogramme (forme intégrée de l'hydrogramme).

- d. La dépendance du terme d'échange souterrain vis-à-vis du taux de remplissage du réservoir de routage a été fixée à une puissance 3,5 au lieu d'une puissance 4.

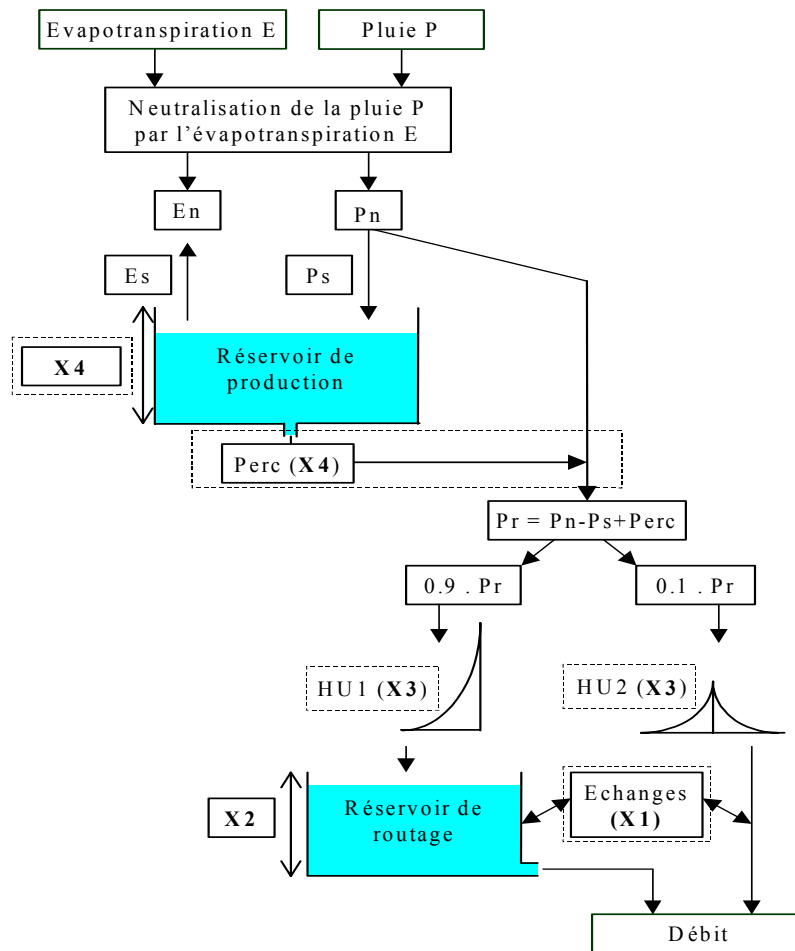


Figure 7.4: Schéma structurel du modèle GR4J', avec en pointillés les modifications introduites par rapport à la version GR3J (X1, X2, X3 et X4 sont les paramètres du modèle)

Ces deux dernières modifications, agissant sur les constantes du modèle permettent des améliorations de 0,2 à 0,6 % en valeur absolue sur les quantiles 0,3 des différents critères, avec cependant une diminution de 0,3 point sur le critère de bilan CR6. Dans le cas de l'hydrogramme, la diminution de la puissance s'est révélée assez bénéfique sur certains bassins étrangers, alors que les résultats sur les bassins français sont restés assez peu sensibles à cette diminution. On a ainsi une forme de l'hydrogramme se rapprochant de l'hydrogramme triangulaire de HBV0 (voir Figure 7.5). En effet, la diminution de la pointe de l'hydrogramme fait que l'on arrive dans certains cas à moins bien restituer les pointes de crues. La variation de la puissance de l'échange permet des améliorations jusqu'à 0,7 point sur le quantile 0,3 du critère CR3 (critère de Nash sur les logarithme des débits). Cette variation induit néanmoins une légère diminution (0,3 point) de la performance moyenne sur le critère CR1 (critère de Nash sur les débits). Les autres paramètres fixes, après tests, ont été conservés.

Remarquons que d'autres structures que celle proposée précédemment ont également paru intéressantes. Parmi les modifications les plus intéressantes, nous pouvons citer:

- l'ajout d'un réservoir exponentiel (tel que celui utilisé dans TOPM, voir Annexe 1) en parallèle de l'actuel réservoir de routage (cette configuration de routage permettant d'obtenir une meilleure restitution des crues);

- l'utilisation de la fonction de production proposée par Nascimento (1995) ou l'utilisation d'une fonction d'échange intermédiaire faisant intervenir le taux de remplissage du réservoir de production dans le cas d'apports en eau (solution intermédiaire entre les fonctions d'échanges de Nascimento, 1995, et Edijatno *et al.*, 1999).

Les versions correspondantes, moins simples et/ou parfois moins performantes suivant certains critères, n'ont pas été retenues ici.

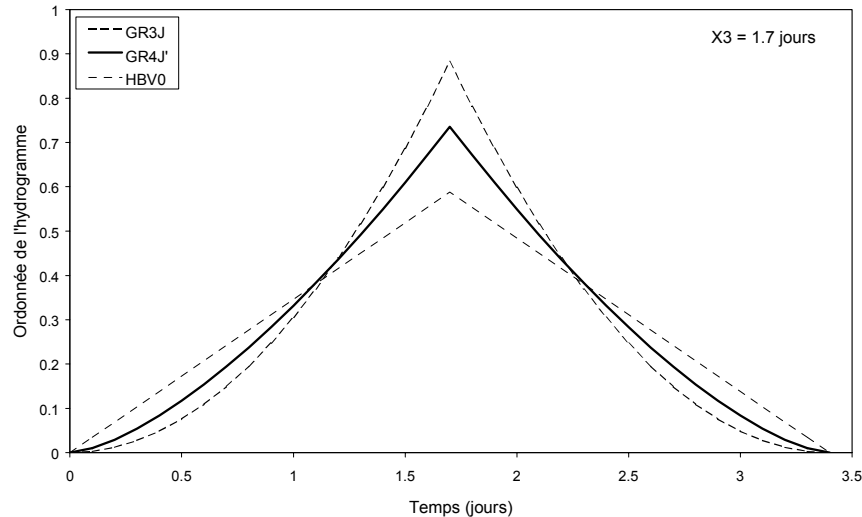


Figure 7.5: Représentation de l'hydrogramme HU2

7.5. Comparaison aux versions antérieures de GR – Critique de la nouvelle structure

7.5.1. Vérification des résultats en utilisant d'autres fonctions objectif

Tous les tests des versions modifiées ont été réalisés en utilisant au calage la même fonction objectif que lors de la comparaison, c'est-à-dire le critère de Nash sur les racines carrées des débits (critère CR2). Nous avons vérifié ici le comportement de la structure choisie en utilisant deux autres fonctions objectif pour le calage du modèle, le critère de Nash sur les débits (critère CR1), qui met davantage l'accent sur les crues et le critère de Nash sur les logarithmes des débits (critère CR3) favorisant davantage les étiages. Ceci avait aussi pour but de vérifier le choix de modifications de paramètres fixes mentionnées précédemment. La Figure 7.6 montre les résultats de sept versions dont les caractéristiques sont décrites dans le Tableau 7.4.

On constate tout d'abord que les performances en contrôle sur un critère particulier sont généralement meilleures lorsque ce même critère a été utilisé pour le calage du modèle, rejoignant ainsi les remarques formulées au chapitre 4. Ceci est vrai excepté pour le critère CR1 (Figure 7.6 a): on obtient en contrôle des résultats de niveau équivalent suivant le critère CR1, que l'on utilise au calage la fonction objectif CR1 ou la fonction objectif CR2. Ceci tendrait à indiquer que le critère de Nash calculé sur les racines carrées des débits est plus robuste que le critère sur les débits pour le calage. On note également que le bilan en eau (critère CR6) est assez nettement moins bon lorsque le calage est effectué avec le critère de Nash calculé sur les logarithmes des débits: en effet, ce critère atténue fortement l'importance des crues, alors que l'essentiel des volumes s'écoule durant ces périodes.

Version	Description	Nombre de paramètres	Exposant de la courbe en S de l'HU	Exposant des échanges
GR3J	Edijatno et al. (1999)	3	3	4
GR4K	idem GR3J + rés. sol optimisé	4	3	4
GR4J'	version hybride avec percolations	4	2,5	3,5
G193	idem GR4J'	4	3	4
G194	idem GR4J'	4	2,5	4
G195	idem GR4J'	4	3	3,5
G192	idem GR4J' + rés. rout. type TOPM	4	2,5	3,5

Tableau 7.4: Description des versions de GR utilisées à titre comparatif

La structure GR4J' obtient généralement de bons résultats par rapport aux autres structures. Dans le cas du critère CR1 cependant, la version G192, avec un réservoir de routage exponentiel, donne des résultats sensiblement meilleurs, ce qui avait déjà été mentionné précédemment. Cependant sur les autres critères, cette version est moins performante. La version G195 obtient aussi de bons résultats sur le critère CR1. Elle ne diffère de GR4J' que par l'exposant de l'hydrogramme. L'exposant 2,5 donne cependant des résultats meilleurs sur les autres critères.

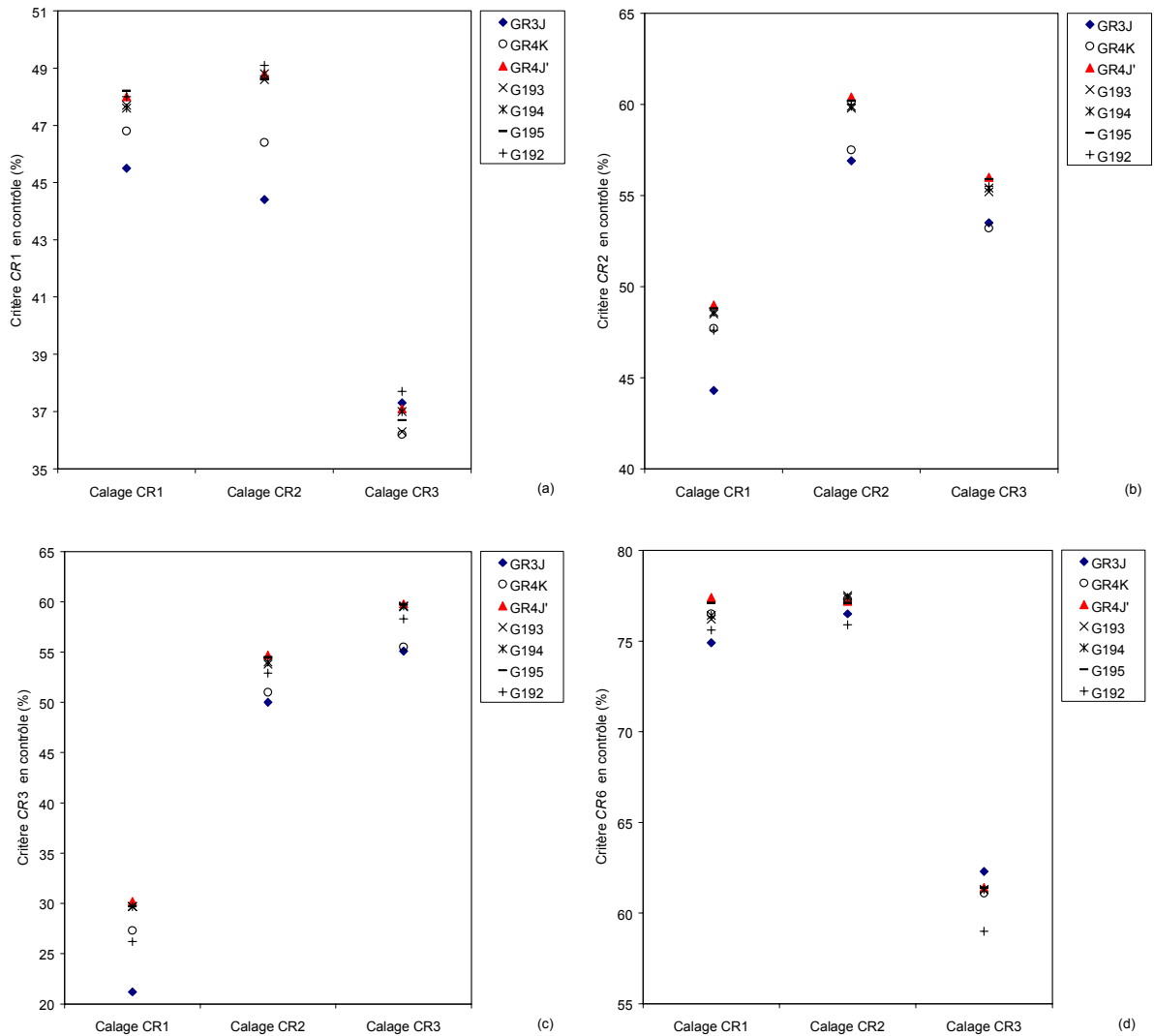


Figure 7.6: Performances comparées en contrôle sur les critères (a) CR1, (b) CR2, (c) CR3 et (d) CR6, de plusieurs structures calées sur différents fonctions objectif

D'autre part, la valeur des paramètres reste en partie dépendante de la fonction objectif utilisée, c'est-à-dire de ce qui est demandé au modèle par l'utilisateur. Ceci est illustré à la Figure 7.7 dans le cas du paramètre X2, capacité du réservoir de routage.

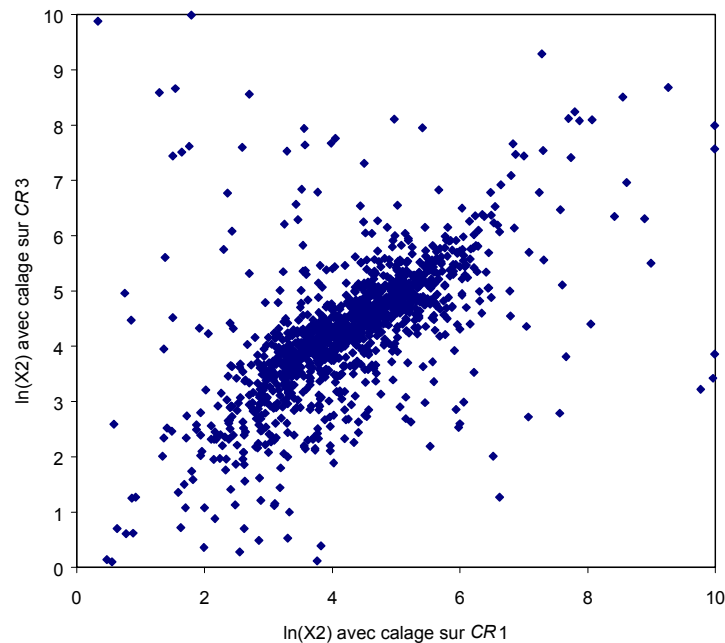


Figure 7.7: Comparaison des valeurs du paramètre X2 obtenues par calage avec les fonctions objectif CR1 et CR3

7.5.2. Version modifiée et versions antérieures du modèle GR journalier

La Figure 7.8 compare les performances (quantiles 0,3) obtenues par la version modifiée GR4J' à celles de cinq autres versions du modèle journalier, à savoir:

- la version GR3J à trois paramètres d'Edijatno (1991), qui fut la première réelle version du modèle journalier;
- la version GR4J à quatre paramètres de Nascimento (1995), qui diffère de la précédente par l'ajout des échanges souterrains;
- la version GR3J à trois paramètres d'Edijatno *et al.* (1999), légèrement différente de celle de Nascimento, et dans laquelle la capacité du réservoir sol a été fixée à 330 mm;
- la version GR4K testée pendant la comparaison, version identique au modèle GR3J d'Edijatno *et al.* (1999) dans lequel la capacité du réservoir sol n'est plus fixe mais optimisée (quatrième paramètre);
- la version GR4J à quatre paramètres de Rakem (1999), qui est une version substantiellement modifiée de la version de Nascimento, proposée à partir d'une reformulation mathématique du modèle.

Les détails de ces différentes versions sont fournis en Annexe 1.

Sur l'échantillon total (Figure 7.8a), sur l'échantillon de bassins français (Figure 7.8b) ou sur l'échantillon de bassins étrangers (Figure 7.8c), la structure GR4J' fournit des améliorations des performances par rapport aux versions antérieures de modèle GR. Au niveau du critère de bilan (CR6), d'autres structures obtiennent cependant d'aussi bons résultats, voire légèrement meilleurs.

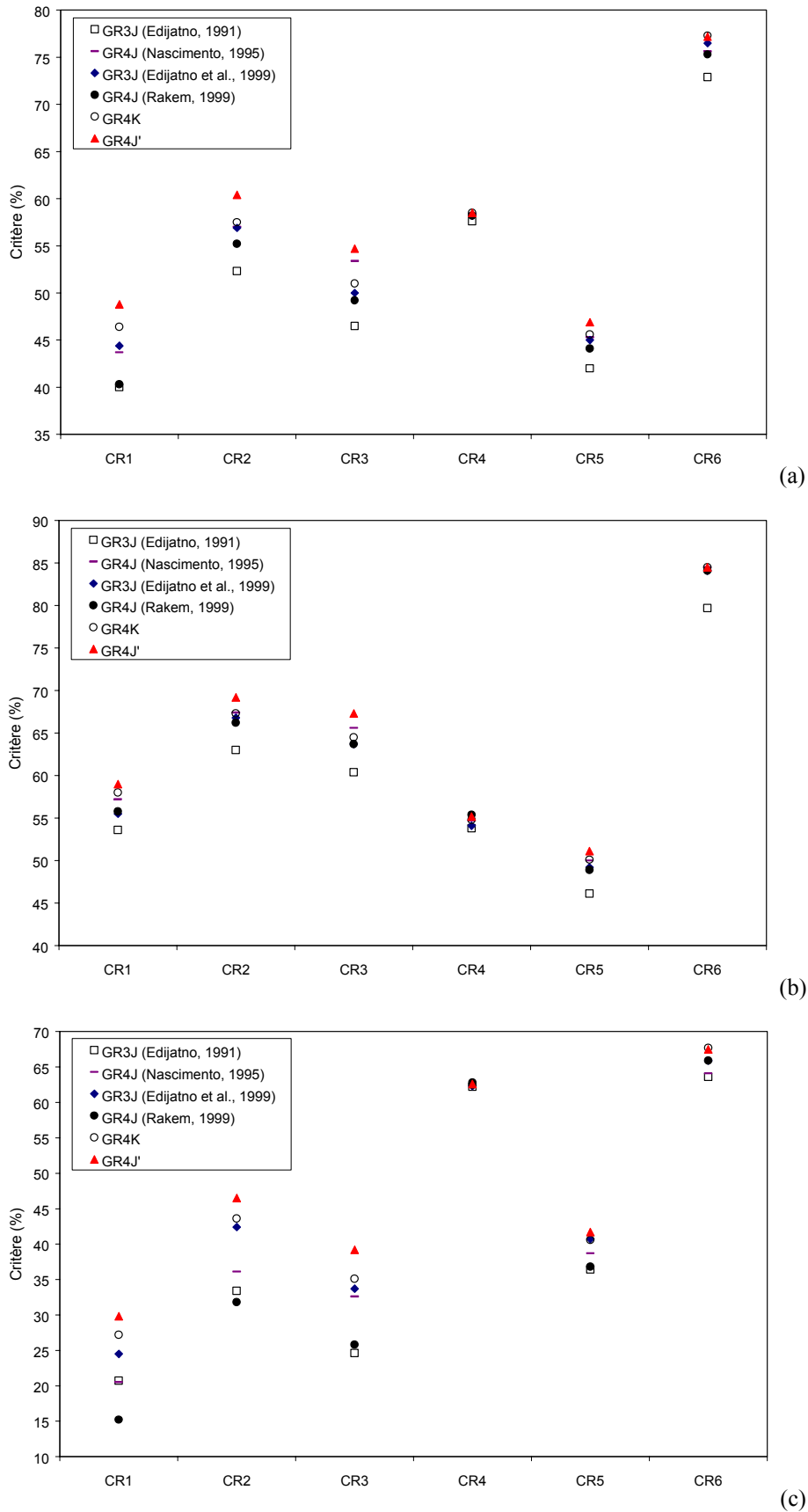


Figure 7.8: Performances comparées (quantiles 0,3) la version hybride GR4J' et de plusieurs versions antérieures du modèle GR journalier à trois ou quatre paramètres sur (a) l'échantillon total de bassins, (b) sur l'échantillon de bassins français et (c) sur l'échantillon de bassins étrangers

Remarquons que la version du modèle à quatre paramètres récemment proposée par Rakem (1999) ne paraît pas ici entièrement satisfaisante par rapport à la version proposée par Nascimento (1995). Nous pouvons avancer une hypothèse de la cause de ces performances relativement moins bonnes. Elle pourrait provenir des modifications introduites dans le module de routage. Rakem (1999) a travaillé, pour proposer une nouvelle version du modèle, sur un échantillon de 60 bassins aux superficies relativement homogènes, peu dispersées autour de la valeur médiane (voir Figure 7.9 et Tableau 7.5). Cette caractéristique de son échantillon de données lui a permis de fixer les ordonnées des hydrogrammes unitaires (dont le temps de base est relativement bien lié à la superficie du bassin). Cette modification pourrait être à la source de difficultés de cette version du modèle sur les bassins les plus petits ou les plus grands de notre échantillon. Rappelons cependant que l'objectif du travail de Rakem (1999) n'était pas d'améliorer le modèle mais d'en proposer simplement une reformulation mathématique.

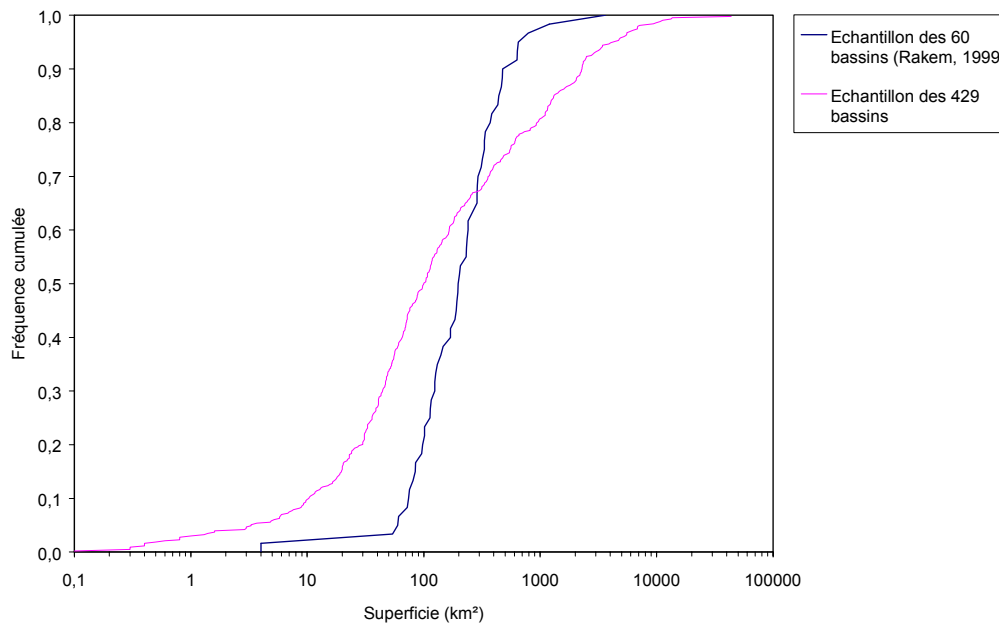


Figure 7.9: Comparaison des distributions des superficies des 60 bassins de l'échantillon de Rakem (1999) et de l'échantillon des 429 bassins utilisés ici.

	Echantillon des 60 bassins	Echantillon des 429 bassins
Superficie du quantile 0,1 (km ²)	74	10
Superficie médiane (km ²)	200	100
Superficie du quantile 0,9 (km ²)	475	2300

Tableau 7.5: Caractéristiques des distributions des superficies des deux échantillons de bassins

Si les performances sont en moyenne meilleures, tous les tests en contrôle ne sont pas systématiquement améliorés, comme le montre la Figure 7.10 (a) sur laquelle sont comparées les performances de GR3J (Edijatno *et al.*, 1999) et GR4J' obtenues en contrôle dans le cas du critère CR1.

Néanmoins, lorsque l'on s'intéresse aux performances moyennes par bassins données aux Figure 7.10 (b) et (c), on s'aperçoit que pour la majorité des bassins, les performances sont améliorées. Dans le cas du critère CR1, les performances diminuent pour 122 bassins (28% de

l'échantillon), cette diminution n'étant supérieure à 2 points que pour 49 d'entre eux (11 % de l'échantillon) et supérieure à 5 points pour 31 bassins (7 % de l'échantillon). Sur ces 49 bassins, la performance de GR3J suivant ce critère n'excédait 60% que pour 11 bassins. Les diminutions notables de performances enregistrées concernent donc principalement des bassins où GR3J avait initialement de grosses difficultés. On retrouve des résultats similaires pour le critère *CR3*.

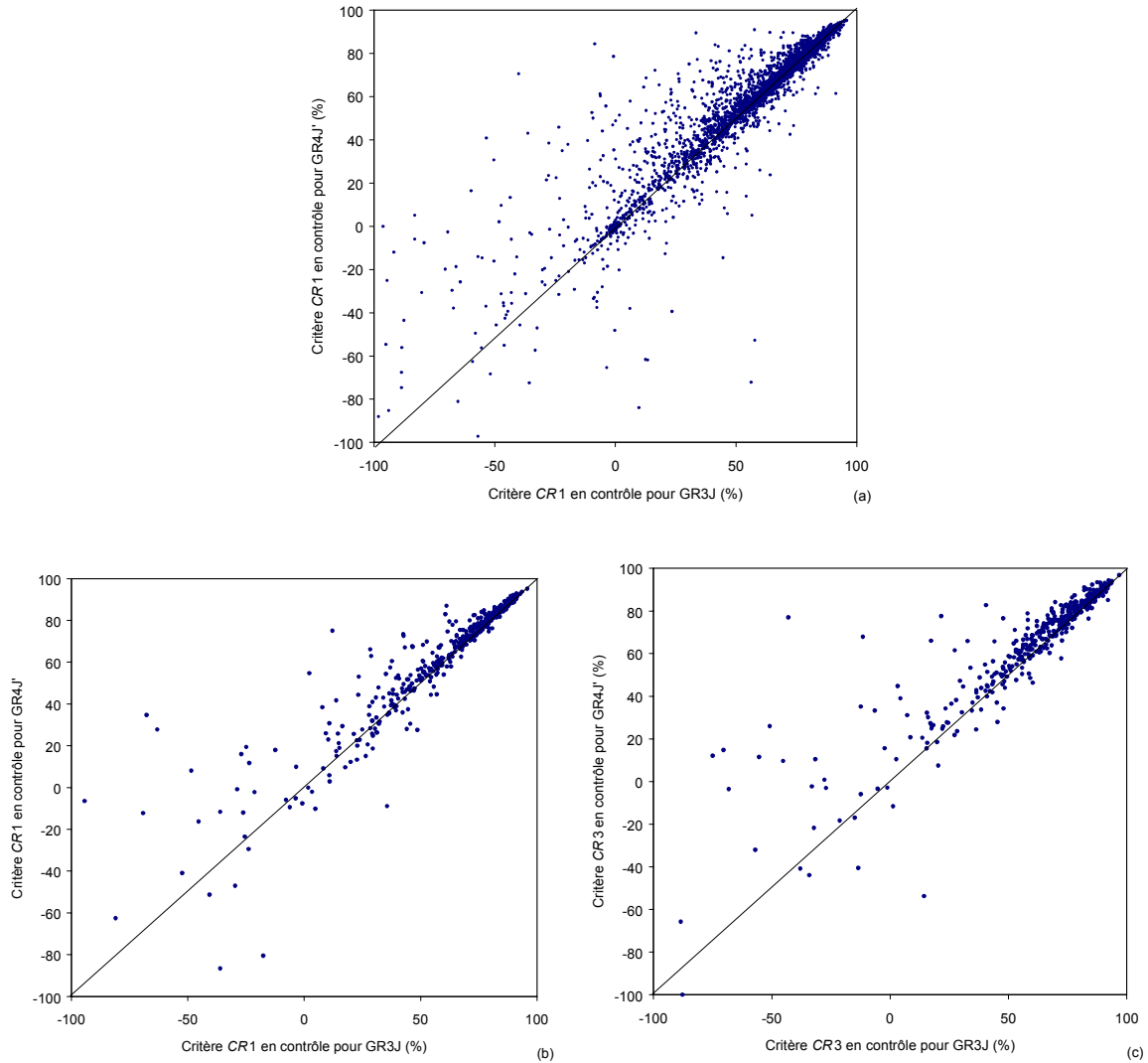


Figure 7.10: Comparaison des performances de GR3J et GR4J' (a) sur tous les contrôles pour le critère *CR1*, sur les performances moyennes par bassin (b) sur le critère *CR1* et (c) sur le critère *CR3*

7.5.3. Comparaison aux autres structures testées

Pour évaluer les résultats de la nouvelle structure, nous avons repris le graphique des distributions de résultats de la Figure 5.14 du chapitre 5. La Figure 7.11 compare, de la même façon, les résultats obtenus par la nouvelle structure avec ceux de différents modèles. On constate, suivant ce critère, que des progrès significatifs ont pu être accomplis par rapport à la meilleure structure testée lors de la comparaison. GR4J' obtient des résultats presque aussi bons que l'association de structures, prouvant ainsi que les efforts d'amélioration consentis ont en partie été fructueux.

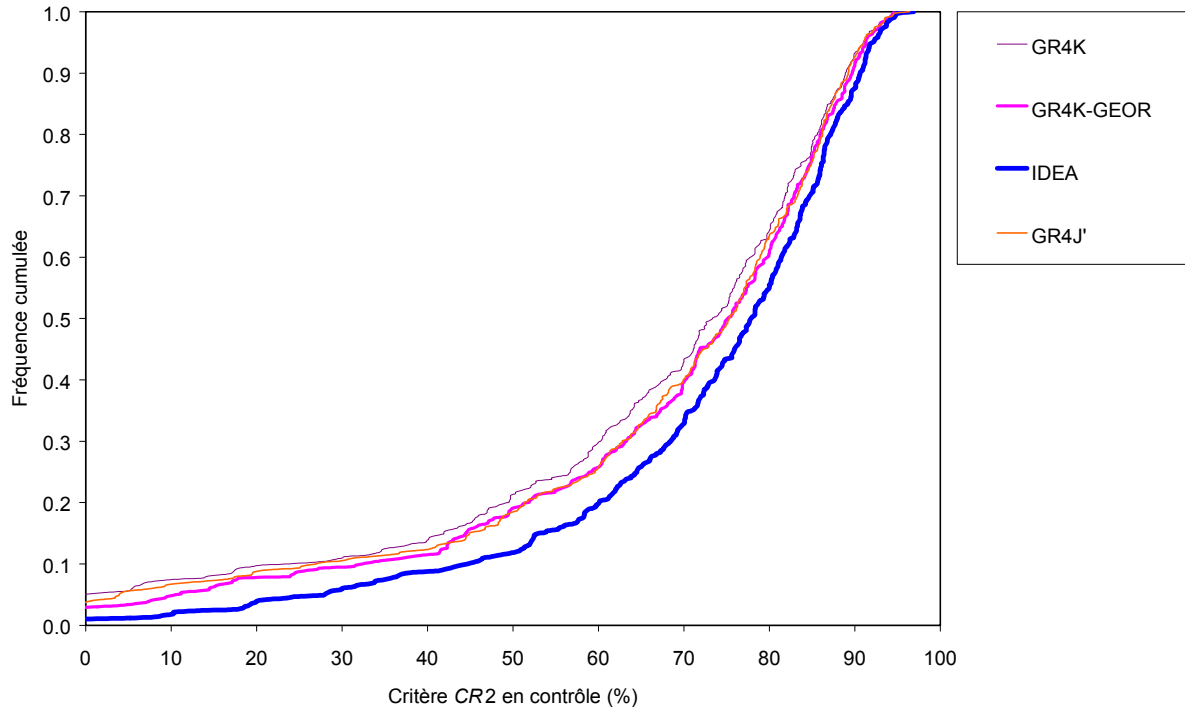
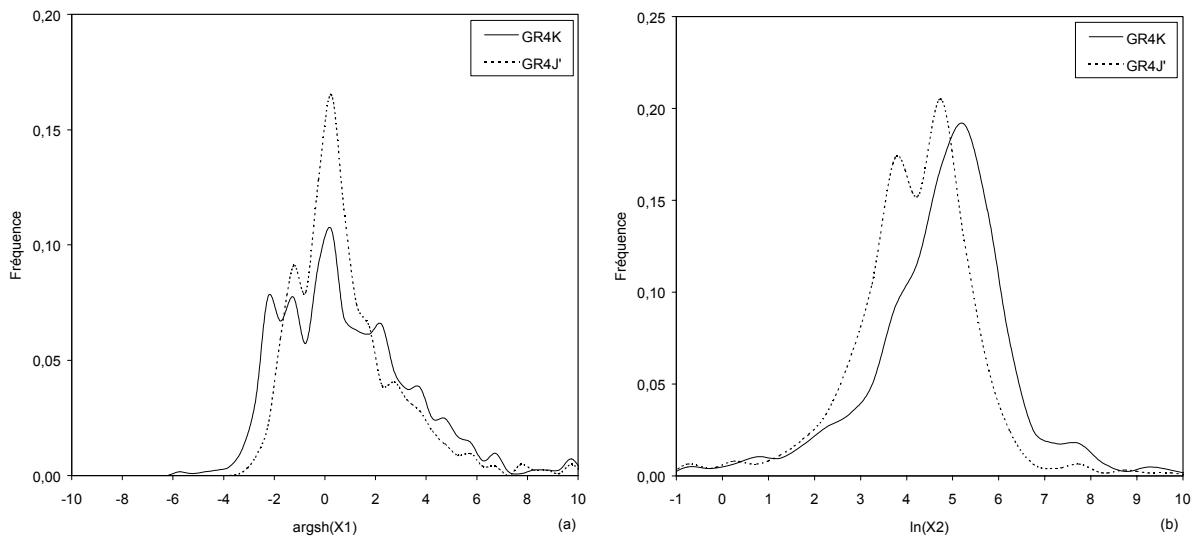


Figure 7.11: Distribution des résultats moyens par bassins suivant le critère $CR2$ pour les modèles GR4K, une association de modèles, le modèle IDEA et le modèle GR4J'

7.5.4. Stabilité des paramètres

Les performances, globalement meilleures pour la nouvelle structure que pour GR3J, peuvent être analysées sous l'aspect de la stabilité des paramètres. La Figure 7.12 présente les densités de probabilité obtenues pour les valeurs des paramètres $X1$ (paramètre d'échange), $X2$ (capacité du réservoir de routage), $X3$ (temps de base de l'hydrogramme) et $X4$ (capacité du réservoir de production) pour le modèle hybride GR4J' et pour le modèle GR4K (version à quatre paramètres de GR3J). Rappelons que dans le cas de la version modifiée, le paramètre $X4$ intervient également dans la vidange du réservoir de production.



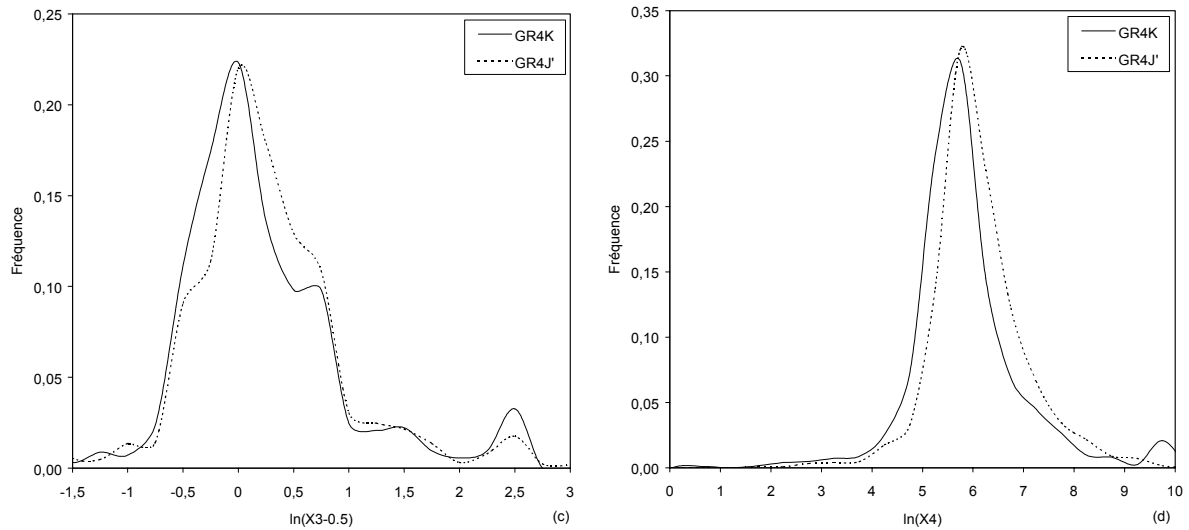


Figure 7.12: Comparaison des fonctions de répartition des quatre paramètres des versions GR4K et GR4J'

D'une manière générale, il semble que l'on ait dans le cas du modèle GR4J', moins de valeurs 'extrêmes' des paramètres (valeurs optimisées aux bornes du domaine de variation) que pour GR4K, en particulier dans le cas du paramètre X4. Ces valeurs extrêmes correspondent à des types de fonctionnement particuliers du modèle (par exemple capacité très petite ou très grande de réservoirs), souvent assez délicats à utiliser. Le nombre moins important de ces valeurs extrêmes pour le modèle GR4J' est certainement en relation avec la diminution du nombre d'échecs du modèle en contrôle. Par ailleurs, les répartition des paramètres sont légèrement différentes pour les deux modèles. Ainsi, on a dans GR4J' des valeurs de X1 davantage resserrées autour de la valeur médiane, des valeurs de X2 en moyenne moins élevées et des valeurs médianes de X3 et X4 légèrement supérieures.

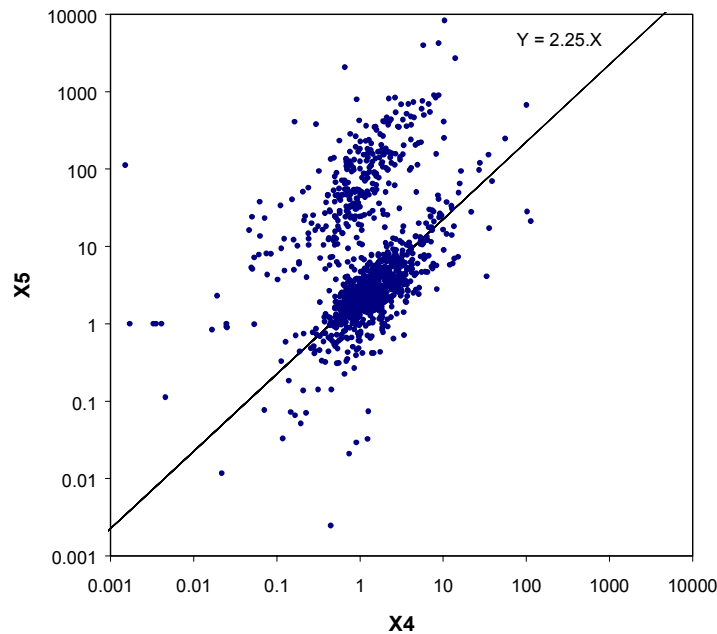


Figure 7.13: Comparaison des valeurs de la capacité du réservoir de production (X4) et de la capacité à un jour (X5) intervenant dans le calcul de la percolation optimisées indépendamment

La Figure 7.13 illustre le lien existant entre la capacité du réservoir de production (X_4) et la capacité à un jour⁷ (X_5) intervenant dans le calcul de la percolation lorsque ces deux paramètres sont optimisés indépendamment (version à cinq paramètres). La relation linéaire choisie dans le modèle GR4J' entre ces deux paramètres ($X_5=2,25.X_4$) semble assez bien vérifiée pour la majorité des couples de paramètres. Cependant, on constate qu'une relation différente existe dans le cas d'un petit nombre de paramètres. La contrainte imposée par cette relation dans la structure hybride GR4J' diminue un peu les performances, mais cette diminution ne justifie cependant pas l'introduction d'un cinquième degré de liberté.

7.6. Performance de la nouvelle structure sur des classes de bassins

A la suite de notre tentative de mise en évidence d'une typologie bassins-modèles au chapitre 5, nous avons essayé d'évaluer ici les performances de la nouvelle structure sur les différentes classes de bassins. La Figure 7.14 montre les performances (quantile 0,3) de toutes les structures modifiées à quatre paramètres sur les cinq sous-échantillons de bassins versants qui avaient été constitués lors de la classification des bassins. A l'image de ce qui avait été montré dans ces précédents travaux, il existe ici, sur chacun des sous-échantillons de bassins, des structures plus performantes que la version GR4J' retenue.

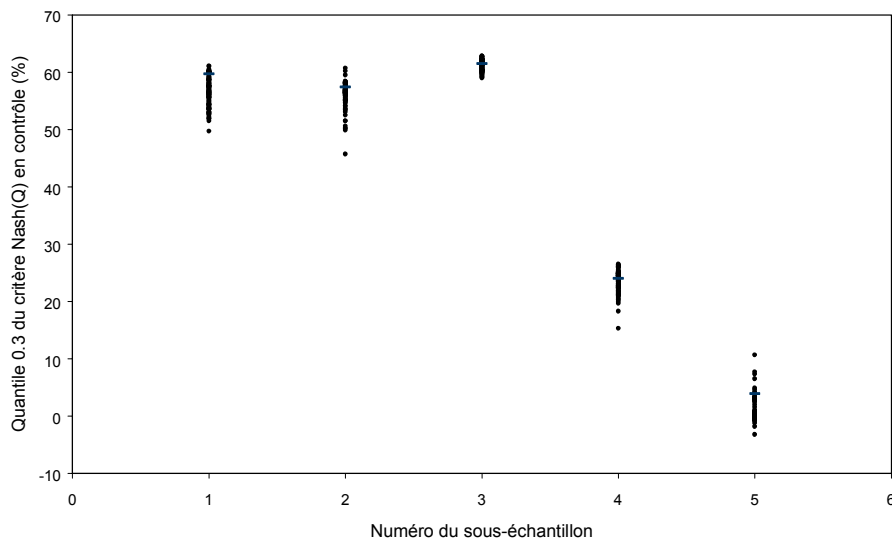


Figure 7.14: Performances des versions modifiées sur les cinq sous-échantillons de bassins (avec signe distinctif pour GR4J')

Sur le sous-échantillon 1, c'est la structure G192 (que nous avons déjà mentionnée) avec réservoir de routage exponentiel en parallèle qui donne les meilleurs résultats. Il est intéressant de constater que c'est TOPM qui donnait les résultats les plus satisfaisants sur ce sous-échantillon. Sur la classe de bassins 5, il a également été constaté que l'utilisation de facteurs de corrections montraient les meilleurs résultats, ceci rejoignant le fait que BOOC ou SIXP, dont les versions testées utilisent toutes deux des facteurs correctifs, obtenaient les résultats les plus satisfaisants sur ce sous-échantillon. Ces résultats tendraient à montrer qu'il existe certains outils mathématiques qui correspondraient mieux à certains types de bassins.

⁷ Ici la capacité à un jour est le contenu qui reste dans le réservoir le jour $j+1$ quand le jour j une entrée de taille infinie s'est produite

Cette correspondance, nette pour les deux classes 1 et 5, entre modèles originaux et versions modifiées, n'a pas pu être mise en évidence aussi clairement dans le cas des autres classes de bassins.

Cependant, il apparaît que dans certains cas, cette adéquation d'outils ne soit pas unique, c'est-à-dire qu'à chaque type de bassins versants pourraient correspondre plusieurs outils obtenant des performances similaires. Ainsi dans le cas de la classe 2, une modification de la fonction de rendement obtient d'aussi bons résultats que l'utilisation d'un coefficient de correction de la surface.

Le Tableau 7.6 présente les résultats moyens obtenus par la structure GR4J' et un modèle composite correspondant au regroupement des meilleurs structures modifiées sur chacune des classes. Ceci correspond à l'association de quatre modèles, l'un d'entre eux s'étant révélé le meilleur sur deux classes. Le passage d'une seule structure (GR4J') à quatre se traduit par un gain substantiel de performances dans le cas du critère CR1. Néanmoins, le gain sur les autres critères reste faible ou nul.

	CR1	CR2	CR3	CR4	CR5	CR6
GR4J'	56,1	65,3	61,4	63,8	54,0	82,1
Modèle composite	57,5	65,2	61,8	63,5	53,3	82,9

Tableau 7.6: Performances moyennes par bassin du modèle modifié et du modèle composite

Comme nous l'avons vu au chapitre 5, la complémentarité entre structures par application de différents modèles en fonction du type de bassin étudié a quelques limites, probablement en raison des incertitudes sur la détermination de groupements de bassins entièrement satisfaisants. La difficulté d'affecter avec certitude un bassin pris au hasard à l'une des classes de bassins limite la fiabilité d'une telle approche de modélisation.

7.7. Conclusion

Nous avons présenté ici la méthodologie et les résultats qui nous ont conduits à la proposition d'une nouvelle structure du modèle du Génie Rural fonctionnant au pas de temps journalier. Notre approche s'est appuyée sur les complémentarités mises en évidence au cours du travail comparatif. De nombreuses modifications de la structure à trois paramètres ont ainsi pu être testées, en adaptant des formulations issues d'autres structures. Un niveau de faible complexité, suggéré par les résultats de notre approche, a été adopté. Il correspond à quatre paramètres optimisés. Les modifications majeures par rapport à la version GR3J d'Edijatno *et al.* (1999) à trois paramètres sont l'introduction d'une percolation issue du réservoir de production et l'optimisation conjointe du paramètre de vidange de cette percolation et de la capacité du réservoir de production. Cette version apporte des améliorations sensibles des résultats, notamment sur la simulation des étiages, et elle est plus robuste que les versions antérieures du modèle GR journalier sur notre échantillon de 429 bassins. A niveau de complexité équivalent, on préférera donc la version GR4J'.

Cependant, même si les progrès réalisés permettent d'accroître la fiabilité du modèle, rien ne garantit que cette version corresponde à l'utilisation optimale de quatre degrés de liberté laissés au modèle. L'introduction d'une fonction de percolation compense peut-être certaines maladresses de la formulation. Le test de nombreuses versions modifiées a permis de montrer qu'il existe plusieurs structures donnant des résultats très proches. Le test du modèle à d'autres pas de temps permettrait peut-être de définir de façon plus sûre certains aspects de la structure. La difficulté de choisir une formulation totalement satisfaisante pour la fonction

d'échange laisse d'ailleurs encore quelques points d'interrogation quant à la forme à donner à cette fonction, qui paraît néanmoins indispensable pour l'application du modèle aux bassins intermittents comme l'a montré Nascimento (1995).

Les progrès réalisés permettent d'obtenir des résultats presque aussi satisfaisants que l'association de deux structures de modèles paraissant les plus complémentaires. Il subsiste cependant un écart de performance important entre cette nouvelle version et le modèle 'idéal' hypothétique qui rassemble les qualités des 38 structures testées dans la comparaison.

Chapitre 8

Chapitre 8

Utilisation dans un modèle hydrologique de l'information contenue dans des descripteurs du bassin versant

8.1. Introduction

La simulation des débits sur des bassins non jaugés était un des objectifs du développement de modèles dont les paramètres peuvent être mesurés sur le terrain. Ainsi, de tels modèles, en utilisant directement l'information contenue dans des descripteurs du bassin, étaient censés être applicables avec fiabilité dans des cas où l'on ne disposait pas de données hydrométriques. Parmi ces modèles, ceux fondés sur la physique utilisent des paramètres généralement mesurables, et leur détermination ne nécessite donc pas en théorie (rarement en pratique) de processus d'optimisation mathématique. Adoptant une approche différente, les modèles conceptuels ou empiriques globaux, qui ne cherchent pas à représenter la complexité et l'hétérogénéité du bassin, dépendent de paramètres qui n'ont généralement pas de correspondant direct sur le terrain a priori. Leur utilisation sur des bassins non jaugés présuppose donc qu'un lien puisse être établi entre la valeur de ces paramètres et des descripteurs mesurables du bassin. Cette démarche est désignée par le terme de régionalisation des paramètres en modélisation hydrologique pluie-débit. Il s'agit en fait d'un type particulier d'application du modèle.

Dans ce chapitre, nous nous proposons de tester une méthodologie améliorée de recherche de liens entre paramètres du modèle et descripteurs du bassin, en partant d'une analyse des approches classiques de régionalisation qui reposent la plupart du temps sur la seule utilisation des régressions multiples. Il s'agit en fait de voir quelle information maximale il est possible d'extraire des descripteurs choisis, l'efficacité étant ici quantifiée par le taux de succès du modèle. Nous utilisons pour cela la version à quatre paramètres du modèle GR proposée au chapitre précédent et des descripteurs climatiques simples du bassin, dans un esprit assez proche des travaux réalisés par Edijatno (1991) ou Makhlouf (1994) avec cette même famille de modèles.

L'objectif principal de ce travail est bien de tester l'efficacité de la méthodologie proposée, l'explication proprement dite des paramètres étant ici limitée par le fait que nous ne disposons pour cette étude que de descripteurs climatiques et non de descripteurs physiographiques qui permettraient peut-être de mieux décrire les spécificités des bassins.

8.2. Approches de régionalisation

Le besoin de connaître les débits sur des sites non jaugés correspond souvent à des préoccupations d'ordre opérationnel: dans bon nombre de projets d'ingénierie par exemple, on ne dispose pas d'une station hydrométrique et d'enregistrements des débits permettant d'évaluer des crues de projet. Le dimensionnement d'ouvrages devient alors sujet aux incertitudes liées aux méthodes d'approximation ou d'interpolation utilisées pour estimer ces débits caractéristiques. La réponse du bassin versant à des événements pluvieux étant en partie le résultat des caractéristiques de celui-ci, il a paru tout à fait naturel de relier les valeurs de certaines variables hydrologiques (crue de période de retour donnée par exemple) à des descripteurs du bassin tels que pédologie, géologie, physiographie ou couvert végétal (voir notamment les travaux de Reimers, 1990 et de Nathan et McMahon, 1992). Lorsqu'un modèle de transformation de la pluie en débit est utilisé, ses paramètres peuvent être liés de diverses façons aux descripteurs du bassin. Nous distinguons ici les deux approches suivantes:

1. quantification de la relation pluie-débit à une échelle régionale, en décrivant les similarités de comportement entre bassins (éventuellement adjacents). Par exemple Nathan et McMahon (1990) ont proposé des méthodes pour scinder un échantillon de 184 bassins du sud-est de l'Australie en groupes de bassins homogènes. Burn et Boorman (1993) ont également utilisé des similarités de caractéristiques de bassins pour identifier des groupes dans un échantillon de 99 bassins de Grande-Bretagne. Dans chaque cas, des méthodes ont été proposées pour transférer les paramètres caractérisant les groupes de bassins à des bassins non jaugés avec des caractéristiques similaires en utilisant des mesures de similarité. Cependant les premiers auteurs mentionnent que la détermination de groupes de bassins est très sensible au type de variables explicatives sélectionnées, comme nous l'avons aussi remarqué dans nos travaux de classification des bassins exposés au chapitre 2. Vandewiele et Elias (1995), eux, ont essayé d'interpoler les paramètres d'un modèle mensuel sur des bassins non jaugés à partir des paramètres de bassins voisins, sur un échantillon de 75 bassins versants en Belgique;
2. étude des liens entre les paramètres du modèle et les caractéristiques du bassin en utilisant des procédés de type régressif (régressions simples ou multiples)

Ces deux approches peuvent être complémentaires. Ici, nous nous intéresserons plus particulièrement à la seconde, qui est la plus couramment appliquée en hydrologie. Nous travaillerons sur des bassins non pas concentrés dans une région relativement homogène ou sous un même climat, mais en diversifiant la localisation des bassins utilisés autant que notre échantillon de base nous le permet. Nous tentons ainsi d'établir des relations avec un spectre de validité aussi large que possible.

Le but de cette seconde approche est d'arriver à extraire un maximum d'information de caractéristiques physiques ou climatiques du bassin pouvant avoir une pertinence hydrologique, autrement dit pouvant influencer la transformation de la pluie en débit. Ce lien peut être fait entre les caractéristiques du bassin et ce qui, dans le modèle, caractérise la transformation pluie-débit sur ce même bassin, c'est-à-dire les paramètres. De nombreux essais adoptant cette approche ont été réalisés avec différents modèles (dont certains ont été testés dans la comparaison) et sous des climats variés. A titre d'exemple, nous répertorions dans le Tableau 8.1 quelques-uns de ces travaux.

Modèle	Etudes de régionalisation	Pays-Région	Nombre de bassins
Stanford	James (1972)	Etats-Unis (Kentucky)	16
	Egbuniwe et Todd (1976)	Nigéria	2*
	Magette <i>et al.</i> (1976)	Etats-Unis (Virginie)	10
CREC et GR3J	Servat et Dezetter (1992, 1993)	Nord de la Côte d'Ivoire	20
HBV	Braun et Renner (1992)	Suisse	5
	Johansson (1994)	Sud de la Suède	11
	Seibert (1999)	Centre de la Suède	11
	Yu et Yang (2000)	Sud de Taïwan	10
HUI de Nash	Tung <i>et al.</i> (1997)	Taiwan	42
IEM4	Pirt et Bramley (1985)	Royaume-Uni (Severn-Trent)	14
IHACRES	Sefton <i>et al.</i> (1995)	Royaume-Uni (Pays-de-Galles)	8
	Post et Jakeman (1996, 1999)	Australie (Victoria)	17
modèle mensuel	Vandewiele <i>et al.</i> (1991)	Nord de la Belgique	60
Sacramento	Weeks et Ashkanasy (1985)	Australie (Queensland)	8
UBC	Micovic et Quick (1999)	Canada (Colombie Britannique)	12
VIC-2L	Abdulla et Lettenmaier (1997)	Centre des Etats-Unis (Arkansas)	34
WATBAL, SFB et SIXPAR	Srikanthan et Goodspeed (1988)	Australie (Nouvelle Galles du Sud)	22

Tableau 8.1: Travaux de régionalisation sur différents modèles pluie-débit (*: travaux en transposition)

Les degrés de succès obtenus ont été assez variés, et souvent trop peu satisfaisants pour pouvoir envisager une application fiable des modèles sur des bassins non jaugés. Les résultats les plus intéressants proviennent des études centrées sur une région particulière ou un grand bassin. Celle réalisée au pas de temps mensuel par Vandewiele *et al.* (1991) fournit également de bons résultats. Dans ces études, la méthode utilisée pour établir des liens entre descripteurs du bassin et paramètres est généralement composée des étapes suivantes:

- (i) sélection d'un jeu de bassins,
- (ii) calage du modèle sur chaque bassin (ou chaque bassin-période) pour obtenir les valeurs des paramètres du modèle dans chaque cas,
- (iii) sélection d'un ensemble de variables physiques (topographie, morphologie, physiographie, couvert végétal, pédologie, géologie...) ou climatiques (décrivant le régime des pluies ou l'évapotranspiration) qui sont supposées pertinentes relativement à leur influence sur la transformation pluie-débit,
- (iv) établissement de relations entre les paramètres du modèle et les variables explicatives sélectionnées en utilisant des régressions simples ou multiples, dans lesquelles paramètres et variables peuvent éventuellement avoir subi des transformations préalables (logarithmique ou puissance par exemple),
- (v) comparaison (pas toujours réalisée) des résultats du modèle obtenus d'une part avec les paramètres calés et d'autre part avec les paramètres prédéterminés par les équations de régionalisation. Parfois cette évaluation est réalisée sur un échantillon de bassins n'ayant pas servi à la mise au point des équations de régression.

Remarquons tout de même l'approche originale adoptée par Yu et Yang (2000) dans leurs travaux de régionalisation des paramètres du modèle HBV: ils ont tout d'abord développé une méthode pour produire des courbes de débits classés sur des bassins non jaugés, et c'est ensuite sur ces courbes que sont calés les paramètres du modèle. Les résultats obtenus sur un échantillon de bassins à Taïwan sont très encourageants.

8.3. Description d'une méthodologie améliorée

La méthode décrite dans le paragraphe précédent par les points (i) à (v) est la plus courante en hydrologie (Tung *et al.*, 1997). Elle fait appel à des outils de régression linéaire pour déterminer des équations de prédétermination des paramètres. C'est en fait une méthode en deux temps, qui tout d'abord caractérise le comportement des bassins à l'aide du modèle par calage des paramètres, et ensuite essaie d'exploiter l'information contenue dans des descripteurs du bassin pour expliquer la variabilité des paramètres calés.

Du fait de ce procédé en deux temps, la dérivation des équations de régression ne tient pas compte directement des non-linéarités générées par le modèle, de la façon dont les paramètres agissent sur les entrées, des éventuelles interactions entre eux⁸ ou de la sensibilité du modèle aux variations de ceux-ci. En effet, les régressions identifient les meilleurs ajustements linéaires pour minimiser l'erreur d'estimation sur les paramètres. Cependant, ces ajustements n'étant jamais parfaits, ils ne garantissent nullement que l'on obtienne en sortie du modèle les meilleurs résultats possibles avec des jeux de paramètres prédéterminés sur l'échantillon de bassins considéré. Ceci vient du fait que la méthode précédemment décrite met l'accent sur la minimisation des erreurs sur les paramètres, et non sur une maximisation des performances du modèle. Or c'est bien ce dernier objectif qui importe dans des cas concrets: sur un bassin non jaugé, on souhaite obtenir les meilleures (les plus fiables) simulations possibles du modèle avec l'information dont on dispose sur le bassin. Les paramètres doivent être considérés ici comme des intermédiaires, vecteurs d'une information apportée au modèle.

La méthode que nous décrivons dans les paragraphes suivants vise donc à faire un meilleur usage de l'information contenue dans les descripteurs du bassin, au regard des seuls résultats du modèle. Bien entendu, plus les relations entre descripteurs et paramètres se rapprochent d'ajustements parfaits (cas rares voire exceptionnels), plus l'objectif de maximisation des performances du modèle coïncide avec celui de minimisation de l'erreur sur les paramètres.

Nous pouvons illustrer la différence entre ces deux approches à l'aide de l'exemple simple suivant. Considérons un modèle interannuel simple qui donnerait sur le bassin j ($1 \leq j \leq N$) le débit moyen interannuel Q_j comme une fonction de la pluie moyenne observée P_j selon la relation:

$$Q_j = \frac{P_j^2}{P_j + A} + \varepsilon_j \quad \text{Eq. (8.1)}$$

où A est le paramètre du modèle et ε_j est l'erreur. Ce modèle est choisi ici arbitrairement pour sa simplicité à des fins d'illustration. Il est inspiré des réflexions de Mouelhi (2000) sur un modèle pluie-débit interannuel. Sur chaque bassin, une estimation de A est donnée par:

$$\tilde{A}_j = P_j \cdot \left(\frac{P_j}{Q_j} - 1 \right) \quad \text{Eq. (8.2)}$$

Dans le cas simple que nous avons choisi, nous obtenons une expression explicite de \tilde{A}_j . Dans le cas général, \tilde{A}_j est estimé grâce à une procédure d'optimisation numérique. Supposons maintenant que le paramètre A est lié à un descripteur S du bassin par la simple relation linéaire suivante:

⁸ Tung *et al.* (1997) remarquent que la dérivation d'équations de régression sur les deux paramètres (N, K) de l'hydrogramme unitaire de Nash ne sont pas fiables lorsque l'on ne tient pas compte des interactions dans le modèle entre ces deux paramètres lors de l'application des procédures de régression.

$$A = \alpha \cdot S + \eta \quad \text{Eq. (8.3)}$$

où α est un coefficient et η est l'erreur de la relation. La valeur de α peut être obtenue par régression, c'est-à-dire en minimisant l'expression *OF1* suivante:

$$OF1 = \sum_{j=1}^N (\tilde{A}_j - \alpha \cdot S_j)^2 \quad \text{Eq. (8.4)}$$

En combinant l'Eq. (8.2) et l'Eq. (8.4) et en dérivant l'expression de *OF1* par rapport à α , on obtient l'expression de $\tilde{\alpha}$ qui minimise l'erreur de la relation. Elle est donnée par:

$$\tilde{\alpha} = \frac{\sum_{j=1}^N P_j \cdot \left(\frac{P_j}{Q_j} - 1 \right)}{\sum_{j=1}^N S_j} \quad \text{Eq. (8.5)}$$

C'est la valeur classiquement obtenue dans les études de régionalisation.

Considérons maintenant la deuxième approche dans laquelle on cherche à optimiser les performances du modèle, en supposant la même forme de relation entre A et S que celle donnée à l'Eq. (8.3). Nous cherchons à optimiser non plus l'erreur de la relation de régression mais la différence entre les débits observés et les débits calculés avec les équations de prédétermination, c'est-à-dire minimiser la fonction *OF2* suivante:

$$OF2 = \sum_{j=1}^N \left(Q_j - \frac{P_j^2}{P_j + \alpha \cdot S_j} \right)^2 \quad \text{Eq. (8.6)}$$

avec les mêmes notations que précédemment. En dérivant cette expression par rapport à α , on obtient l'équation suivante qui donne une expression implicite de $\tilde{\alpha}$:

$$\sum_{j=1}^N \frac{P_j^2 \cdot Q_j}{(P_j + \tilde{\alpha} \cdot S_j)^2} = \sum_{j=1}^N \frac{P_j^4}{(P_j + \tilde{\alpha} \cdot S_j)^3} \quad \text{Eq. (8.7)}$$

De toute évidence, l'expression de $\tilde{\alpha}$ donnée à l'Eq. (8.5) n'est pas solution de l'Eq. (8.7). La valeur $\tilde{\alpha}$, qui est optimale pour l'équation de régression, ne l'est donc pas forcément par voie de conséquence pour obtenir les meilleurs résultats possibles du modèle de simulation des débits.

La méthode classique privilégie donc l'ajustement des paramètres au détriment (généralement) des performances du modèle. Elle présente cependant l'avantage d'être relativement facile d'application. Dans la méthode que nous utiliserons, les coefficients des équations de régression liant les paramètres du modèle aux descripteurs du bassin ne peuvent plus être calculés automatiquement, mais doivent être déterminés par essai-erreur au cours d'une procédure itérative faisant intervenir le modèle. Les équations ainsi déterminées donnent généralement de moins bons ajustements entre paramètres et descripteurs, mais inversement il en résulte de meilleures performances du modèle.

En guise d'initialisation de cette nouvelle méthode, on reprend les premières étapes de l'approche classique correspondant aux points (i) à (iv): il est ainsi possible de déterminer quels descripteurs peuvent être pertinents et sous quelle forme ils interviennent dans les équations qui les lient aux paramètres. On poursuit alors la procédure par les points suivants:

- (v') appliquer le modèle aux bassins de l'échantillon avec les jeux de paramètres prédéterminés par les équations de régression, et déterminer les performances moyennes du modèle (ou la distribution des performances) qui serviront de performances de référence;
- (vi) sur chaque bassin de l'échantillon, tester des formulations modifiées des équations de régression, dans lesquelles les valeurs des coefficients initiaux ont subi de petites variations. On teste ainsi différentes modifications des coefficients des équations ainsi que diverses combinaisons d'équations;
- (vii) lorsque tous les bassins ont été testés, c'est-à-dire lorsque successivement sur chacun d'eux on a appliqué toutes les formulations modifiées des équations, on retient la combinaison de formulations modifiées qui permet d'obtenir les meilleures performances du modèle.

Les étapes (vi) et (vii) sont réitérées pour tester de nouvelles formulations des équations jusqu'à atteindre une convergence vers des valeurs de coefficients des équations de régression qui ne permettent plus d'améliorer les résultats du modèle.

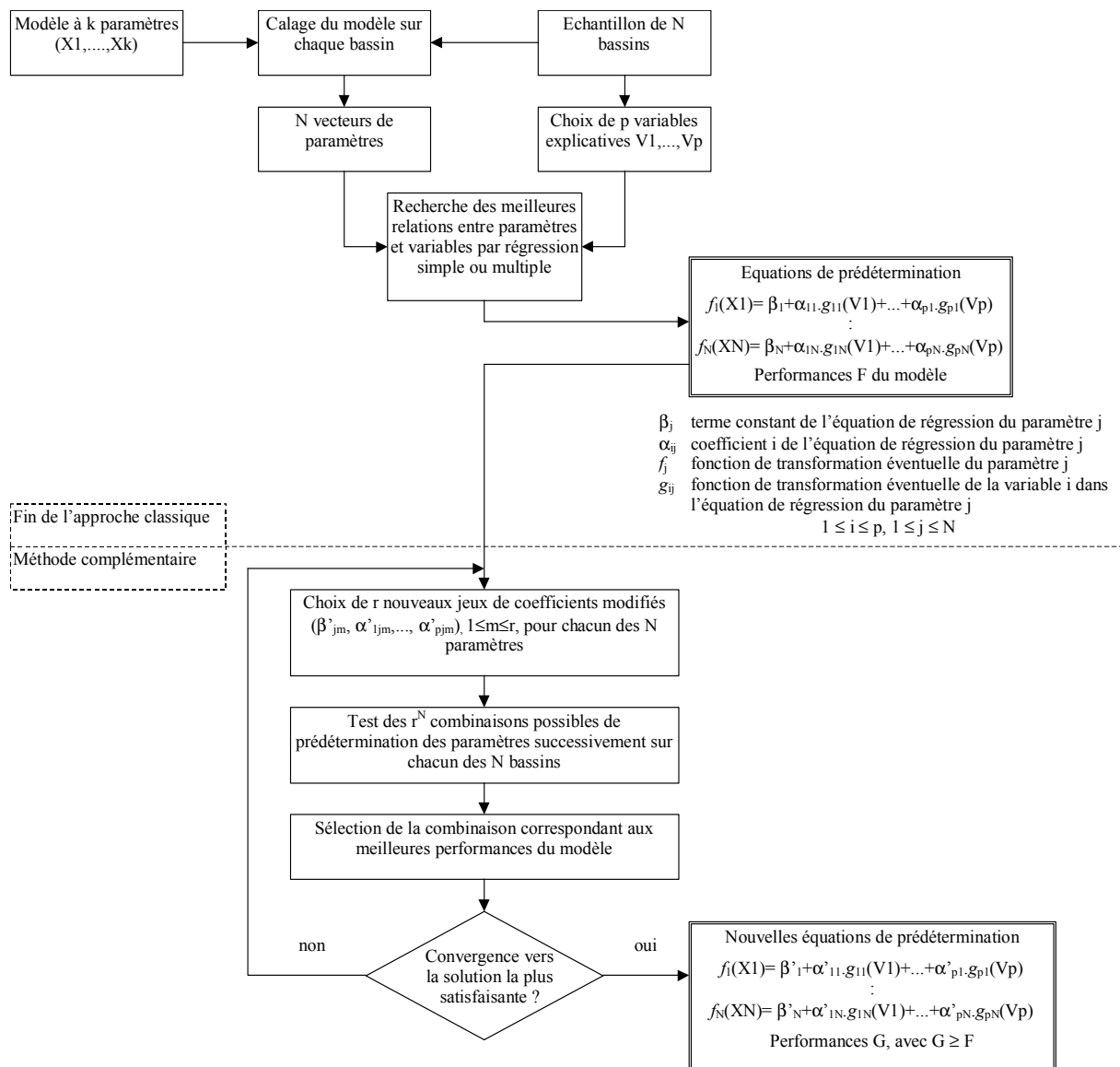


Figure 8.1: Représentation schématique des procédures de régionalisation classique et améliorée

Dans l'étape (vi), l'avantage de tester toutes les formulations sur un bassin après l'autre est de n'avoir ainsi qu'une phase de lecture de données par bassin, ce qui permet de réduire le temps nécessaire pour mener à bien cette étape de l'algorithme. La solution de tester une formulation après l'autre sur l'ensemble des bassins est beaucoup trop longue. La procédure complète est illustrée à la Figure 8.1.

Cette procédure itérative revient en fait à réaliser une optimisation des coefficients des équations de régression, en prenant comme fonction objectif les performances du modèle. Les étapes (vi) et (vii) sont répétées jusqu'à ce que l'on n'obtienne plus d'amélioration des performances du modèle, ce qui représente une masse importante de calculs. Ici, nous avons opté pour une procédure essentiellement manuelle, qui se déroule par essai-erreur: à la fin de chaque étape (vii), on choisit de nouveaux coefficients des équations, dont on teste l'efficacité en revenant à l'étape (vi). La méthode est d'autant plus longue que le nombre de paramètres du modèle et le nombre de coefficients des équations de régression sont élevés. Notons qu'il est possible par cette approche de tester différentes formes d'équations de régression, par exemple des équations déterminées sur les paramètres transformés ou sur les paramètres réels.

La méthode itérative peut être automatisée, en utilisant une méthode d'optimisation pour déterminer un jeu optimum de coefficients des équations de régression, la fonction objectif reposant sur les performances du modèle. Cela suppose néanmoins d'importants moyens de calcul lorsque l'on s'intéresse à un échantillon de bassins conséquent. Notons par ailleurs que pour une équation de prédétermination donnée, l'ajustement des paramètres ne peut se faire au hasard: pour une même équation, les variations des valeurs des différents coefficients ne peuvent pas être réalisées indépendamment, sous peine de faire perdre tout réalisme aux équations de régression initialement déterminées.

8.4. Choix du modèle et travaux antérieurs

Comme l'ont noté certains auteurs (par exemple Srikanthan et Goodspeed, 1988 ou Post et Jakeman, 1999), la détermination non ambiguë des valeurs des paramètres est un pré-requis à la mise en oeuvre d'études de régionalisation. En effet la sur-paramétrisation des modèles complexes entraîne la non-unicité d'un jeu optimum de paramètres. L'incertitude sur la valeur de ces paramètres est alors d'autant plus grande, ce qui diminue la fiabilité des équations de prédétermination des paramètres. Il est donc préférable d'avoir un modèle dans lequel on se garde d'introduire des fonctions qui se révéleraient peu pertinentes et dont la détermination des paramètres serait difficile. Le chapitre précédent indique qu'en adoptant cette attitude de minimalité, il est difficile de dépasser quatre paramètres. Une telle parcimonie, en diminuant les problèmes d'identifiabilité, est donc recommandable pour permettre des études de régionalisation, dans la mesure tout de même où le modèle reste fiable pour caractériser la transformation pluie-débit sur le bassin.

Par ailleurs, l'applicabilité de la méthode précédemment décrite est d'autant plus aisée que le modèle utilisé a peu de paramètres. Par conséquent, le modèle GR à quatre paramètres, dans sa version améliorée présentée au chapitre précédent, semble bien adapté pour tester cette méthode sur notre échantillon de bassins versants. Des travaux d'explication des paramètres des modèles GR ont déjà été menés sur d'autres versions. Nous nous proposons ici d'en rappeler les principaux résultats, pour mieux situer le contexte de notre démarche.

Les notations adoptées pour les paramètres du modèle GR journalier sont les suivantes: A , capacité du réservoir de production; B , capacité à un jour du réservoir de routage; C , temps de base du principal hydrogramme unitaire; D , paramètre d'échange.

Edijatno (1991), après avoir mis au point la première version à trois paramètres GR3J du modèle journalier, s'est intéressé au lien entre les valeurs des paramètres du modèle et des descripteurs du bassin. Il disposait pour cela d'un échantillon de 93 bassins français (sur un échantillon initial de 110), où le modèle obtenait des critères de Nash supérieurs à 40 % et sur lesquels étaient disponibles la superficie, la pluie journalière de fréquence décennale, la pluie annuelle moyenne et la température annuelle moyenne. Pour 39 de ces bassins, des informations physiques avaient également été rassemblées (indice global de pente, longueur de thalwegs, indice de "torrentialité" et pourcentage de surface boisée). Quelques conclusions intéressantes ressortent de cette étude:

- les trois paramètres A , B et C du modèle apparaissent très faiblement liés voire indépendants, ce qui tend à renforcer l'idée que ces trois paramètres sont nécessaires pour la simulation;
- seules trois variables (pluie moyenne, température moyenne et surface) ont contribué à apporter une explication significative des paramètres;
- les relations obtenues sont d'une efficacité limitée, en particulier pour les paramètres A et B et ne permettent pas en l'état une application fiable à des bassins non jaugés;
- il n'existe pas de tendance régionale nette dans la similarité des valeurs des paramètres, à part en Bretagne, en Charente et dans les Vosges où semble se dessiner une certaine homogénéité (la répartition géographique des paramètres, utilisée également par Kabouya, 1990, dans le cas de 70 bassins d'Algérie septentrionale avec le modèle mensuel à trois paramètres GR3M, n'avait pas permis non plus de mettre en évidence des tendances régionales marquées).

Les relations d'Edijatno (1991), établies par régression multiple sur les transformées logarithmiques des paramètres, sont les suivantes pour les paramètres vrais:

$$A = 0,55.SF^{0,1}.TAM^{2,3} \quad \text{Eq. (8.8)}$$

$$B = \frac{1}{1800}.SF^{0,3}.PAM^{1,6} \quad \text{Eq. (8.9)}$$

$$C = 3500.SF^{0,3}.PAM^{-0,6}.TAM^{2,4} \quad \text{Eq. (8.10)}$$

où SF est la superficie du bassin (km^2), PAM la pluie annuelle moyenne (mm) et TAM la température annuelle moyenne ($^{\circ}\text{C}$). Les coefficients de détermination R^2 obtenus pour ces ajustements sont respectivement de 0,25, 0,25 et 0,58.

Utilisant cette même version du modèle journalier, Servat et Dezetter (1993) ont lié les paramètres A , B et C à des caractéristiques de 20 bassins de Côte d'Ivoire. Les relations établies font intervenir des paramètres d'occupation du sol, de pluie cumulée sur certaines périodes de l'année et un indice de compacité du bassin dans le cas du paramètre C . Les coefficients de détermination obtenus pour les trois paramètres sont respectivement de 0,32, 0,44 et 0,87. Comparé à une version simplifiée du modèle CREC ayant sept paramètres, GR3 semble présenter une meilleure aptitude à la régionalisation des paramètres. La conclusion de cette étude souligne l'intérêt d'avoir un modèle à peu de paramètres pour ce genre d'approche.

Partant notamment des constatations d'Edijatno (1991) sur la relative homogénéité des paramètres du modèle GR3 sur certaines régions française, Makhlouf (1994) s'est intéressé à l'explication des quatre paramètres de la version GR4J proposée par Nascimento (1995) (qui a ajouté le paramètre d'échanges D). Cette étude a notamment porté sur un groupe de 23 bassins versants situés en Bretagne (avec 13 variables dont superficie, paramètres climatiques,

morphologiques, géologiques, et de végétation), sur un groupe de 34 bassins versants du bassin de la Moselle (avec 14 variables explicatives dont superficie, altitude, pente, paramètres d'occupation du sol, de perméabilité et de lithologie), puis sur un échantillon de 349 bassins français répartis sur le territoire national.

Dans le premier cas, des relations assez satisfaisantes (faisant intervenir la pente, la pluie moyenne annuelle, le pourcentage d'affleurement de granite et le nombre moyen de jours de pluie par an) ont été mises au point pour les paramètres B et C , et dans une moindre mesure pour le paramètre A . En revanche aucune relation satisfaisante n'a pu être établie avec le paramètre d'échange D . Des relations intéressantes ont été obtenues pour B et C lorsque le paramètre d'échange est fixé à une valeur nulle et le paramètre A à une valeur moyenne pour tous les bassins. Le modèle régional ainsi construit (deux paramètres calculés, deux paramètres fixes) a donné des résultats satisfaisants, meilleurs sur 17 de ces bassins bretons que des résultats obtenus par les modèles Martine et Amande du BRGM (Mazenc *et al.*, 1984) dont les paramètres avaient été optimisés sur chaque bassin.

Sur les bassins de Moselle, des relations acceptables avec les descripteurs du bassin ont été de nouveau établies pour les paramètres B et C , le paramètre A restant plus faiblement expliqué. Ici cependant, le paramètre D a pu être lié significativement à l'altitude ($R^2=0,33$).

Pour pallier la difficulté de relier les paramètres de bilan (A et D) du modèle avec les caractéristiques du bassin, Makhoulf a mis au point un modèle mensuel à deux paramètres pour lequel de bons résultats en régionalisation ont été obtenus sur les bassins bretons. Ensuite, grâce à la parenté des structures des modèles mensuel et journalier, des relations significatives ont pu être établies entre les paramètres des deux modèles, ouvrant ainsi la voie à une explication en deux temps des paramètres du modèle journalier. Cette approche est intéressante puisqu'elle montre qu'un modèle à plus grand pas de temps, en accordant une plus grande importance aux fonctions de rendement, peut permettre d'établir des relations de régionalisation plus fiables sur ces fonctions que dans un modèle journalier, où la partie routage du modèle a plus de poids. Remarquons d'ailleurs que de bons succès ont également été obtenus par Vandewiele *et al.* (1991) avec un modèle mensuel. Ceci souligne l'intérêt de construire des familles de modèles à différents pas de temps dont les structures présentent des similarités.

Reprenant enfin la démarche d'Edijatno (1991) sur un échantillon étendu de 349 bassins répartis en France, Makhoulf retrouve des conclusions similaires aux études sur les bassins de Bretagne et de Moselle, avec des explications assez satisfaisantes pour B et C et moins bonnes pour A et D . Les relations suivantes ont été proposées:

$$A = 4,1.SF^{0,3} . TAM^4 . PJ10^{-1} \quad \text{Eq. (8.11)}$$

$$B = \frac{1}{2700} . R^{-1,7} . PAM^{1,8} \quad \text{Eq. (8.12)}$$

$$C = 6,8.SF^{0,2} . PAM^{-0,4} . R^{-0,2} \quad \text{Eq. (8.13)}$$

$$D = 4,5.TAM . PAM^{-0,6} \quad \text{Eq. (8.14)}$$

où $PJ10$ est la pluie journalière décennale (mm) et R un coefficient régional introduit dans la méthode Crupedix d'estimation de la crue décennale (CTGREF, 1980), avec les mêmes notations que précédemment pour les autres variables. Les coefficients de détermination R^2 obtenus pour ces ajustements sont respectivement de 0,26, 0,42, 0,54 et 0,19. Dans cette dernière phase, Makhoulf a utilisé la totalité puis des sous-ensembles de son échantillon de bassins, suivant la disponibilité des variables explicatives et avec un classement suivant les

performances obtenues par le modèle. Dans certains cas, de façon surprenante, la qualité des ajustements (suivant R^2) baissait lorsque l'échantillon était restreint aux bassins où le modèle marchait le mieux. En fait, cet effet est lié à une diminution de la plage de variation des paramètres; l'erreur standard, elle, diminue toujours lorsque les performances du modèle augmente.

Enfin, les travaux de Zermani (1998) se sont intéressés à des aspects de spatialisation, avec régionalisation des paramètres de plusieurs modèles sur des bassins versants du bassin de la Charente. Pour le modèle GR4J, des relations satisfaisantes ont pu être obtenues pour les paramètres A , B et C en considérant des paramètres pédologiques, géologiques ou agronomiques (réserve utile). En revanche, le paramètre D est resté difficilement explicable par les variables considérées. Cependant, le nombre de bassins utilisés dans cette étude reste assez faible.

Des études précédentes, nous pouvons retenir les points suivants:

1. la méthodologie utilisée s'appuie toujours sur une analyse par régression multiple entre descripteurs du bassin et paramètres du modèle;
2. les relations les plus satisfaisantes avec des descripteurs du bassin ont été obtenues pour les paramètres de routage du modèle (B et C). B a pratiquement dans tous les cas été relié à un paramètre de pluie moyenne annuelle (ou pluie cumulée sur une période de l'année). C a été systématiquement lié à la surface ou à un descripteur physiographique (pente, indice de compacité...) faisant référence à un temps de concentration sur le bassin. Les paramètres de bilan (A et D dans le cas de GR4J) ont généralement été plus difficilement liés aux descripteurs du bassin, en particulier pour le paramètre d'échange D ;
3. même lorsqu'un nombre important de variables explicatives est disponible en début d'analyse, seules quelques-unes apparaissent significatives pour expliquer la variance des paramètres;
4. les relations établies par Edijatno (1991) et Makhlouf (1994) données précédemment présentent d'importantes similarités pour les paramètres A , B et C ; elles ne peuvent cependant pas être exploitées avec fiabilité en l'état sur des bassins non jaugés;
5. au niveau régional, des relations souvent plus satisfaisantes ont pu être établies, ceci provenant vraisemblablement de la plus grande homogénéité des comportements des bassins et des influences climatiques. Cette tendance a également été remarquée par Johansson, (1994) qui mentionne qu'en Finlande et en Norvège les relations régionales, spécifiques à une zone, sont moins difficiles à établir que des relations beaucoup plus générales pour l'ensemble du territoire;
6. les modèles GR3J et GR4J ont montré des performances aussi voire plus satisfaisantes que d'autres modèles avec un plus grand nombre de paramètres lors d'études de régionalisation.

Notons par ailleurs que les équations de prédétermination des paramètres font souvent intervenir des variables climatiques. Il peut paraître surprenant que de telles caractéristiques, variables de forçage du bassin, jouent un rôle aussi important que les caractéristiques intrinsèques telles que pédologie ou géologie. Ceci indiquerait qu'il existe une interaction forte entre le bassin et le climat auquel il est soumis, et que ces deux éléments sont nécessaires pour caractériser le bassin d'un point de vue hydrologique, à la manière dont le fait un modèle pluie-débit. Cela provient probablement aussi du fait que les paramètres du modèle sont calés à partir de séries de données hydroclimatiques.

Il faut par ailleurs souligner qu'il est très difficile, d'après ces études, d'arriver à des explications satisfaisantes de paramètres, même dans le cas d'un modèle très simple où les problèmes de détermination des valeurs des paramètres par optimisation sont minimisés. A cette difficulté, nous pouvons essayer d'avancer plusieurs explications:

- comme nous l'avons expliqué à plusieurs reprises, le fait que la structure ne soit qu'une grossière approximation de la réalité limite la correspondance avec le système réel;
- chaque paramètre du modèle est intégrateur de l'effet conjoint de beaucoup de caractéristiques du bassin, apparentes ou cachées;
- de la même façon que les paramètres du modèle hydrologique sont globaux, les variables explicatives utilisées lors de travaux de régionalisation sont globales: elles décrivent en un chiffre des caractéristiques hétérogènes sur le bassin et parfois variables dans le temps.

Il se pose donc un double problème de pertinence, des paramètres du modèle d'une part et des variables explicatives d'autre part, quant à leur pouvoir de description du système modélisé. Les limites des travaux de régionalisation proviennent donc probablement de la difficulté de mettre en adéquation deux systèmes simples de caractérisation du bassin, l'un fondé sur une approche modélisatrice, l'autre sur une approche descriptive (climat-géographie). Si l'on accepte le modèle mathématique comme base de travail, c'est essentiellement l'adéquation des descripteurs utilisés qui conditionne le succès ou l'échec de la démarche de régionalisation, cette pertinence étant à la fois relative au bassin et au modèle hydrologique.

Nous allons maintenant appliquer la méthode proposée dans la partie précédente, en reprenant successivement les différentes étapes.

8.5. Application

8.5.1. Choix des bassins

8.5.1.1. Calage du modèle et sélection des bassins

Sur chacun des 429 bassins dont nous disposons (307 (71,6 %) en France et 122 (28,4 %) à l'étranger), la plus longue chronique ininterrompue disponible a été sélectionnée. Le modèle GR4J a été calé sur l'ensemble de l'échantillon avec un critère de Nash calculé sur les racines carrées des débits (critère *CR2*). Pour l'analyse, nous n'avons retenu dans un premier temps que les bassins où le modèle est exploitable, partant de l'idée que l'inefficacité du modèle sur un bassin se traduit par la non-significativité de ses paramètres. Nous avons ainsi retenu les bassins pour lesquels le modèle répond à la double exigence d'avoir un critère de calage (*CR2*) et un critère de Nash sur les débits (*CR1*) supérieurs ou égaux à 80 %. 131 bassins répondaient à ces conditions dont 115 (87,8 %) en France et 16 (12,2 %) à l'étranger. La proportion de bassins étrangers est plus faible que dans l'échantillon initial, ceci étant lié au fait que de moins bonnes performances ont été obtenues en moyenne sur cet échantillon. Nous ne nous plaçons donc pas ici dans le cas d'une étude à l'échelle régionale, où l'on peut avoir une certaine homogénéité dans les caractéristiques et la réponse des bassins.

8.5.1.2. Analyse de l'échantillon des paramètres

Une première étape de l'analyse vise à vérifier l'indépendance des paramètres. Le Tableau 8.2 fournit la matrice de corrélation des paramètres transformés. Les paramètres *A*, *B*, *C* et *D* sont largement indépendants, avec des coefficients de corrélations faibles. Seule une légère liaison entre les deux paramètres de routage (*B* et *C*) apparaît.

Les essais de mise en relation des paramètres par régression multiple ne donnent, dans aucun des cas, des équations avec des rapports de Student significatifs, sauf pour les paramètres B et C où l'on atteint tout juste le seuil de significativité. Ces résultats sont légèrement différents de ceux donnés par Edijatno (1991), qui dans le cas de la première version du modèle GR3, faisait état d'une légère corrélation entre A et B et une absence de corrélation de C avec les deux autres paramètres. Cette différence s'explique probablement par la présence ici du quatrième paramètre. Elle peut aussi provenir de la différence de fonction objectif utilisée pour caler les paramètres des modèles: critère de Nash sur les débits dans l'étude de Edijatno (1991), critère de Nash sur les racines carrées des débits ici (la dépendance entre paramètres et fonction objectif ayant été illustrée au chapitre précédent). Remarquons d'ailleurs que les relations établies entre paramètres du modèle et descripteurs du bassin dans des études de régionalisation seront dépendantes de la fonction objectif choisie, ceci dépendant bien sûr de la sensibilité des paramètres à cette fonction objectif. Ainsi la forme des équations de régression et les variables pertinentes pour l'explication des paramètres calés sur un critère donnant un poids préférentiel aux crues seront probablement différentes de celles obtenues lorsque le modèle est calé avec un critère plus général.

	$\ln(A)$	$\ln(B)$	$\ln(C-0,5)$	$\text{argsh}(D)$
$\ln(A)$	1,00			
$\ln(B)$	0,04	1,00		
$\ln(C-0,5)$	0,05	-0,25	1,00	
$\text{argsh}(D)$	-0,07	0,07	-0,12	1,00

Tableau 8.2: Matrice de corrélation entre paramètres

Ces résultats de corrélation confirment que les paramètres du modèle ont peu de liens entre eux et que les problèmes d'interaction sont donc faibles, rejoignant les remarques formulées au chapitre 6 sur les interactions entre les quatre paramètres du modèle GR4J. Le Tableau 8.3 donne quelques caractéristiques de l'échantillon des paramètres (en valeur transformée et en valeur réelle) obtenus par calage sur les 131 bassins. Si les paramètres transformés varient sur des intervalles restreints (c'était le but des transformations !), les paramètres vrais peuvent en revanche varier sur plusieurs ordres de grandeur (par exemple pour les capacités A et B des réservoirs).

Paramètres transformés	$\ln(A)$	$\ln(B)$	$\ln(C-0,5)$	$\text{argsh}(D)$
	Moyenne	6,03	4,62	0,31
Ecart-type	0,63	0,74	0,55	1,27
Médiane	5,98	4,68	0,21	-0,51
Minimum	3,05	2,18	-0,99	-4,60
Maximum	8,23	7,02	1,62	2,18
Paramètres réels	A	B	C	D
	Moyenne	515,6	133,4	2,10
Ecart-type	455,1	127,4	1,03	5,51
Médiane	395,4	107,8	1,73	-0,53
Minimum	21,1	8,8	0,87	-49,74
Maximum	3751,8	1118,8	5,55	4,37

Tableau 8.3: Caractéristiques de l'échantillon des 131 jeux de paramètres

8.5.2. Choix et analyse de variables explicatives

Les descripteurs retenus sont simples et correspondent au niveau d'information dont nous avons disposé au cours de ce travail sur la majorité des bassins étudiés. Les quatre variables explicatives choisies sont:

- la superficie du bassin SF , en km^2 ,
- la pluie annuelle moyenne PAM , en mm ,
- l'ETP annuelle moyenne EAM , en mm ,
- le coefficient d'irrégularité saisonnière des pluies, CP (comme il est défini au chapitre 2).

Sur les 131 bassins, les plages de variation de ces variables sont données par le Tableau 8.4. A titre d'information, nous avons également indiqué le coefficient de rendement (rapport du débit à la pluie). Cet échantillon présente une assez grande diversité de tailles de bassins et de conditions hydro-climatiques.

	SF (km^2)	PAM (mm)	EAM (mm)	CP	$REND$ (%)
Médiane	147	1043,3	719,1	179,1	0,45
Minimum	0,1	662,1	635,6	145,5	0,01
Maximum	50600	1940,3	2045,5	399,7	1,01

Tableau 8.4: Caractéristiques des bassins utilisés pour établir les équations de prédétermination des paramètres

Dans la suite de l'analyse, nous avons parfois utilisé des variables composites issues de plusieurs des variables précédentes, comme par exemple:

- le rapport de la pluie moyenne PAM à l'ETP moyenne EAM ;
- le produit du coefficient d'irrégularité CP par la variable PAM , qui représente en fait la différence de pluie entre le mois le plus arrosé et le mois le moins arrosé (avec un coefficient 12);
- le produit de CP et PAM divisé par la variable EAM , qui représente la différence de pluie entre le mois le plus arrosé et le mois le moins arrosé rapportée à l'ETP moyenne mensuelle; cette variable composite permet d'associer en une même variable les deux causes principales qui peuvent générer la variabilité saisonnière du régime des débits.

La matrice de corrélation entre les variables explicatives (voir Tableau 8.5) montre qu'elles ne sont pas toutes indépendantes. En particulier, il existe une corrélation assez élevée entre EAM et CP . Ce lien est illustré à la Figure 8.2. C'est la seule relation significative qui peut être établie entre les variables. Elle indique en fait que, sur ces 131 bassins, plus l'ETP moyenne est élevée, plus il y a des saisons sèche et pluvieuse contrastées.

	EAM	PAM	CP	SF
EAM	1,00			
PAM	-0,04	1,00		
CP	0,81	0,14	1,00	
SF	0,04	0,04	0,05	1,00

Tableau 8.5: Matrice de corrélation entre variables explicatives

Les variables sélectionnées ne comprennent pas de caractéristiques physiques (en dehors de la superficie) relatives à la pédologie, la géologie, la physiographie ou au couvert végétal des bassins. Du fait de l'absence de telles variables, la tentative d'explication des paramètres que nous allons réaliser se trouvera nécessairement limitée. En effet, comme nous l'avons mentionné précédemment, le bassin versant est un système dont les caractéristiques résultent d'une interaction entre un substrat physique et le climat auquel il est soumis. Il paraît donc logique que ces deux types de descripteurs soient nécessaires pour bien décrire le bassin d'un point de vue hydrologique.

A un niveau régional, il est plus facile de collecter ces données, et ce supplément d'information, combiné avec une plus grande homogénéité des conditions hydro-climatiques,

permet généralement d'obtenir des corrélations plus satisfaisantes entre paramètres du modèle et variables explicatives, que celles que nous présentons dans la suite (voir par exemple les travaux de Makhlouf, 1994).

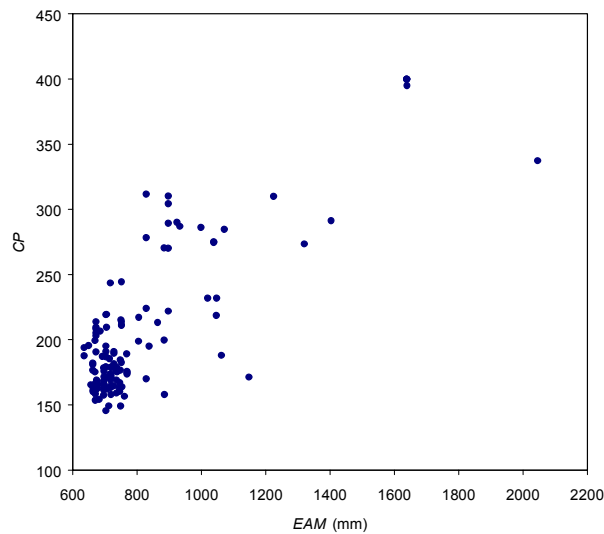


Figure 8.2: Lien entre les variables *EAM* et *CP*

Cependant, l'intérêt de notre étude repose davantage ici sur l'évaluation de l'efficacité de la méthodologie proposée que sur la recherche approfondie de variables explicatives pertinentes, même si cette dernière reste indispensable dans la démarche globale de régionalisation.

8.5.3. Recherche de relations entre paramètres du modèle et variables explicatives

8.5.3.1. Méthodologie de recherche

Cette phase de la démarche constitue l'étape essentielle de l'approche classique de régionalisation. Dans l'approche proposée, elle ne servira cependant que de base pour détecter les variables pertinentes à utiliser et la forme générale des équations de régression. Des régressions multiples ont été réalisées d'une part sur les transformées des paramètres telles qu'elles sont utilisées par l'algorithme de calage du modèle et d'autre part sur les valeurs réelles des paramètres telles qu'elles sont utilisées dans le modèle. Abdulla et Lettenmaier (1997) font également intervenir diverses transformations — sur les paramètres ou les variables explicatives — dans la recherche des équations de régression. Dans la mesure du possible, nous avons choisi ici de ne retenir qu'une seule variable explicative par paramètre pour avoir des équations simples de régression, quitte à utiliser les variables composites précédemment décrites.

Dans le même esprit que les travaux menés par Edijatno (1991) et Makhlouf (1994), nous utilisons des outils de régression linéaire multiple fondés sur le modèle linéaire général, où l'on cherche à expliquer une variable Y (paramètre du modèle) en fonction de k variables x_1, \dots, x_k (descripteurs du bassin), à partir de n observations (Michel, 1989). Le modèle s'écrit alors:

$$\begin{cases} Y_1 = a_0 + a_1 x_{11} + \dots + a_k x_{k1} + e_1 \\ Y_2 = a_0 + a_1 x_{12} + \dots + a_k x_{k2} + e_2 \\ \vdots \\ Y_n = a_0 + a_1 x_{1n} + \dots + a_k x_{kn} + e_n \end{cases} \quad \text{Eq. (8.15)}$$

où a_0, \dots, a_k sont les paramètres du modèle et e_1, \dots, e_n les termes d'erreurs. Les paramètres sont optimisés de manière à maximiser le coefficient de détermination (carré du coefficient de corrélation de la relation linéaire), dont nous utilisons ici l'estimateur non débiaisé. La qualité de l'ajustement peut aussi être évaluée par l'erreur standard. La significativité de la relation peut être estimée par le coefficient de Fischer qui fixe les minima acceptables pour le coefficient de détermination. Ici nous avons utilisé la variable $|t|$ de Student pour évaluer la significativité de chaque variable dans les relations. Si le rapport entre la valeur du coefficient de la régression et son écart-type (a_i/σ_{a_i}) est supérieur au quantile 0,99 de la distribution théorique de $|t|$ (qui dans notre cas est de l'ordre de 3), on a ainsi une chance sur 100 de considérer à tort le coefficient a_i comme significativement différent de zéro. Nous avons également utilisé les illustrations graphiques et l'erreur standard des régressions pour évaluer la qualité de l'ajustement, le coefficient de corrélation pouvant être assez trompeur du fait de la présence de *horsains* (un point en dehors du nuage peut influencer fortement la valeur de R^2).

L'utilisation du modèle linéaire correspond à une première approche simple, qui permet de dégager des tendances dans les liens entre variables explicatives et paramètres. Si l'on trouve des résultats assez satisfaisants, on peut alors se permettre d'introduire des sophistications dans ce modèle rudimentaire. Remarquons que l'utilisation de paramètres transformés dans la recherche de ces relations permet d'obtenir des relations non-linéaires entre paramètres et descripteurs.

8.5.3.2. Régressions sur les paramètres transformés

Pour les quatre paramètres A , B , C et D , nous retrouvons à peu près les mêmes ordres de grandeur des coefficients de corrélation que dans les travaux d'Edijatno (1991) et Makhlof (1994). On peut noter (résultat attendu) que les erreurs standards sur les paramètres sont inférieures aux écarts-types correspondants du Tableau 8.3. Nous avons appliqué ici aux variables une transformation logarithmique préalable.

Paramètre A

Le Tableau 8.6 récapitule les relations entre paramètre A et variables explicatives, dont le seuil de significativité a été jugé satisfaisant d'après les valeurs de la variable de Student.

Paramètre	Formule de régression	Coefficients de régression	Rapport de Student	Erreur standard	Coefficient de détermination
ln(A)	a_0	1.17	0.83	0.57	0.20
	$+ a_1 \ln(SF)$	0.10	3.94		
	$+ a_2 \ln(CP)$	0.86	3.36		
	$+ a_3 \ln(PAM/EAM)$	-0.49	2.97		
ln(A)	a_0	-0.09	0.06	0.59	0.14
	$+ a_1 \ln(SF)$	0.08	3.41		
	$+ a_2 \ln(CP)$	1.08	4.29		
ln(A)	a_0	5.91	0.06	0.59	0.13
	$+ a_1 \ln(SF)$	0.07	3.41		
	$+ a_2 \ln(PAM/EAM)$	1.08	4.29		
ln(A)	a_0	-2.53	1.38	0.58	0.16
	$+ a_1 \ln(SF)$	0.09	3.60		
	$+ a_2 \ln(EAM)$	1.22	4.59		

Tableau 8.6: Relations entre paramètre transformé $\ln(A)$ et variables explicatives

Ici, même si les seuils de significativité sont à peu près acceptables, les ajustements graphiques correspondants révèlent que pour aucune des relations proposées une réelle

tendance générale d'ajustement n'est observable pour le nuage de points. Ceci est illustré à la Figure 8.3, où sont représentées les valeurs du paramètre calculées avec la première équation du Tableau 8.6: il apparaît que c'est davantage l'effet des *horsains* qui contribue à rendre le coefficient de détermination significatif. Pour cette raison, il nous a semblé difficile de retenir une quelconque équation de prédétermination pour le paramètre transformé A , et nous nous en tiendrons donc à prendre la valeur moyenne du paramètre.

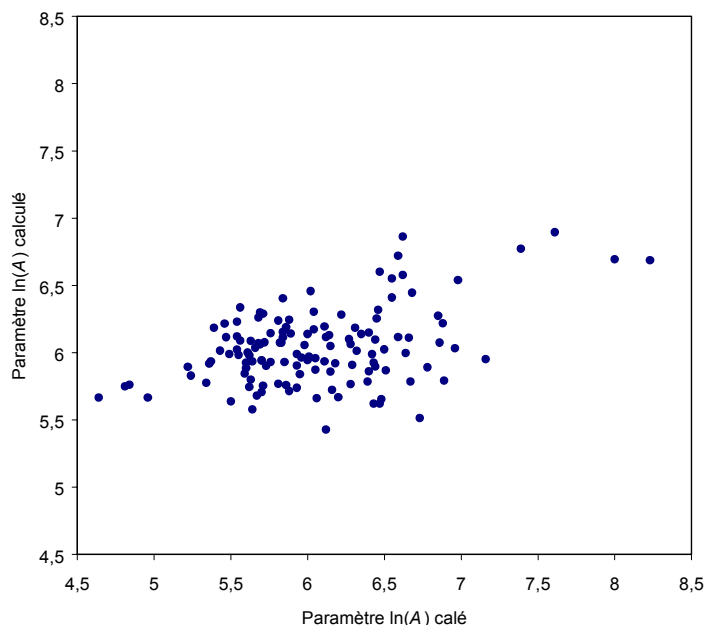


Figure 8.3: Graphe de corrélation entre valeurs transformées de A calées et calculées par l'équation de prédétermination

Paramètre B

Paramètre	Formule de régression	Coefficients de régression	Rapport de Student	Erreur standard	Coefficient de détermination
$\ln(B)$	a_0	-2.54	1.14	0.59	0.38
	$+ a_1 \ln(EAM)$	-1.72	4.33		
	$+ a_2 \ln(PAM)$	0.95	4.41		
	$+ a_3 \ln(CP)$	2.26	5.94		
$\ln(B)$	a_0	-5.33	4.03	0.60	0.36
	$+ a_1 \ln(PAM/EAM)$	1.16	6.86		
	$+ a_2 \ln(CP)$	1.81	7.35		
$\ln(B)$	a_0	-8.57	4.62	0.63	0.28
	$+ a_1 \ln(PAM)$	1.17	5.24		
	$+ a_2 \ln(CP)$	0.95	3.88		
$\ln(B)$	a_0	-8.49	4.59	0.63	0.28
	$+ a_1 \ln(CP.PAM)$	1.07	7.10		
$\ln(B)$	a_0	-4.57	2.81	0.67	0.20
	$+ a_1 \ln(PAM)$	1.31	5.66		
$\ln(B)$	a_0	-1.18	1.07	0.70	0.13
	$+ a_1 \ln(CP)$	1.10	4.37		
$\ln(B)$	a_0	-1.18	50.32	0.71	0.09
	$+ a_1 \ln(PAM/EAM)$	1.10	3.66		

Tableau 8.7: Relations entre paramètre transformé $\ln(B)$ et variables explicatives

Parmi les relations proposées au Tableau 8.7 pour le paramètre B , celle avec CP et PAM semble la plus satisfaisante parmi les formules les plus simples:

$$\ln(B) = -8.49 + 1.07 \cdot \ln(CP \cdot PAM) \quad \text{Eq. (8.16)}$$

soit $B = \frac{(CP \cdot PAM)^{1.07}}{4866}$ Eq. (8.17)

Ici, contrairement au cas précédent du paramètre A , les *horsains* semblent limiter la valeur des coefficients de détermination, alors qu'il existe une tendance plus nette du nuage, comme le montre la Figure 8.4.

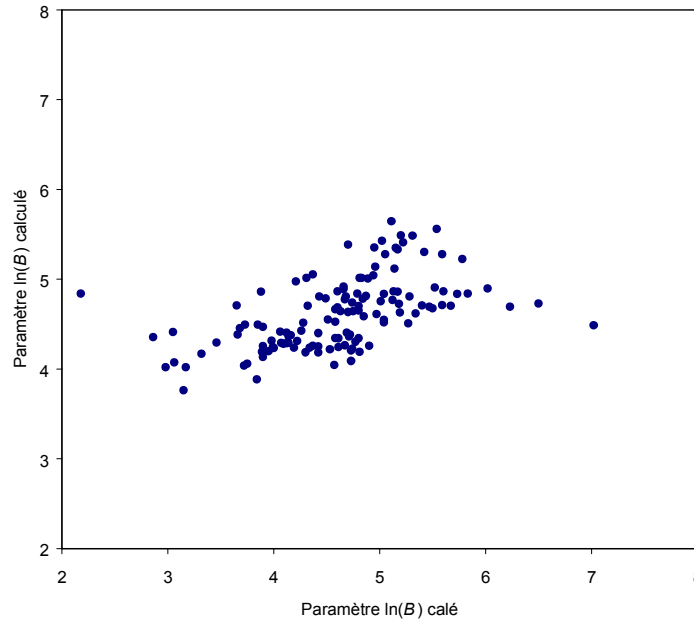


Figure 8.4: Graphe de corrélation entre valeurs transformées de B calées et calculées par l'équation de prédétermination

Paramètre C

Paramètre	Formule de régression	Coefficients de régression	Rapport de Student	Erreur standard	Coefficient de détermination
ln(C-0,5)	a_0	7.13	6.36	0.36	0.59
	$+ a_1 \ln(CP)$	-0.56	3.57		
	$+ a_2 \ln(PAM)$	-0.65	5.11		
	$+ a_3 \ln(SF)$	0.13	8.72		
ln(C-0,5)	a_0	4.65	5.06	0.38	0.55
	$+ a_1 \ln(PAM)$	-0.74	5.63		
	$+ a_2 \ln(SF)$	0.16	11.20		
ln(C-0,5)	a_0	3.40	3.66	0.39	0.51
	$+ a_1 \ln(CP)$	-0.71	4.22		
	$+ a_2 \ln(SF)$	0.13	7.57		
ln(C-0,5)	a_0	4.01	13.50	0.38	0.55
	$+ a_1 \ln(SF/PAM^4)$	0.16	12.50		
ln(C-0,5)	a_0	-0.50	5.64	0.42	0.44
	$+ a_1 \ln(SF)$	0.16	10.01		
ln(C-0,5)	a_0	7.10	7.49	0.47	0.29
	$+ a_1 \ln(CP)$	-1.29	7.17		
ln(C-0,5)	a_0	5.39	4.20	0.53	0.11
	$+ a_1 \ln(PAM)$	-0.73	3.96		

Tableau 8.8: Relations entre paramètre transformé ln(C-0,5) et variables explicatives

De nombreuses relations (voir Tableau 8.8) ont pu être établies pour le paramètre C transformé. Ce sont les trois variables SF , PAM et CP qui permettent d'apporter les meilleures explications du paramètre C . Une bonne partie de la variance de ce paramètre semble pouvoir être logiquement expliquée par la superficie SF du bassin. La relation simple regroupant SF et PAM , illustrée à la Figure 8.5 a été retenue:

$$\ln(C - 0,5) = 4,01 + 0,16 \cdot \ln(SF / PAM^4) \quad \text{Eq. (8.18)}$$

$$\text{soit } C = 0,5 + 55,1 \cdot \left(\frac{SF}{PAM^4} \right)^{0,16} \quad \text{Eq. (8.19)}$$

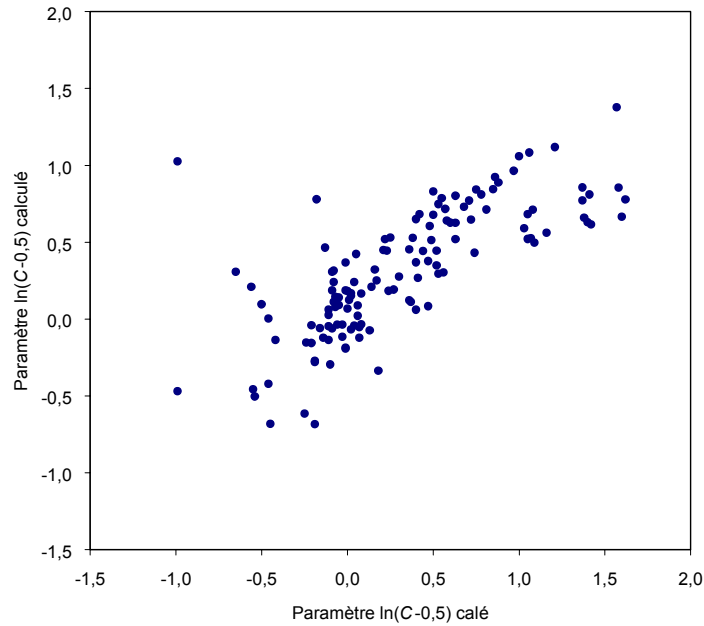


Figure 8.5: Graphe de corrélation entre valeurs transformées de C calées et calculées par l'équation de prédétermination

Paramètre D

Les deux seuls ajustements obtenus pour le paramètre D sont très faiblement significatifs (voir Tableau 8.9). La Figure 8.6 montre l'absence de tendance générale, le coefficient de détermination non nul étant seulement dû à la présence de quelques points excentrés du nuage. Par conséquent, comme dans le cas du paramètre A , aucune relation n'a été retenue ici. Seule la valeur moyenne du paramètre sera utilisée.

Paramètre	Formule de régression	Coefficients de régression	Rapport de Student	Erreur standard	Coefficient de détermination
argsh(D)	a_0 + $a_1 \ln(PAM/EAM)$	-0.87 0.82	5.64 2.53	1.24	0.05
argsh(D)	a_0 + $a_1 \ln(EAM)$	11.14 -1.76	3.42 3.60	1.22	0.09

Tableau 8.9: Relations entre paramètre transformé argsh(D) et variables explicatives

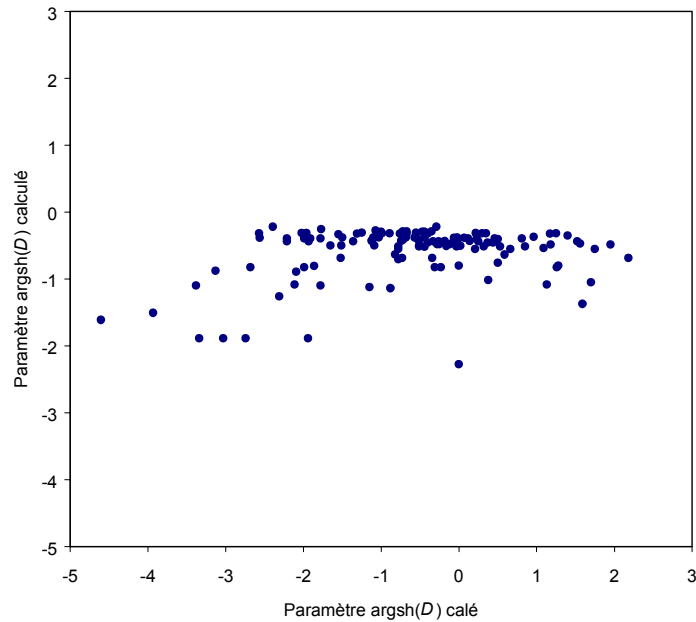


Figure 8.6: Graphe de corrélation entre valeurs transformées de D calées et calculées par l'équation de prédétermination

8.5.3.3. Régressions sur les paramètres réels

Comme précédemment pour les paramètres transformés, nous donnons maintenant les résultats des régressions multiples obtenues sur l'échantillon des paramètres réels. Il est important de noter qu'ici, les *horsains* (voir par exemple Figure 8.7) influencent davantage les coefficients de détermination des régressions que dans le cas précédent, du fait que les paramètres réels varient sur plusieurs ordres de grandeur. Les paramètres ont été reliés à des fonctions puissance des variables, les exposants ayant été choisis dans le seul souci d'efficacité.

Paramètre A

Le Tableau 8.10 récapitule les caractéristiques des meilleures relations obtenues pour le paramètre A . Si quelques-unes des relations proposées peuvent paraître satisfaisantes quant à la valeur du coefficient de détermination, le tracé graphique de ces relations révèle que cela ne provient en fait que de la position excentrée de quelques points, comme cela est illustré à la Figure 8.7 pour la première relation du Tableau 8.10.

Paramètre	Formule de régression	Coefficients de régression	Rapport de Student	Erreur standard	Coefficient de détermination
A	a_0	-1880.00	4.24		
	$+ a_1 SF^{0.5}$	7.22	7.59	363.10	0.38
	$+ a_2 CP^{0.3}$	454.00	5.04		
A	a_0	383.50	9.68	394.20	0.26
	$+ a_1 SF^{0.6}$	2.49	6.75		
A	a_0	-1450.00	2.75	435.60	0.10
	$+ a_1 CP^{0.3}$	403.00	3.73		
A	a_0	1180.00	5.20	443.70	0.06
	$+ a_1 EAM^{-1}$	-507400.00	2.97		

Tableau 8.10: Relations entre paramètre A et variables explicatives

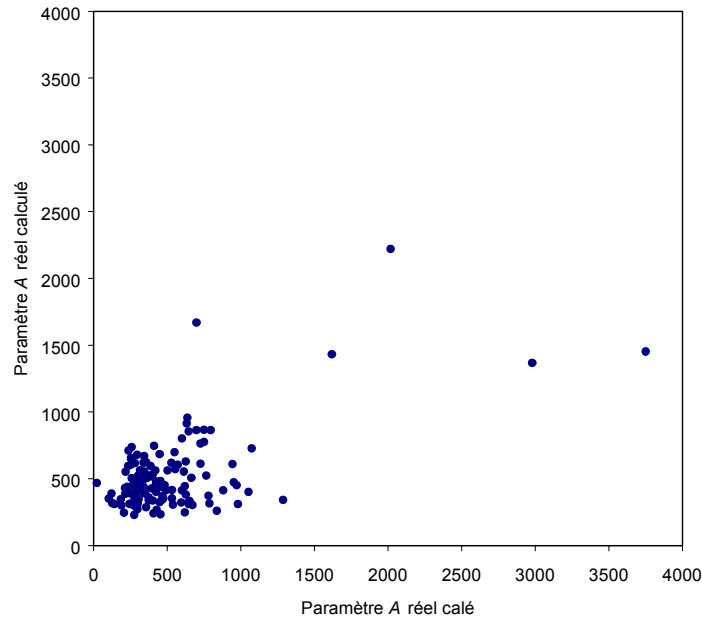


Figure 8.7: Graphe de corrélation entre valeurs réelles de A calées et calculées par l'équation de prédétermination

Devant la faiblesse des ajustements, nous avons préféré, comme dans le cas des régressions sur le paramètre A transformé, nous en tenir à la prise en compte de la valeur moyenne du paramètre.

Paramètre B

Paramètre	Formule de régression	Coefficients de régression	Rapport de Student	Erreur standard	Coefficient de détermination
B	a_0	-129.00	1.17		
	$+ a_1 CP^{-1,5}$	-432700.00	4.33	117.40	0.17
	$+ a_2 (PAM/EAM)^{0,3}$	384.90	3.77		
B	a_0	333.90	4.96	124.10	0.07
	$+ a_1 CP^{-0,8}$	-13500.00	3.01		
B	a_0	-520.00	2.42	124.00	0.07
	$+ a_1 PAM^{0,2}$	161.20	3.04		
B	a_0	29.10	0.72	124.90	0.05
	$+ a_1 (PAM/EAM)$	71.30	2.70		
B	a_0	-1340.00	4.19	118.90	0.14
	$+ a_1 (CP.PAM/EAM)^{0,1}$	839.00	4.61		

Tableau 8.11: Relations entre paramètre B et variables explicatives

Parmi les ajustements du Tableau 8.11, l'équation utilisant la variable composée de CP , PAM et EAM a été retenue pour le paramètre B :

$$B = -1340 + 839 \cdot \left(\frac{CP \cdot PAM}{EAM} \right)^{0,1} \tag{Eq. (8.20)}$$

Comme dans le cas du paramètre transformé, la présence de *horsains* limite la valeur du coefficient de détermination (voir Figure 8.8).

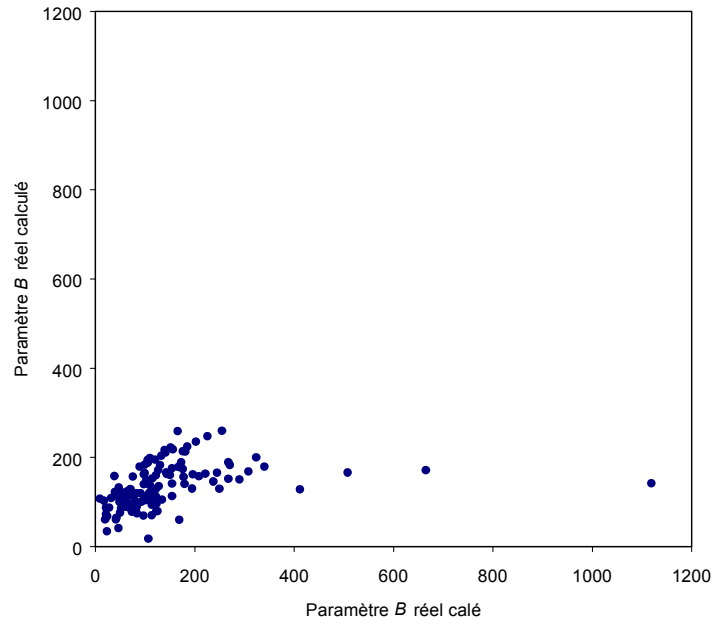


Figure 8.8: Graphe de corrélation entre valeurs réelles de B calées et calculées par l'équation de prédétermination

Paramètre C

Paramètre	Formule de régression	Coefficients de régression	Rapport de Student	Erreur standard	Coefficient de détermination
C	a_0	-16.64	6.22		
	$+ a_1 PAM^{-0,1}$	26.62	5.06	0.75	0.48
	$+ a_2 SF^{0,05}$	4.23	9.58		
C	a_0	-2.13	3.30		
	$+ a_1 CP^{-0,5}$	38.70	4.14	0.78	0.44
	$+ a_2 SF^{0,16}$	0.59	7.38		
C	a_0	0.26	1.48	0.72	0.51
	$+ a_1 (SF/PAM^5)^{0,15}$	150.00	11.60		
C	a_0	-1.38	3.44	0.81	0.38
	$+ a_1 SF^{0,08}$	2.27	8.84		
C	a_0	-2.21	2.90	0.92	0.20
	$+ a_1 CP^{-0,5}$	60.25	5.70		
C	a_0	-3.65	2.60	0.97	0.12
	$+ a_1 EAM^{-0,3}$	42.39	4.10		

Tableau 8.12: Relations entre paramètre C et variables explicatives

Des relations données au Tableau 8.12, celle associant les variables SF et PAM a été choisie pour le paramètre C :

$$C = 0,26 + 150. \left(\frac{SF}{PAM^5} \right)^{0,15} \quad \text{Eq. (8.21)}$$

Cette association, bien qu'insolite, permet de limiter le nombre de coefficients de l'équation de régression. La représentation graphique donnée à la Figure 8.9 est relativement satisfaisante.

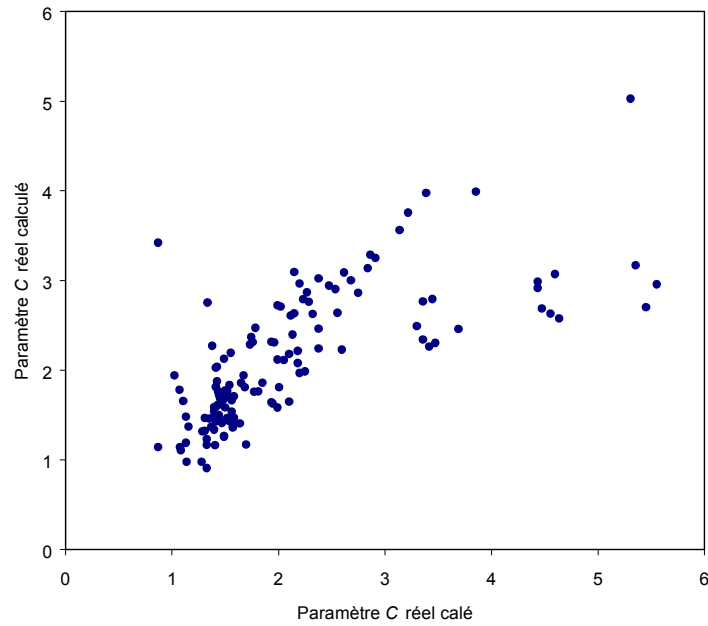


Figure 8.9: Graphe de corrélation entre valeurs réelles de C calées et calculées par l'équation de prédétermination

Paramètre D

Les régressions établies pour le paramètre D sont données au Tableau 8.13. On rencontre un problème identique à celui observé avec les régressions sur les paramètres transformés, à savoir la faible significativité des ajustements, la valeur du coefficient de détermination étant rehaussée par la présence de points hors du nuage principal (voir la Figure 8.10 dans le cas de la première équation du Tableau 8.13). Nous avons donc choisi de ne retenir pour D que la valeur moyenne

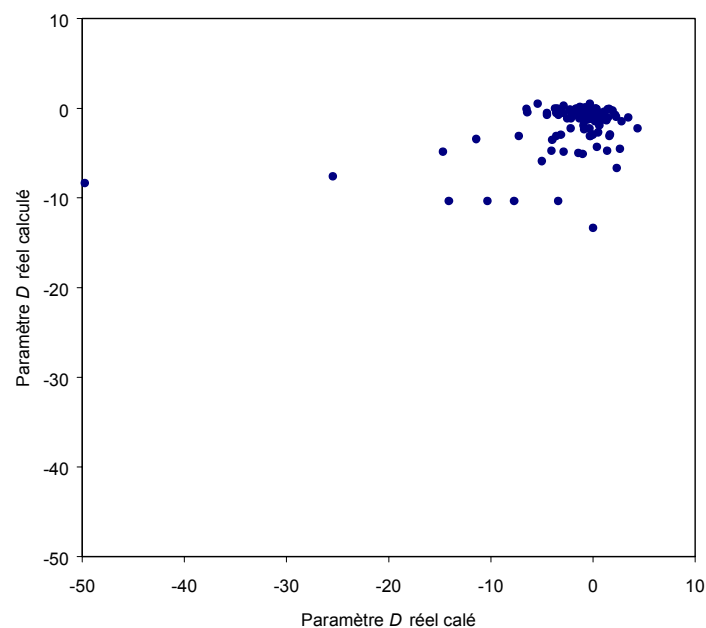


Figure 8.10: Graphe de corrélation entre valeurs réelles de D calées et calculées par l'équation de prédétermination

Paramètre	Formule de régression	Coefficients de régression	Rapport de Student	Erreur standard	Coefficient de détermination
D	a_0	33.40	5.40	4.96	0.20
	$+ a_1 EAM^{0.3}$	-4.75	5.69		
D	a_0	3.92	2.81	5.20	0.13
	$+ a_1 (PAM/EAM)^{-0.8}$	-7.06	4.28		

Tableau 8.13: Relations entre paramètre D et variables explicatives

Cette première étape de recherche d'équations pour relier les paramètres du modèle à des descripteurs simples du bassin n'a permis de proposer des relations satisfaisantes que pour les deux paramètres de routage B et C . Pour les deux paramètres de rendement (A et D), la faiblesse des relations obtenues, que ce soit sur les paramètres transformés ou réels, n'autorise pas leur exploitation et nous préférons donc leur affecter les valeurs moyennes du Tableau 8.3, c'est-à-dire des équations de régression se réduisant à un terme constant. Ces problèmes d'explication des paramètres de rendement du modèle journalier rejoignent ceux exposés dans les travaux antérieurs cités précédemment.

Les relations précédentes ne sont pas exploitables en l'état dans des applications de régionalisation. Elles nous serviront ici seulement à tester la méthode proposée. Une des raisons de ce faible degré de satisfaction est que les variables explicatives utilisées ne renseignent que sur les entrées du système et pas sur le bassin lui-même. Même s'il existe des relations entre ces deux aspects, les caractéristiques pédologiques et géologiques du bassin pourraient apporter des compléments d'information significatifs pour l'explication des paramètres.

8.5.4. Résultats de l'application de la méthode proposée

Dans les paragraphes suivants, nous allons appliquer la méthode présentée en début de chapitre en distinguant plusieurs phases:

- test des relations retenues, pour déterminer les performances du modèle qui auraient été obtenues sur les 131 bassins par une méthode classique de régionalisation;
- détermination du jeu de paramètres constants (unique pour l'ensemble des bassins) permettant d'obtenir les résultats les plus satisfaisants, ceci fournissant une référence pour estimer l'intérêt d'utiliser des formules de prédétermination des paramètres;
- recherche des valeurs des coefficients des régressions permettant d'obtenir les meilleurs résultats du modèle. Nous nous sommes placés ici dans trois cas différents, suivant que l'estimation des paramètres est réalisée à partir des régressions établies sur les paramètres transformés, sur les paramètres réels, ou de manière combinée à partir de l'une ou l'autre de ces solutions suivant les paramètres;
- vérification des relations établies sur l'échantillon des 429 bassins.

L'évaluation des résultats a été réalisée sur les distributions des performances sur l'échantillon des 131 bassins, préférées aux performances moyennes du modèle qui sont, sur ce faible nombre de tests, peu significatives à cause des valeurs fortement négatives du critère. La sélection a donc été réalisée à partir des quantiles de distribution, donnant ainsi des résultats plus satisfaisants. L'utilisation d'un seul de ces quantiles n'est pas très fiable, du fait de l'entrecroisement possible des courbes de distribution. Nous avons donc choisi de prendre en compte ici plusieurs quantiles qui correspondent aux proportions de valeurs du critère respectivement supérieures à 60, 70, 75 et 80 %. Une combinaison linéaire de ces quantiles a

été prise comme critère de sélection des solutions les plus satisfaisantes, c'est-à-dire en fait comme fonction objectif à minimiser dans notre procédure d'optimisation manuelle. Ce critère arbitrairement choisi ici et qui ne se veut pas de portée générale, est donné par:

$$Cr = \frac{1}{8}(2.F(60) + 3.F(70) + 2.F(75) + F(80)) \quad \text{Eq. (8.22)}$$

où $F(X)$ est la fréquence au non dépassement pour la valeur X du critère de Nash. Un jeu de coefficients de régression était donc préféré à un autre dans la mesure où il obtenait une plus forte valeur de ce critère composite. Dans la suite, en plus des valeurs de ce critère, nous utilisons les représentations graphiques des distributions (construites à partir de quelques points) pour illustrer les résultats. Remarquons que lorsque les paramètres sont calés, le critère donné à l'Eq. (8.22) est nul sur l'échantillon des 131 bassins.

Nous sommes partis de quatre solutions suivantes (détaillées dans le Tableau 8.14) pour estimer la valeur des paramètres:

- solution 1: paramètres moyens calculés sur l'échantillon des paramètres transformés,
- solution 2: idem avec utilisation des équations de prédétermination établies sur les paramètres transformés pour B et C ,
- solution 3: paramètres moyens calculés sur l'échantillon des paramètres réels,
- solution 4: idem avec utilisation des équations de prédétermination établies sur les paramètres réels pour B et C ,

	Paramètre A	Paramètre B	Paramètre C	Paramètre D
Solution 1	415,7	101,5	1,86	-0,62
Solution 2	415,7	$\frac{(CP.PAM)^{1,07}}{4866}$	$0,5 + 55,1 \cdot \left(\frac{SF}{PAM^4}\right)^{0,16}$	-0,62
Solution 3	515,6	133,4	2,10	-1,72
Solution 4	515,6	$-1340 + 839 \cdot \left(\frac{CP.PAM}{EAM}\right)^{0,1}$	$0,26 + 150 \cdot \left(\frac{SF}{PAM^5}\right)^{0,15}$	-1,72

Tableau 8.14: Solutions de prédétermination des paramètres

On peut ensuite essayer toutes les combinaisons possibles des paramètres A , B , C et D en choisissant l'une des quatre solutions. Ce sont donc au total 4^4 soit 256 possibilités qui peuvent être testées. Après une première évaluation de ces solutions initiales, on peut changer les différentes possibilités d'estimation des paramètres, en modifiant notamment les valeurs des paramètres constants et celles des coefficients des équations de prédétermination. On détermine ainsi les coefficients les plus adaptés et les meilleures solutions combinées.

8.5.4.1. Vérification directe

Le premier test réalisé correspond à l'étape (v) de l'approche classique, c'est-à-dire à la simulation des débits sur les bassins de l'échantillon en utilisant les équations de prédétermination. Nous avons donc testé les quatre solutions initiales proposées dans le Tableau 8.14.

Nous avons tracé à la Figure 8.11 les distributions des résultats obtenus avec ces quatre solutions. Pour référence, nous avons également tracé la distribution des performances du modèle obtenues en calage. Les valeurs correspondantes du critère Cr sont données au Tableau 8.15.

	Solution 1	Solution 2	Solution 3	Solution 4
Cr	0,275	0,244	0,312	0,295

Tableau 8.15: Valeur du critère Cr pour les quatre solutions initiales

Des quatre solutions de base, la deuxième faisant intervenir les équations de prédétermination de B et C établies sur les paramètres transformés apparaît la plus satisfaisante. Les solutions 3 et 4, déterminées à partir des paramètres réels, donnent de moins bons résultats. La solution 1 qui utilise des paramètres constants paraît assez satisfaisante. Ce premier test souligne le fait que le choix a priori de la forme sous laquelle doivent être utilisés les paramètres du modèle pour établir les relations de prédétermination ne va pas de soi. Ici, il semble plus judicieux d'utiliser les paramètres transformés, c'est-à-dire la forme des paramètres utilisée par l'algorithme d'optimisation.

Par ailleurs, l'utilisation de valeurs constantes pour tous les paramètres (solutions 1 et 3) peut fournir de relativement bons résultats, en particulier pour la solution 1. Cependant, les valeurs moyennes ne sont pas forcément les mieux adaptées. Nous chercherons donc dans la partie suivante à déterminer les valeurs des paramètres constants conduisant aux meilleurs résultats.

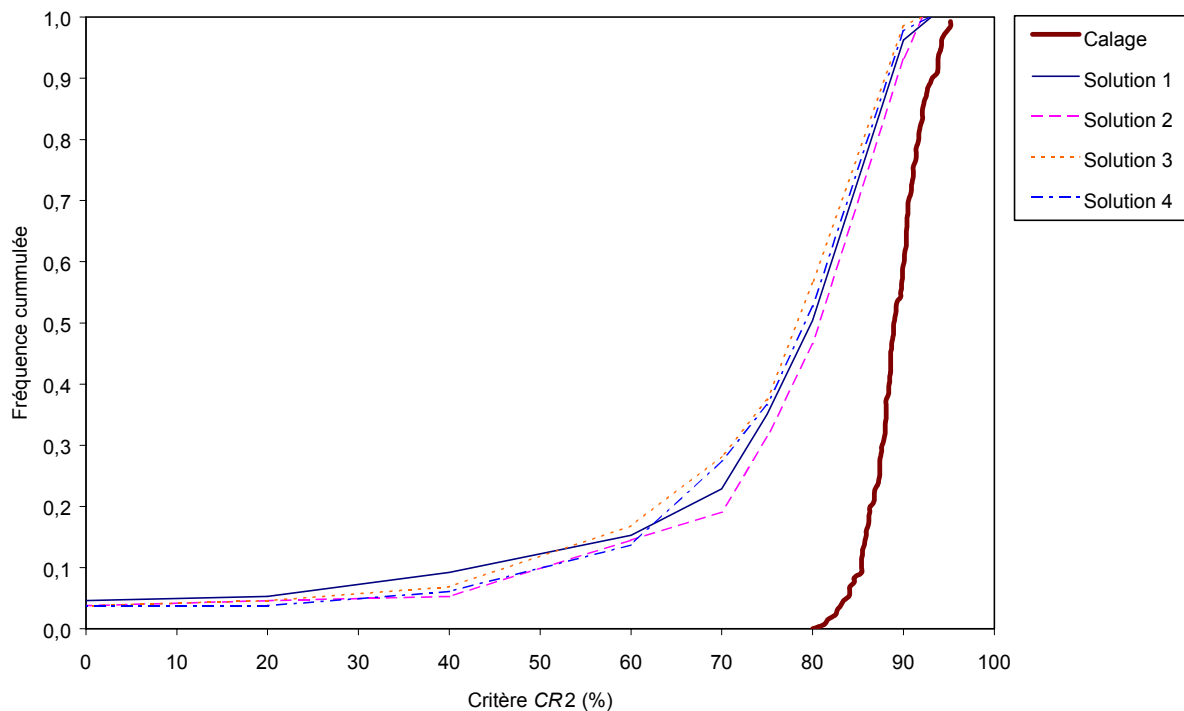


Figure 8.11: Distributions des résultats obtenus avec les paramètres calés, constants et prédéterminés

8.5.4.2. Recherche des coefficients constants les plus efficaces

Par la procédure itérative proposée, nous avons cherché à obtenir les résultats les plus satisfaisants du modèle avec des paramètres constants sur notre échantillon de 131 bassins. Le quadruplet (A, B, C, D) obtenu est donné au Tableau 8.16. Ces valeurs sont assez proches de celles de la solution 1 (415,7, 101,5, 1,86, -0,62), sauf dans le cas de A , dont la valeur est intermédiaire entre les solutions 1 et 3. La Figure 8.12 montre que la distribution correspondante (solution (1-3)') est meilleure que celles obtenues avec les paramètres moyens des solutions 1 et 3. La solution de prendre des paramètres moyens n'est donc pas optimale. Les gains par rapport à la solution 1 restent cependant relativement modestes.

	Paramètre <i>A</i>	Paramètre <i>B</i>	Paramètre <i>C</i>	Paramètre <i>D</i>	<i>Cr</i>
Solution (1-3)'	460	110	1,78	-0,58	0,249

Tableau 8.16: Solution après ajustement des paramètres constants

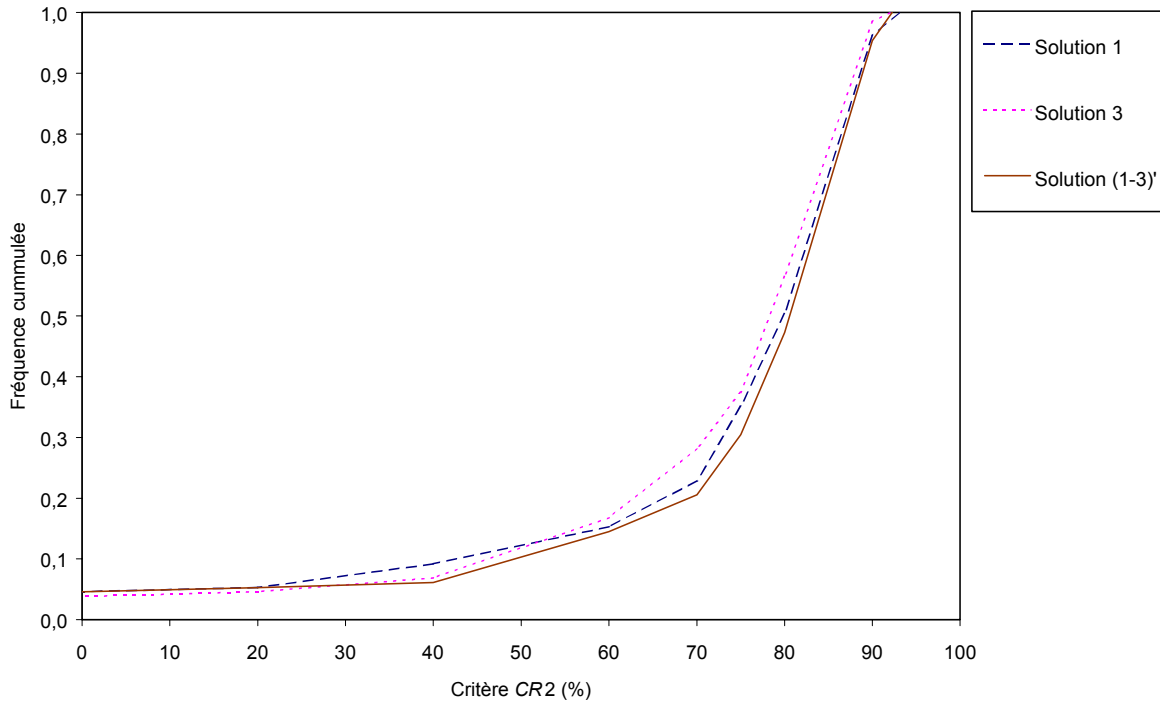


Figure 8.12: Distributions des résultats obtenus avec des paramètres constants

8.5.4.3. Recherche de coefficients de régression plus efficaces

Nous examinons ici s'il est possible d'améliorer les résultats obtenus avec les solutions 2 et 4, en appliquant la méthode proposée. Nous essayons également de voir si la combinaison d'équations des solutions 2 et 4 permet d'améliorer les résultats.

A chaque fois, nous partons des paramètres constants pour *A* et *D* et des équations de prédétermination des paramètres pour *B* et *C*, et nous appliquons la méthode en faisant varier les valeurs des constantes et des coefficients des régressions. Nous nous trouvons en fait confrontés ici à un problème d'optimisation dans un espace à six dimensions (les 2 paramètres constants et les quatre coefficients des équations de régression sont à optimiser). Il s'agit donc d'une problématique complexe, mais nous sommes aidés dans notre démarche par le fait que l'optimum recherché se situe probablement dans le voisinage de la solution initiale.

Cas 1: Paramètres déterminés selon les relations établies sur les paramètres transformés (solution 2)

L'application de la méthode dans ce premier cas de figure nous a conduit à la solution 2' donnée au Tableau 8.17. Cette solution est peu différente de la solution initiale. Les modifications principales ont été effectuées sur les paramètres constants. L'équation de prédétermination de *B* est restée inchangée.

	Paramètre <i>A</i>	Paramètre <i>B</i>	Paramètre <i>C</i>	Paramètre <i>D</i>	<i>Cr</i>
Solution 2'	535	$\frac{(CP \cdot PAM)^{1,07}}{4866}$	$0,5 + 47,5 \cdot \left(\frac{SF}{PAM^4}\right)^{0,16}$	-0,40	0,221

Tableau 8.17: Solution 2' après ajustement des coefficients

La Figure 8.13 compare les distributions des résultats pour les deux solutions 2 et 2'. Comme dans le cas des paramètres constants, il existe une amélioration des performances, indiquant que la solution 2 n'est pas optimale. Les progrès restent cependant ici assez modestes.

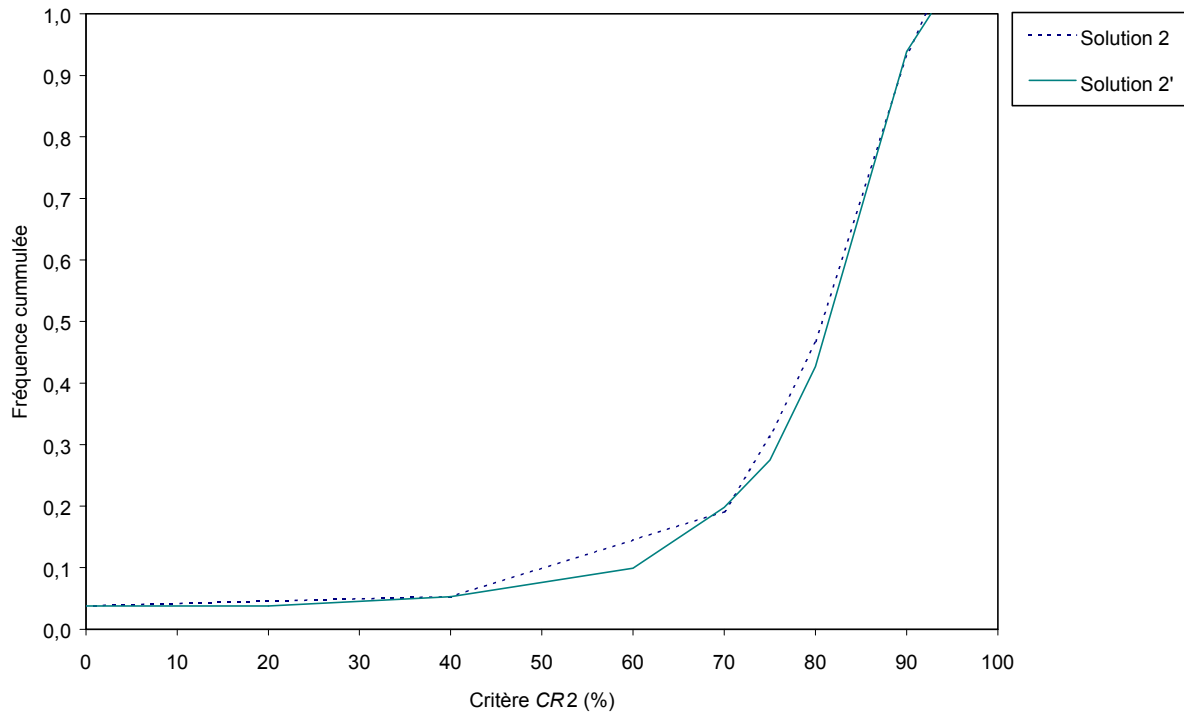


Figure 8.13: Distributions des résultats avant et après ajustement des coefficients des formulations de prédétermination de la solution 2

Cas 2: Paramètres déterminés selon les relations établies sur les paramètres transformés (solution 2)

Nous sommes partis de la solution 4 et nous avons appliqué la procédure d'ajustement des coefficients des formules de prédétermination. Nous avons obtenu la solution 4' dont les caractéristiques sont données au Tableau 8.18. Dans ce cas, la différence entre la solution initiale et la solution avec coefficients ajustés est plus marquée que dans le cas précédent. On peut remarquer que la valeur du paramètre *D* correspond à la valeur moyenne issue de l'échantillon des paramètres transformés.

	Paramètre <i>A</i>	Paramètre <i>B</i>	Paramètre <i>C</i>	Paramètre <i>D</i>	<i>Cr</i>
Solution 4	495	$-1150 + 720 \cdot \left(\frac{CP \cdot PAM}{EAM}\right)^{0,1}$	$0,6 + 120 \cdot \left(\frac{SF}{PAM^5}\right)^{0,15}$	-0,62	0,234

Tableau 8.18: Solution 4' après ajustement des coefficients

Les améliorations constatées en terme de performances du modèle sont plus sensibles que dans le cas précédent (solution 2 et 2'), comme le montre la Figure 8.14.

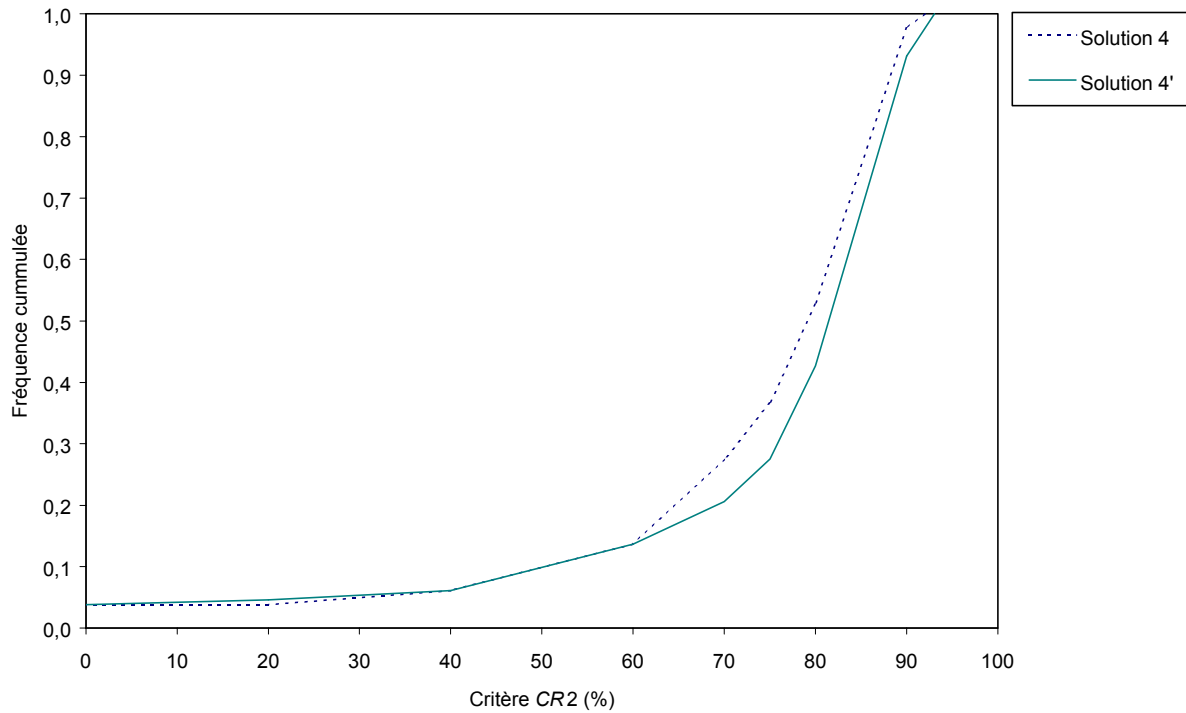


Figure 8.14: Distributions des résultats avant et après ajustement des coefficients des formulations de prédétermination de la solution 4

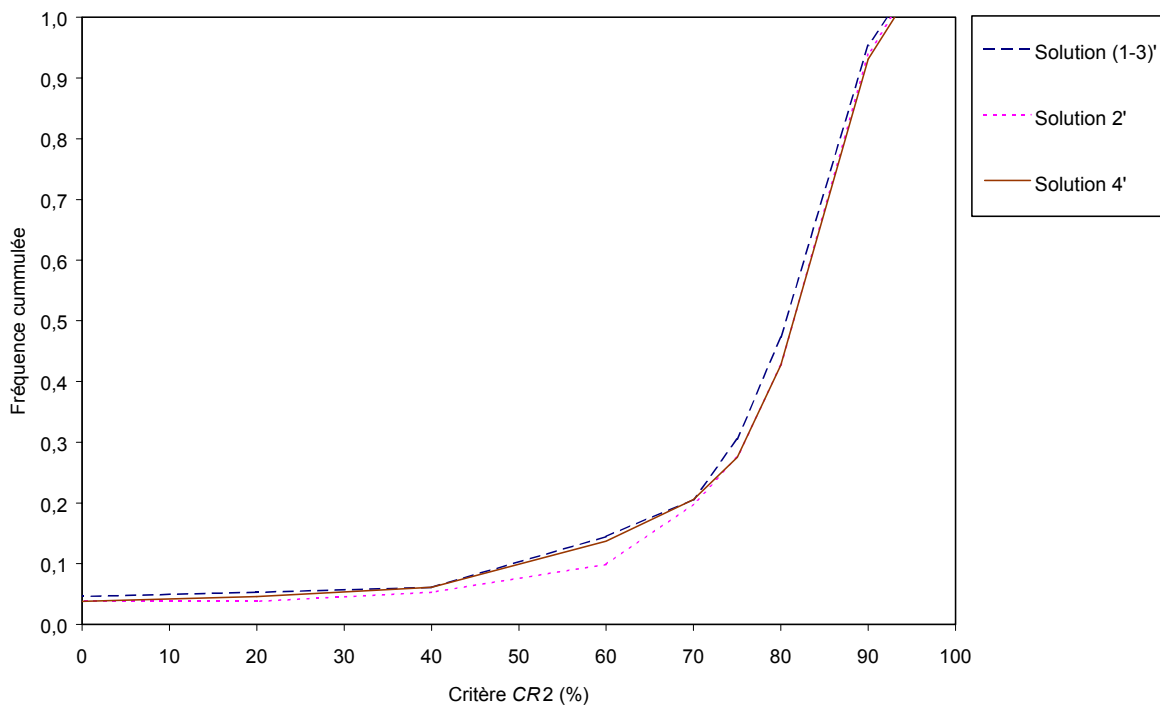


Figure 8.15: Comparaison des distributions des résultats avec les solutions de prédétermination après ajustement des coefficients, sur l'échantillon des 131 bassins

La Figure 8.15 compare, sur les 131 bassins, les différentes solutions ajustées que nous venons d'établir. Elles sont beaucoup plus proches les unes des autres que ne l'étaient les

solutions initiales (voir Figure 8.11). La meilleure reste, avec un faible avantage, celle établie sur les paramètres transformés (solution 2').

Cas 3: Paramètres déterminés selon les relations établies sur les paramètres réels ou transformés

La méthode proposée laisse la possibilité de choisir des formulations combinées pour les paramètres. En maintenant toujours A et D constants, B et C peuvent être déterminés suivant les équations établies sur les paramètres réels pour l'un et sur ceux transformés pour l'autre. Les tests que nous avons effectués ici, notamment en combinant les solutions établies précédemment, n'ont pas permis d'obtenir de meilleurs résultats que ceux présentés à la Figure 8.15. Nous nous contenterons donc dans la suite d'utiliser les cas simples où B et C sont prédéterminés tous les deux soit à partir des relations établies sur les paramètres transformés, soit à partir de celles établies sur les paramètres réels.

8.5.4.4. Vérification sur l'échantillon des 429 bassins

Nous avons réalisé une comparaison des solutions initiales et des solutions après ajustement sur l'ensemble des 429 bassins versants, en appliquant ces solutions de prédétermination des paramètres à l'ensemble des 1284 bassins-périodes utilisées dans les tests de comparaison du chapitre 5. Ceci permet de vérifier la robustesse des relations établies sur un échantillon élargi de bassins par rapport à celui utilisé pour établir les relations. La Figure 8.16 montre les distributions des résultats obtenus en utilisant les solutions initiales de prédétermination. Comme dans le cas de l'échantillon des 131 bassins, c'est la solution 2 établie sur les paramètres transformés, qui est la plus satisfaisante. Cependant, on peut remarquer qu'ici, la distribution correspondant à la solution 1 (avec paramètres constants) est très peu différente de celle obtenue avec la solution 2.

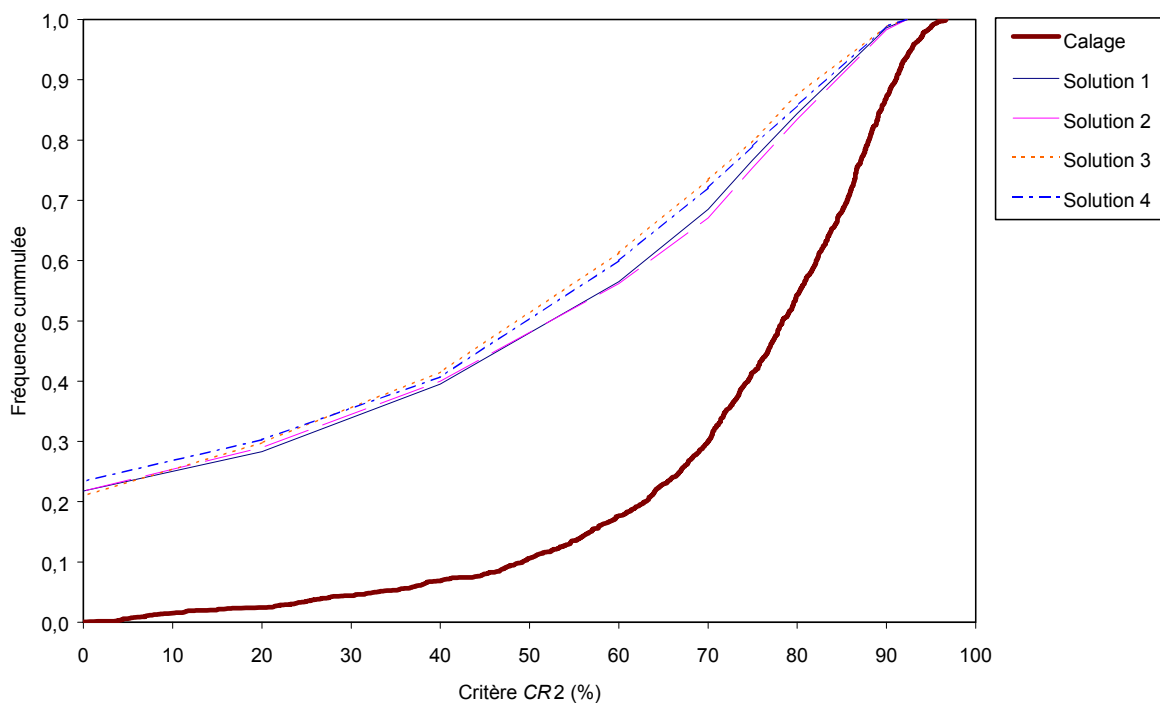


Figure 8.16: Distributions des résultats obtenus sur les 429 bassins en utilisant les solutions initiales de prédétermination

Ceci est confirmé au Tableau 8.19 par les valeurs très proches du critère Cr pour ces deux solutions (pour référence, la valeur de Cr lorsque les paramètres sont calés est de 0,328).

	Solution 1	Solution 2	Solution 3	Solution 4	Solution 2'	Solution (1-3)'	Solution 4'
Cr	0,695	0,685	0,738	0,725	0,693	0,688	0,684

Tableau 8.19: Valeurs du critère Cr pour les différentes solutions sur l'échantillon des 429 bassins

Comme cela a été fait à la Figure 8.15 pour les 131 bassins, nous donnons à la Figure 8.17 les distributions des résultats obtenus sur l'échantillon total de bassins pour les solutions de prédétermination après ajustement. A titre de référence, nous avons aussi inclus la meilleure distribution (solution 2) de la Figure 8.16. Les trois solutions ajustées obtiennent ici des résultats très similaires, avec des distributions très peu différentes les unes des autres. Elles sont également très voisines de la distribution de la solution 2 initiale. L'écart entre les différentes solutions n'apparaît pas suffisamment significatif pour être interprété de façon consistante.

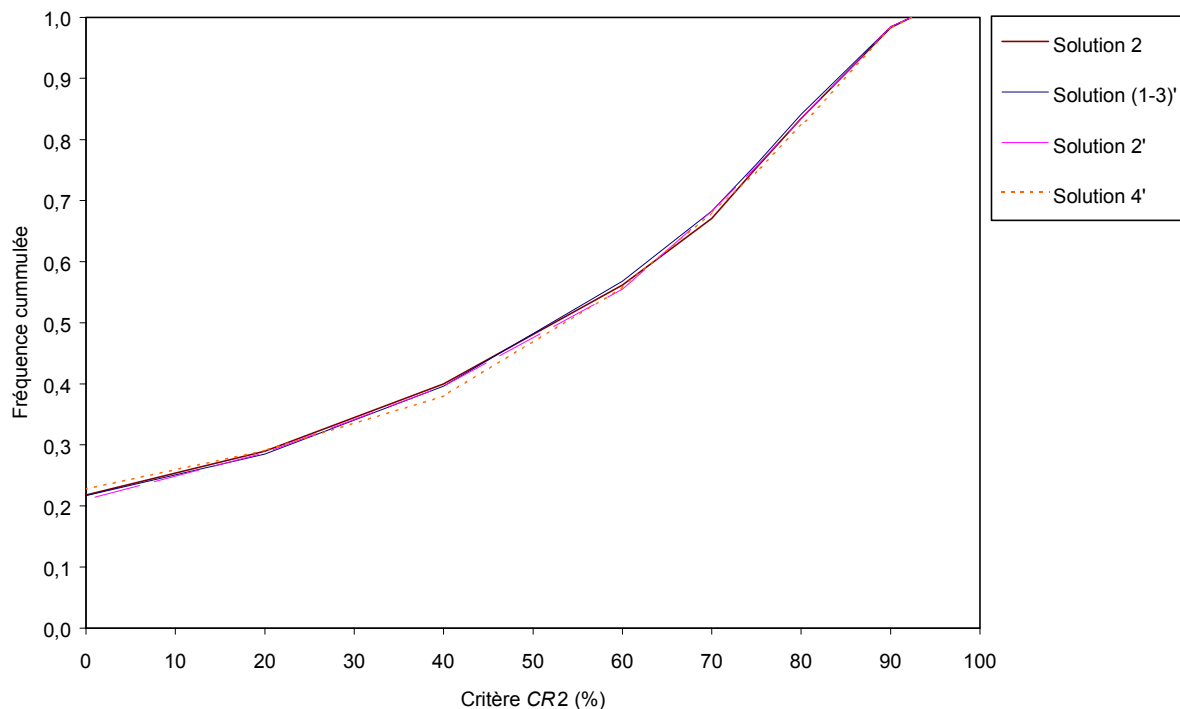


Figure 8.17: Distributions des résultats obtenus sur les 1284 périodes de calage après ajustement des formules de prédétermination

8.5.5. Discussion

Au travers des résultats exposés précédemment se dégagent quelques points importants:

- L'application de la méthode que nous nous sommes proposés de tester a démontré que la démarche classique de régionalisation des paramètres n'est pas optimale pour maximiser les performances du modèle sur les échantillons de bassins utilisés pour établir les équations de prédétermination des paramètres. Que ce soit dans le cas des relations établies sur paramètres constants, transformés ou réels, il a été possible à chaque fois d'identifier des relations permettant d'obtenir de meilleures performances du modèle. Ceci

indique donc que l'on réussit ainsi à extraire une information plus pertinente de descripteurs simples du bassin que dans les démarches habituelles.

- Si cette amélioration est sensible dans le cas des relations sur paramètres réels, elle est cependant apparue plus faible avec les relations sur paramètres transformés. Les progrès restent d'envergure limitée dans le cas des équations initiales de régression les plus satisfaisantes (régression sur paramètres transformés). Ils sont plus importants dans les autres cas. La méthode d'ajustement permet d'obtenir des solutions grossièrement équivalentes dans les trois cas étudiés, un léger avantage se maintenant pour la solution sur les paramètres transformés.
- L'application de la méthode, par une démarche d'essai-erreur, est assez longue en dépit de la simplicité du cas étudié (modèle à quatre paramètres, équations simples de régression). Il est en effet difficile d'identifier manuellement la zone de convergence dans l'espace à six dimensions dans lequel nous avons travaillé. D'ailleurs, nous ne pouvons assurer que les solutions que nous avons proposées sont optimales. L'automatisation de la procédure d'optimisation serait souhaitable, mais les besoins en calcul sont très importants dans le cas d'un large échantillon comme celui utilisé ici.
- Lorsque l'on essaie de valider les relations sur un échantillon de bassins plus large, en incluant des bassins n'ayant pas servi à l'établissement des équations de régression, on s'aperçoit qu'après ajustement, les solutions obtenues dans les trois cas donnent des résultats quasiment identiques. Elles ne montrent pas d'amélioration significative par rapport à la meilleure solution initiale établie sur les paramètres transformés.

Sur l'échantillon des 131 bassins, le faible avantage des solutions régionalisées par rapport à celle utilisant des paramètres constants peut être rapproché des observations formulées par Seibert (1999): sur onze bassins du centre de la Suède, l'utilisation de formules régionalisées pour les paramètres donne des résultats à peine meilleurs que l'utilisation d'un débit calculé comme la moyenne arithmétique des débits sur ces onze bassins. Cette situation s'explique en partie par le fait que, si les rapports de Student obtenus lors de l'établissement des formules de régression apparaissaient acceptables, les erreurs standard de ces relations indiquent que la confiance qui peut leur être accordée est limitée. Cette situation est d'autant plus vraie sur l'échantillon total des 429 bassins.

Elle indique par ailleurs que sur notre échantillon, l'utilisation des seuls descripteurs climatiques retenus ici avec la superficie du bassin (ou la forme sous laquelle ils ont été utilisés) ne permet pas d'extraire une information suffisamment pertinente pour renseigner le modèle sur les caractéristiques de la transformation pluie-débit

Dans une dernière étape de ce travail, nous avons voulu comparer les résultats obtenus par le modèle GR4J' avec des paramètres constants et les résultats obtenus par les modèles testés dans la comparaison. La Figure 8.18 montre que les résultats obtenus sont nettement meilleurs que ceux obtenus par le modèle de référence TSYK et que l'on réussit avec le modèle GR4J' sans paramètres optimisés (paramètres fixés aux valeurs médianes sur l'échantillon des 429 bassins) à obtenir des résultats relativement satisfaisants.

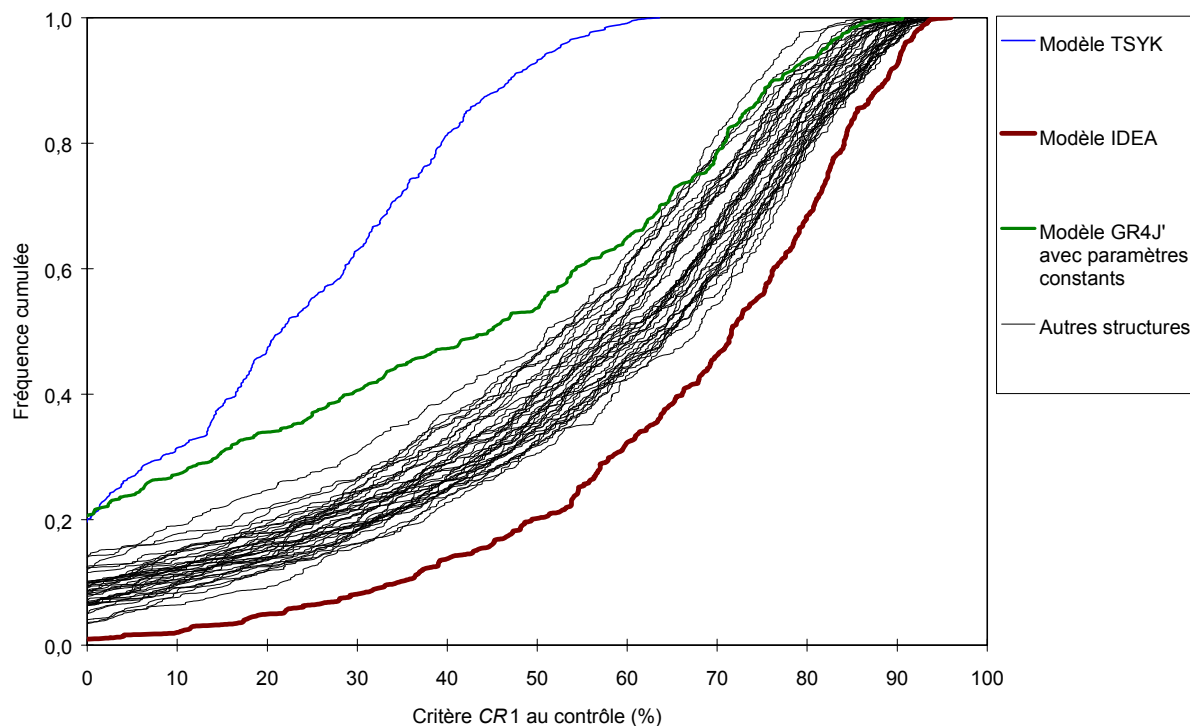


Figure 8.18: Distribution des résultats obtenus au contrôle pour les modèles testés dans la comparaison et le modèle GR4J' avec paramètres constants

8.6. Conclusion

Nous avons recherché dans ce travail à explorer une des voies d'application des modèles globaux, la régionalisation, qui a pour but ultime de pouvoir utiliser ces modèles sur des sites non-jaugés. Classiquement, il s'agit d'établir des liens entre paramètres du modèle et descripteurs du bassin, par des procédures de type régressif. Ici, nous avons proposé une méthode qui paraît plus apte à optimiser les performances du modèle lorsque les valeurs des paramètres sont prédéterminées à partir de ces descripteurs. La nouvelle méthodologie fait intervenir le modèle explicitement au cours de la recherche de la forme des équations de régression. Les coefficients de ces formules de prédétermination sont ajustés non pas pour optimiser les corrélations entre paramètres calés et prédéterminés, comme cela est fait classiquement, mais pour optimiser directement les performances du modèle.

Nous avons appliqué la méthodologie sur un sous-échantillon de 131 bassins sur lesquels la nouvelle version du modèle GR à quatre paramètres obtient de bons résultats. Il est apparu, comme cela avait été montré dans de précédentes études sur ces modèles, que les deux paramètres de rendement du modèle sont très difficiles à prédéterminer d'après de simples descripteurs climatiques du bassin dont nous avons disposé. L'utilisation d'un modèle mensuel avec une structure comparable à celle du modèle journalier pourrait conduire à une estimation plus satisfaisante de ces paramètres, comme l'a montré Makhlof (1994). Par conséquent, les deux paramètres de rendement ont été maintenus à des valeurs constantes dans notre étude.

Avec les descripteurs que nous avons utilisés, il paraît préférable de considérer les valeurs transformées des paramètres (telles qu'elles sont utilisées dans l'algorithme de calage) pour conduire la procédure de régionalisation. Dans ce cas, les gains apportés par la méthode que

nous avons appliquée sont relativement faibles. Ils sont plus nets lorsque l'on part des régressions établies sur paramètres réels. Lorsque l'on essaie d'étendre l'application des formules à l'échantillon total des 429 bassins, les solutions de prédétermination après ajustement n'apparaissent cependant pas meilleures que la solution la plus satisfaisante avant ajustement. L'utilisation de paramètres constants donne dans ce cas des résultats tout aussi satisfaisants.

Le relatif échec dans l'établissement d'équations de régression qui puissent apporter un gain par rapport à l'utilisation de paramètres constants réside probablement dans le fait que l'on est certainement très loin de pouvoir identifier toutes les caractéristiques qui trahiraient le comportement réel d'un bassin vis-à-vis de la transformation pluie-débit, et qui puissent en même temps être pertinente pour le modèle considéré.

Remarquons enfin que le modèle utilisant des paramètres constants a paru capable de donner des résultats relativement robustes en comparaison du modèle 'boîte noire' testé comme modèle de référence dans la comparaison. Ceci indique que même en l'absence d'ajustements des paramètres, la structure du modèle peut fournir de relativement bons résultats.

Le degré de simplicité auquel nous nous sommes situés ici ne nous a pas permis d'établir de relations fiables permettant d'envisager l'application du modèle sur des bassins non-jaugés. Ce n'était d'ailleurs pas l'objectif de ce chapitre. Il semble en outre que la phase d'optimisation des paramètres reste indispensable. Des relations de prédétermination pourraient être utilisées comme premiers estimateurs des paramètres qui seraient ensuite affinés par une procédure d'optimisation sur un nombre très limité de données (données historiques par exemple).

Conclusion générale

Conclusion générale

Notre démarche de recherche s'est inscrite dans la continuation des travaux de modélisation pluie-débit poursuivis depuis une quinzaine d'années au Cemagref d'Antony, dont un des objectifs est de mettre au point des modèles hydrologiques simples, fiables, opérationnels et nécessitant peu de données, les rendant ainsi applicables dans des contextes très variés. Notre approche a été centrée sur l'étude de modèles globaux conceptuels ou empiriques fonctionnant en continu au pas de temps journalier (pas de temps bien adapté à l'étude de nombreux bassins et pour lequel les données sont abondantes). Les discussions présentées au chapitre 1 font ressortir que de tels modèles semblent actuellement les plus aptes à apporter des réponses à une grande partie des questions d'hydrologie opérationnelle dans les domaines de la gestion de la ressource en eau, de la prévision des risques et de l'ingénierie hydrologique. Comme d'autres types de modèles, ils sont cependant encore peu capables de prévoir des effets de changements d'occupation des sols ou de climat. Par ailleurs, leur caractère global, qui leur permet de se soustraire aux problèmes de prise en compte de la forte hétérogénéité spatiale du bassin, ne leur permet pas de représentations des flux ou des chemins de l'eau au sein du bassin. Il implique par ailleurs une détermination numérique (calage) de leurs paramètres qui n'ont pas de signification physique a priori. Nos recherches se sont plus particulièrement intéressées au modèle du Génie Rural (GR), succédant ainsi à celles réalisées autour de Claude Michel par Edijatno (1991), Yang (1993), Makhlof (1994) et Nascimento (1995).

L'objectif général de notre travail était d'évaluer les qualités des modèles existants, de situer parmi eux les modèles GR, et d'explorer les voies d'amélioration possibles de ces derniers. La recherche de modèles plus satisfaisants a en effet des enjeux importants. Elle permet tout d'abord une meilleure compréhension du comportement hydrologique du bassin versant. Elle garantit ensuite une meilleure fiabilité aux applications qui utilisent les simulations du modèle dans un contexte opérationnel. Partant des réflexions de Nascimento (1995), qui soulignent la nécessité de réaliser des études comparatives de modèles, nous avons mis au point un cadre d'évaluation pour tester de nombreux modèles pluie-débit sur un vaste échantillon de bassins.

Mise en place d'un cadre de comparaison

Le foisonnement actuel des modèles conceptuels ou empiriques constitue une richesse que nous avons voulu exploiter. Nous avons regroupé des formulations mathématiques variées correspondant à des concepts et perceptions très diversifiés sur les processus de génération des débits à l'échelle du bassin versant. Ainsi 38 structures dérivées de modèles existants et contenant de trois à neuf paramètres ont été comparées. L'établissement d'une procédure d'évaluation uniforme a permis de tester ces structures dans des conditions rigoureusement identiques, avec les mêmes données d'entrée et la même procédure d'estimation des paramètres, les mettant ainsi sur un pied d'égalité.

Leur capacité à reproduire les débits observés a été mesurée sur des bassins aux caractéristiques très variées, représentant une large gamme de conditions climatiques. Un vaste échantillon de 429 bassins versants a en effet servi de support à notre étude, ceux-ci étant situés majoritairement en France, mais également aux Etats-Unis, en Australie, en Côte-d'Ivoire et au Brésil. Une telle diversification, en conférant à nos résultats une plus grande généralité, leur assure ainsi une plus grande robustesse que si nous nous étions restreints à l'étude de quelques bassins particuliers.

L'utilisation des modèles sur chacun des bassins requiert une estimation mathématique préalable des valeurs de leurs paramètres. Cette phase d'optimisation a été conduite à l'aide d'une méthode de recherche locale, la méthode 'pas-à-pas'. L'évaluation de ses performances réalisée au chapitre 3 indique qu'elle est capable d'identifier un optimum conduisant à des simulations acceptables des modèles, et que son utilisation préférentiellement à d'autres méthodes plus élaborées (procédures multi-départs notamment) n'introduit pas de biais dans notre comparaison au détriment des structures les plus complexes. Il ressort que les éventuels problèmes de stabilité des paramètres rencontrés lors de l'optimisation, qui influencent a posteriori la robustesse des modèles, sont davantage à mettre sur le compte des caractéristiques structurelles de ceux-ci que sur de possibles faiblesses de l'algorithme de calage. En effet, les difficultés d'identification d'un optimum proviennent de l'existence pour la majorité des modèles d'optima essentiellement équivalents au regard de la fonction objectif utilisée.

Toute procédure numérique de détermination des paramètres présuppose le choix d'un critère quantitatif, adopté par l'utilisateur en fonction des objectifs d'application du modèle (prévision de crues, d'étiages, etc.). Le cadre comparatif de notre travail, qui ne se voulait pas centré sur une application particulière, nous a conduits à choisir un critère d'optimisation intermédiaire, ne favorisant particulièrement ni la simulation des crues ni celle des étiages. Par ailleurs, nous avons sélectionné un ensemble de critères quantitatifs permettant d'évaluer plusieurs aspects de la qualité de simulation des modèles lors de la phase de contrôle. Ces mesures de performances peuvent permettre à l'utilisateur de quantifier la crédibilité du modèle vis-à-vis des différentes applications qu'il conduira à partir de celui-ci.

Cette recherche de critères nous a conduits à nous intéresser à certaines mesures de quantification de l'erreur relative du modèle, dont la formulation classique n'est pas entièrement satisfaisante. Le critère de Nash, en particulier, apparaît peu pertinent dans les cas où les débits observés présentent une faible variance. Nous avons pu en proposer une nouvelle formulation faisant intervenir les variables climatiques d'entrée du modèle. Elle réduit largement les cas d'obtention de valeurs fortement négatives obtenues avec le critère de Nash, qui sont difficilement interprétables et parfois peu justifiées au vu des hydrogrammes simulés. La formulation proposée, en étant moins influencée par les caractéristiques propres des bassins, permet de comparer de façon plus pertinente les performances du modèle d'un bassin à l'autre ou d'une période à l'autre. Par ailleurs, une nouvelle formulation du critère de bilan, donnant de façon symétrique une même importance aux surestimations et aux sous-estimations du bilan en eau a également été proposée. Au total, six critères de qualité, construits sur diverses formes de l'erreur du modèle et diverses prises en compte du débit par des transformations mathématiques préalables, ont été utilisés pour évaluer les performances des modèles.

Résultats des tests comparatifs

Notre réflexion s'est fondée sur les résultats des 38 structures obtenus en phase de contrôle. Leur analyse nous a conduits à distinguer deux aspects majeurs de modélisation qui conditionnent la capacité des modèles conceptuels et empiriques globaux à simuler la

transformation pluie-débit, à savoir la complexité et la formulation mathématique de leur structure.

Une première conclusion de ces tests est que ces modèles apparaissent tous grossièrement équivalents lorsqu'ils sont comparés à une approche de modélisation plus rudimentaire ne comportant pas de procédure conceptuelle de suivi d'humidité du bassin.

La complexité du modèle a été analysée sous l'angle du nombre de paramètres optimisés. Les résultats de la comparaison, ainsi que le test de structures modifiées du modèle GR, ont indiqué que, si l'accroissement du nombre de paramètres améliore logiquement en phase de calage la capacité du modèle à obtenir de bons résultats, des structures simples avec seulement trois à cinq paramètres optimisés obtiennent en revanche en phase de contrôle des résultats aussi satisfaisants que des structures plus complexes. Le manque de robustesse des structures les plus complexes provient principalement de leur sur-paramétrisation. Les gains de performance obtenus en calage par ajout de paramètres supplémentaires sont perdus en contrôle par l'effet des incertitudes plus importantes que ces derniers apportent, comme nous l'avons vu au chapitre 6.

Par ailleurs, les résultats des chapitres 5 et 7 indiquent également qu'à niveau de complexité identique, il peut exister d'importants écarts de performances entre des structures différentes. Notre analyse fait donc ressortir le deuxième point important de la construction des modèles, à savoir la formulation mathématique de la structure, c'est-à-dire le choix et l'agencement des réservoirs et des fonctions intervenant dans les modules de production et de routage du modèle. Or, à partir de l'approche comparative, nous avons montré que certaines structures de modèle, utilisant des fonctions différentes, sont complémentaires (en terme de performance) dans leur capacité à simuler les débits. Un modèle 'idéal' hypothétique regroupant toutes les qualités des 38 structures testées obtient d'ailleurs des performances nettement plus satisfaisantes que les meilleures des 38 structures testées. Ces deux éléments suggèrent que des progrès peuvent être faits dans la formulation des structures des modèles, conduisant à des structures plus performantes.

Vers une amélioration du modèle GR

Dans notre étude, les modèles GR sont apparus parmi les plus fiables et les plus robustes. Leur qualité provient à la fois de leur parcimonie et de l'approche empirique adoptée dans leur développement, qui a permis de choisir les différents éléments de leur structure avec un seul souci d'efficacité. L'analyse des complémentarités à partir de la structure la plus simple (celle du modèle GR3J) a mis en évidence des voies possibles de modifications de cette structure. Si l'optimisation de paramètres fixes du modèle est une solution simple de complexification, elle n'apparaît cependant pas être la manière la plus judicieuse d'introduire un degré de liberté supplémentaire, qui semble nécessaire, au sein de la structure. Nous avons adopté une approche empirique de modification du modèle, dans laquelle les changements sont acceptés seulement au regard des améliorations de performance obtenues sur l'échantillon de bassins versants. Une structure modifiée du modèle GR, intégrant une percolation issue du réservoir de production, a paru être parmi les plus satisfaisantes. Cette version à quatre paramètres permet d'améliorer assez sensiblement les performances sur les différents critères de qualité par rapport aux versions antérieures du modèle: elle réalise un progrès substantiel sur la simulation des étiages et une légère amélioration sur la représentation des crues. Les améliorations ainsi obtenues permettent d'avoir une confiance accrue dans le modèle et garantissent de ce fait une meilleure fiabilité aux applications hydrologiques qui l'utilisent.

Une autre voie de recherche de complémentarités entre structures a également été explorée par la mise au point d'une typologie entre structures et types de bassins étudiés. Celle-ci permettrait d'appliquer sur des groupes de bassins présentant des similarités de comportement et de caractéristiques hydrologiques, les modèles qui leur semblent les plus adaptés. Une telle approche ne permet cependant pas d'obtenir de meilleurs résultats que l'approche précédente visant à exploiter les complémentarités au sein d'une structure unique. Ses limitations proviennent, d'après les résultats que nous avons obtenus, de la difficulté de déterminer des groupes homogènes de bassins. Cependant, cette approche distinctive a été menée dans notre étude indépendamment de toute référence à la localisation géographique des bassins. Le travail à une échelle régionale aurait peut-être permis d'établir une typologie plus fiable, à l'image des approches de régionalisation où il est souvent plus facile d'identifier des homogénéités hydrologiques sur une échelle réduite.

Dans la continuité des travaux d'Edijatno (1991) et Makhoul (1994), nous avons ensuite cherché à voir quelle information peut être extraite de descripteurs simples du bassin pour calculer les valeurs des paramètres du modèle, dans la perspective ultime de réaliser des simulations de débit sans calage préalable sur des séries temporelles. Pour cela, on établit, dans une démarche de régionalisation, des équations liant les valeurs des paramètres à des descripteurs du bassin. Nous avons proposé, dans le cadre du chapitre 8, une procédure différente de celle adoptée classiquement, qui permet de maximiser les performances du modèle. Les résultats obtenus indiquent que la procédure permet en effet d'obtenir des résultats du modèle légèrement meilleurs que par l'approche classique utilisée pour établir les relations de prédétermination. Cependant, la vérification sur un échantillon plus vaste de bassins ne confirme pas ce gain et montre que l'utilisation de paramètres constants est aussi satisfaisante. L'échec relatif de l'utilisation de solutions de prédétermination réside en partie dans la difficulté de trouver des variables effectives qui soient à la fois pertinentes pour le modèle et pour décrire le bassin. Les descripteurs simples, facilement accessibles, qui sont communément utilisés, sont souvent maladroits pour à la fois caractériser le bassin et répondre aux besoins du modèle. D'autres, relatifs aux parties sol et sous-sol du bassin par exemple, jouent un rôle essentiel dans la transformation pluie-débit mais restent actuellement plus difficiles d'accès.

Quelques perspectives de recherche

Les progrès accomplis durant cette thèse dans la construction du modèle GR journalier ne doivent cependant pas occulter les points d'interrogation qui subsistent. Dans notre démarche de recherche de modifications de la structure, il est apparu très difficile d'introduire au sein du modèle des degrés de complexité supplémentaires donnant lieu à des gains substantiels de performance. Nous pensons que cela vient avant tout du fait que les modèles hydrologiques développés actuellement restent relativement éloignés du comportement réel du bassin versant et qu'ils ne peuvent de ce fait supporter qu'un faible degré de complexité. Même s'il a montré des performances satisfaisantes en comparaison aux autres structures, le modèle GR ne semble pas échapper à cette observation générale. Ce manque de pertinence du modèle ne nous permet de savoir que de façon grossière le degré de justesse de sa structure. En particulier, il est très difficile de dire si les choix de modifications que nous avons adoptés dans ce travail sont réellement judicieux et s'ils ne sont pas apparus satisfaisants seulement dans la mesure où ils compensent d'autres maladrotes du modèle. La simplicité du modèle, tant dans sa formulation que dans le nombre de paramètres utilisés, reste cependant un garde-fou efficace à la multiplication de ces maladrotes.

Nascimento (1995), en conclusion de ses travaux, soulignait la nécessité de valider de façon plus approfondie la fonction d'échanges en eau qu'il avait introduite dans le modèle pour une

meilleure simulation sur les bassins intermittents. Cette fonction s'était en fait révélée bénéfique à d'autres bassins sans caractère d'intermittence. Nos tests étendus à l'échantillon des 429 bassins versants ont confirmé la nécessité d'une telle fonction dans le modèle. Plusieurs formulations sont apparues équivalentes sans qu'il soit possible de déterminer avec certitude la forme la plus pertinente. Par ailleurs, il est délicat d'expliquer clairement le lien entre cette fonction et le fonctionnement du réservoir de production qui sont tous deux impliqués dans la gestion des bilans en eau dans le modèle. A l'occasion des travaux d'analyse d'incertitudes, nous avons également remarqué qu'il subsiste des ambiguïtés sur la détermination du paramètre d'échange. Ces questions sont certainement à mettre en lien avec la simulation des pertes par évaporation dans le modèle et avec la pertinence de l'évapotranspiration potentielle comme variable de prise en compte de la demande évaporative. Une réflexion globale sur ces différents aspects de la fonction de production devrait être engagée, éventuellement en lien avec une modélisation à un pas de temps plus grand (mensuel par exemple).

Nos travaux d'explication des paramètres, s'ils n'ont pas apporté tous les progrès escomptés, ont cependant permis de mettre en évidence quelques points intéressants. Dans les études de régionalisation, il est particulièrement souhaitable de disposer d'un modèle dont les paramètres sont déterminés avec une faible incertitude. Le modèle GR semble à ce titre bien se prêter à de tels exercices. Nous avons observé une grande difficulté pour relier les paramètres de rendement du modèle aux descripteurs climatiques du bassin, qui est peut-être à mettre en lien avec les remarques précédentes sur la définition interne du modèle. Cette difficulté avait été également soulignée par Makhlouf (1994), qui avait pourtant disposé de descripteurs morphologiques, géologiques ou pédologiques du bassin en plus des variables climatiques. Makhlouf (1994) avait contourné cette difficulté par l'utilisation d'un modèle à plus grand pas de temps dans lequel les fonctions de production ont une importance dominante sur celles de transfert. Cette possibilité renforce l'intérêt de chercher à établir des familles de modèles à différents pas de temps qui aient une forte parenté de structure, permettant ainsi d'établir des liens entre les paramètres des différentes structures. Des travaux sont actuellement menés dans ce sens par Mouelhi (2000).

Il semble cependant que l'utilisation de telles relations de prédétermination des paramètres ne permettent pas à l'heure actuelle de garantir la fiabilité requise pour l'utilisation du modèle dans des applications hydrologiques opérationnelles sur des bassins non jaugés. Cette étape de prédétermination des paramètres apparaît donc comme une manière d'avoir un premier estimateur grossier de leurs valeurs. Une procédure de calage reste nécessaire. Il serait donc intéressant de déterminer quel niveau d'information (en terme de données hydrologiques) sur le bassin serait nécessaire pour affiner ces premières estimations grossières. L'enregistrement de données sur une courte période, ainsi que l'utilisation de quelques valeurs historiques pourraient être suffisantes pour augmenter de façon importante la fiabilité des paramètres du modèle initialement prédéterminés par de simples équations régionales et ouvrir ainsi la possibilité d'utiliser le modèle dans ces conditions.

Perspective d'application

A côté de ces perspectives de recherches, ce travail doit permettre d'engager avec plus de confiance le développement d'outils hydrologiques s'appuyant sur le modèle et destinés à être utilisés dans le domaine opérationnel. Il existe en effet une forte demande dans ce sens en France et à l'étranger de la part des ingénieurs, des concepteurs d'ouvrages, des gestionnaires de la ressource et des décideurs, dans un contexte où l'eau devient l'objet de plus en plus d'enjeux. Proposer des applications du modèle pour l'ingénierie permettrait de faire des progrès en comparaison des méthodes utilisées actuellement, encore souvent trop

rudimentaires. Si notre approche de modélisation ne peut aujourd'hui pas répondre à tous les problèmes d'hydrologie quantitative, elle se situe, selon nous, parmi les plus fiables et les plus facilement utilisables pour un large éventail de questions posées.

Bibliographie

Bibliographie

- [1] **Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. et Rasmussen, J. (1986).** An introduction to the European Hydrological System - Système Hydrologique Européen 'SHE'. *Journal of Hydrology*, **87**, 45-77.
- [2] **Abdulla, F.A. et Lettenmaier, D.P. (1997).** Development of regional parameter estimation equations for a macroscale hydrologic model. *Journal of Hydrology*, **197**, 230-257.
- [3] **Abdulla, F.A., Lettenmaier, D.P. et Liang, X. (1999).** Estimation of the ARNO model baseflow parameters using daily streamflow data. *Journal of Hydrology*, **222**, 37-54.
- [4] **Aitken, A.P. (1973).** Assessing systematic errors in rainfall-runoff models. *Journal of Hydrology*, **20**, 121-136.
- [5] **Allred, B. et Haan, C.T. (1991).** Variability of optimized parameter estimates based on observed record length. *Trans. of Amer. Soc. Agr. Eng.*, **34**(6), 2421-2426.
- [6] **Ambroise, B. (1998).** La dynamique du cycle de l'eau dans un bassin versant - Processus, Facteurs, Modèles. *H*G*A* (Ed.), Bucarest, 200 p.
- [7] **Ambroise, B., Perrin, J.L. et Reutenauer, D. (1995).** Multicriterion validation of a semi-distributed conceptual model of the water cycle in the Fecht catchment (Vosges Massif, France). *Water Resources Research*, **31**(6), 1467-1481.
- [8] **Andréassian, V. (2001).** Impact de l'évolution du couvert forestier sur le comportement hydrologique des bassins versants. *Thèse de Doctorat, Université Pierre et Marie Curie Paris VI, Cemagref, Thèse en cours.*
- [9] **Arikan, A. (1988).** MODALP: a deterministic rainfall-runoff model for large karstic areas. *Hydrological Sciences Journal*, **33**(4), 401-414.
- [10] **Arnell, N., Stanescu, V.A., Krasovskaia, I., Servat, E., Mahé, G., Galéa, G., Leblois, E., Sauquet, E. et Van der Wateren de Hoog, B. (1997).** Chapter 2: Regimes and regional hydrology. In: *FRIEND projects H-5-5 and I.1. Third report: 1994-1997, Cemagref (Ed.)*, 63-128.
- [11] **Baudez, J.C. (1997).** Déterminants hydrologiques régionaux pour la gestion et la prévision des ressources en eau. *Mémoire d'ingénieur, ENGEES/Cemagref*, 88 p.
- [12] **Bergström, S. (1991).** Principles and confidence in hydrological modelling. *Nordic Hydrology*, **22**, 123-136.
- [13] **Bergström, S. (1995).** The HBV model. In: *Computer Models of Watershed Hydrology, Chapter 13, V. P. Singh (Ed.)*, Water Resources Publications, 443-476.
- [14] **Bergström, S. et Forsman, A. (1973).** Development of a conceptual deterministic rainfall-runoff model. *Nordic Hydrology*, **4**, 147-170.
- [15] **Beven, K. (1986).** Hillslope runoff processes and flood frequency characteristics. In: *Hillslope Processes, A. D. Abrahams (Ed.)*, 187-202.

- [16] **Beven, K. (1987).** Towards a new paradigm in hydrology. *Water for the future : Hydrology in Perspective, Proceedings of the Rome Symposium, April 1987, IAHS Publication n° 164*, 393-403.
- [17] **Beven, K. (1989).** Changing ideas in hydrology - The case of physically-based models. *Journal of Hydrology*, **105**, 157-172.
- [18] **Beven, K. et Binley, A. (1992).** The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*, **6**(3), 279-298.
- [19] **Beven, K.J. (1993).** Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, **16**, 41-51.
- [20] **Beven, K.J., Calver, A. et Morris, E.M. (1987).** The Institute of Hydrology Distributed Model. *Report n° 8, Institute of Hydrology, Wallingford (UK)*.
- [21] **Beven, K.J. et Kirkby, M.J. (1979).** A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, **24**(1), 43-69.
- [22] **Blackie, J.R. et Eeles, C.W.C. (1985).** Lumped catchment models - Chapter 11. In: *Hydrological forecasting, M. G. Anderson et T. P. Burt (Ed.), John Wiley & Sons Ltd*, 311-345.
- [23] **Bobba, A.G. et Lam, D.C.L. (1985).** Application of linearly distributed surface run-off model for watershed application problems. *Proc. Can. Hydrol. Symp.*, **15**, 381-398.
- [24] **Bonvoisin, N.J. et Boorman, D.B. (1992).** Daily rainfall-runoff modelling as an aid to the transfer of hydrological parameters. *Report to MAFF, Institute of Hydrology, Wallingford, UK*.
- [25] **Box, G.E.P. et Cox, D.R. (1964).** An analysis of transformations. *J. Royal Stat. Soc.*, **26**, 211-252.
- [26] **Box, G.E.P. et Jenkins, G.M. (1976).** Time series analysis - Forecasting and control. *Holden Day (Ed.), Oakland, California*, 575 p.
- [27] **Braun, L.N. et Renner, C.B. (1992).** Application of a conceptual runoff model in different physiographic regions of Switzerland. *Hydrological Sciences Journal*, **37**(3), 217-231.
- [28] **Brazil, L.E. et Hudlow, M.D. (1980).** Calibration procedures used with the National Weather Service River Forecast System. *Proceedings of IFAC Conference on Water and Related Land Resources Systems, Cleveland, USA*, 457-466.
- [29] **Brazil, L.E. et Krajewski, W.F. (1987).** Optimization of complex hydrologic models using random search methods. *Engineering Hydrology Proceedings, Williamsburg, Virginia, USA, August 3-7, Hydraulics Division, ASCE*, 726-731.
- [30] **Bronstert, A., Glüsing, B. et Plate, E. (1998).** Physically based hydrological modelling on the hillslope and micro-catchment scale: examples of capabilities and limitations. *IAHS Publication n° 248*, 207-215.
- [31] **Bultot, F. et Dupriez, G.L. (1976).** Conceptual hydrological model for an average-sized catchment area. I. Concepts and relationships. *Journal of Hydrology*, **29**, 251-272.
- [32] **Burn, D.H. (1997).** Catchment similarity for regional flood frequency analysis using seasonality measures. *Journal of Hydrology*, **202**, 212-230.
- [33] **Burn, D.H. et Boorman, D.B. (1993).** Estimation of hydrological parameters at ungauged catchments. *Journal of Hydrology*, **143**, 429-454.
- [34] **Burn, D.H. et Goel, N.K. (2000).** The formation of groups for regional flood frequency analysis. *Hydrological Sciences Journal*, **45**(1), 97-112.
- [35] **Burnash, R.J.C., Ferral, R.L. et McGuire, R.A. (1973).** A generalized streamflow simulation system - Conceptual modelling for digital computers, *U.S. Department of Commerce, National Weather Service and State of California, Department of Water Resources*.
- [36] **Cameron, D.S., Beven, K.J., Tawn, J., Blazkova, S. et Naden, P. (1999).** Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty). *Journal of Hydrology*, **219**, 169-187.

- [37] **Cans, R. (1994).** La Bataille de l'Eau. *Le Monde (Ed.), Paris*, 219 p.
- [38] **Chapman, T.G. (1996a).** Common unitgraphs for sets of runoff events. Part 1: unitgraph identification from streamflow data. *Hydrological Processes*, **10**(6), 773-782.
- [39] **Chapman, T.G. (1996b).** Common unitgraphs for sets of runoff events. Part 2: comparisons and inferences for rainfall loss models. *Hydrological Processes*, **10**(6), 783-792.
- [40] **Chiew, F. et McMahon, T. (1994).** Application of the daily rainfall-runoff model MODHYDROLOG to 28 Australian catchments. *Journal of Hydrology*, **153**, 383-416.
- [41] **Chiew, F.H.S. et McMahon, T.A. (1991).** The applicability of Morton's and Penman's evapotranspiration estimates in rainfall-runoff modeling. *Water Resources Bulletin*, **27**(4), 611-620.
- [42] **Chiew, F.H.S. et McMahon, T.A. (1993).** Detection of trend or change in annual flow of Australian rivers. *International Journal of Climatology*, **13**, 643-653.
- [43] **Chiew, F.H.S., Pitman, A.J. et McMahon, T.A. (1996).** Conceptual catchment scale rainfall-runoff models and AGCM land-surface parameterisation schemes. *Journal of Hydrology*, **179**, 137-157.
- [44] **Chiew, F.H.S., Stewardson, M.J. et McMahon, T.A. (1993).** Comparison of six rainfall-runoff modelling approaches. *Journal of Hydrology*, **147**, 1-36.
- [45] **Chiew, F.H.S., Whetton, P.H., McMahon, T.H. et Pittock, A.B. (1995).** Simulation of the impacts of climate change on runoff and soil moisture in Australian catchments. *Journal of Hydrology*, **167**, 121-147.
- [46] **Clarke, R.T. (1973).** A review of some mathematical models used in hydrology, with observations on their calibration and use. *Journal of Hydrology*, **19**, 1-20.
- [47] **Cooper, V.A., Nguyen, V.T.V. et Nicell, J.A. (1997).** Evaluation of global optimization methods for conceptual rainfall-runoff model calibration. *Water Sci. Tech.*, **36**(5), 53-60.
- [48] **Cormary, Y. et Guilbot, A. (1973).** Etude des relations pluie-débit sur trois bassins versants d'investigation. *IAHS Madrid Symposium, IAHS Publication n° 108*, 265-279.
- [49] **Coulibaly, P., Anctil, F. et Bobbée, B. (2000).** Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *Journal of Hydrology*, **230**, 244-257.
- [50] **Crawford, N.H. et Linsley, R.K. (1963).** A conceptual model of the hydrologic cycle. *IAHS Publication n° 63*, 573-587.
- [51] **CTGREF (1980).** Estimation du débit de crue décennal sur un bassin versant non jaugeé - La méthode CRUPEDIX, *CTGREF-Informations techniques, Cahier 40, n° 3*.
- [52] **Da Ros, D. et Borga, M. (1997).** Adaptive use of a conceptual model for real time forecasting. *Nordic Hydrology*, **28**(3), 169-188.
- [53] **Dawdy, D.R. et O'Donnell, T. (1965).** Mathematical models of catchment behavior. *American Society of Civil Engineers Proceedings*, **91**(HY4), 123-137.
- [54] **Dawson, C.W. et Wilby, R. (1998).** An artificial neural network approach to rainfall-runoff modelling. *Hydrological Sciences Journal*, **43**(1), 47-66.
- [55] **Dezetter, A. (1991).** Modélisation globale de la relation pluie débit. Application en zone de savanes soudanaises (Nord-Ouest de la Côte d'Ivoire). *Thèse de Doctorat, Université de Montpellier II*, 422 p.
- [56] **Diskin, M.H. et Simon, E. (1977).** A procedure for the selection of objective functions for hydrologic simulation models. *Journal of Hydrology*, **34**, 129-149.
- [57] **Donnelly-Makowecki, L.M. et Moore, R.D. (1999).** Hierarchical testing of three rainfall-runoff models in small forested catchments. *Journal of Hydrology*, **219**(3/4), 136-152.
- [58] **Dooge, J.C.I. (1986).** Looking for hydrological laws. *Water Resources Research*, **22**(9), 46S-58S.
- [59] **Dooge, J.C.I. (1988).** Hydrology past and present. *Journal of Hydraulic Research*, **26**(1), 5-26.

- [60] **Duan, Q., Gupta, V.K. et Sorooshian, S. (1993).** Shuffled Complex Evolution approach for effective and efficient global minimization. *Journal of Optimization Theory and Applications*, **76**(3), 163-168.
- [61] **Duan, Q., Sorooshian, S. et Gupta, V.K. (1992).** Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, **28**(4), 1015-1031.
- [62] **Duband, D., Obled, C. et Rodriguez, J.Y. (1993).** Unit hydrograph revisited: an alternate iterative approach to UH and effective precipitation identification. *Journal of Hydrology*, **150**, 115-149.
- [63] **Eagleson, P.S. (1991).** Preface. In: *Opportunities in the Hydrologic Sciences*, N. R. Council (Ed.), National Academy Press, Washington, 348 p.
- [64] **Edijatno (1991).** Mise au point d'un modèle élémentaire pluie-débit au pas de temps journalier. *Thèse de Doctorat, Université Louis Pasteur/ENGEES, Strasbourg*, 242 p.
- [65] **Edijatno, Nascimento, N.O., Yang, X., Makhlof, Z. et Michel, C. (1999).** GR3J: a daily watershed model with three free parameters. *Hydrological Sciences Journal*, **44**(2), 263-277.
- [66] **Egbuniwe, N. et Todd, D.K. (1976).** Application of the Stanford Watershed Model to Nigerian watersheds. *Water Resources Bulletin*, **12**(3), 449-460.
- [67] **Farnsworth, R.K., Thompson, E.S. et Peck, E.L. (1982).** Evaporation atlas for the contiguous 48 United States, *National Oceanic and Atmospheric Administration, National Weather Service, NOAA technical report NWS n° 33, Washington, D.C.*
- [68] **Faucher, V., Ferry, M. et Michel, C. (1996).** Traitement de base des données hydrologiques. Note de présentation des logiciels "PV", "PF" et "TRANS". *Note interne, Cemagref, Division Hydrologie*.
- [69] **Fletcher, R. et Powell, M.J.D. (1963).** A rapidly convergent descent method for minimization. *The Computer Journal*, **6**, 163-168.
- [70] **Fortin, J.P., Charbonneau, R., Lefevre, J. et Girard, G. (1971).** Proposition et analyse de quelques critères adimensionnels d'optimisation. *IAHS Publication n° 101*, 548-557.
- [71] **Franchini, M. (1996).** Use of a genetic algorithm combined with a local search method for the automatic calibration of conceptual rainfall-runoff models. *Hydrological Sciences Journal*, **41**(1), 21-39.
- [72] **Franchini, M. et Galeati, G. (1997).** Comparing several genetic algorithm schemes for the calibration of conceptual rainfall-runoff models. *Hydrological Sciences Journal*, **42**(3), 357-379.
- [73] **Franchini, M., Galeati, G. et Berra, S. (1998).** Global optimization techniques for the calibration of conceptual rainfall-runoff models. *Hydrological Sciences Journal*, **43**(3), 443-458.
- [74] **Franchini, M. et Pacciani, M. (1991).** Comparative analysis of several conceptual rainfall-runoff models. *Journal of Hydrology*, **122**, 161-219.
- [75] **Franchini, M., Wendling, J., Obled, C. et Todini, E. (1996).** Physical interpretation and sensitivity analysis of the TOPMODEL. *Journal of Hydrology*, **175**, 293-338.
- [76] **Gan, T.Y. et Biftu, G.F. (1996).** Automatic calibration of conceptual rainfall-runoff models: optimization algorithms, catchment conditions, and model structure. *Water Resources Research*, **32**(12), 3513-3524.
- [77] **Gan, T.Y., Dlamini, E.M. et Biftu, G.F. (1997).** Effects of model complexity and structure, data quality and objective function on hydrologic modeling. *Journal of Hydrology*, **192**, 81-103.
- [78] **Garrick, M., Cunnane, C. et Nash, J.E. (1978).** A criterion of efficiency for rainfall-runoff models. *Journal of Hydrology*, **38**, 375-381.
- [79] **Georgakakos, K.P. et Baumer, O.W. (1996).** Measurement and utilization of on-site soil moisture data. *Journal of Hydrology*, **184**, 131-152.

- [80] **Girard, G. (1970).** Essai pour un modèle hydropluviométrique conceptuel et son utilisation au Québec. *Cahiers ORSTOM, Série Hydrologique*, **VII**(2).
- [81] **Girard, G., Ledoux, E. et Villeneuve, J.P. (1981).** Le modèle couplé - Simulation conjointe des écoulements de surface et des écoulements souterrains sur un système hydrologique. *Cah. ORSTOM, Série Hydrologique*, **18**(4), 195-280.
- [82] **Girard, G., Morin, G. et Charbonneau, R. (1972).** Modèle précipitations-débits à discrétisation spatiale. *Cahiers ORSTOM, Série Hydrologie*, **IX**(4), 35-52.
- [83] **Grayson, R.B., Moore, I.D. et McMahon, T.A. (1992).** Physically-based modelling. 2. Is the concept realistic. *Water Resources Research*, **26**(10), 2659-2666.
- [84] **Güntner, A., Uhlenbrook, S., Seibert, J. et Leibundgut, C. (1999).** Multi-criterial validation of TOPMODEL in a mountainous catchment. *Hydrological Processes*, **13**(11), 1603-1620.
- [85] **Gupta, H.V., Sorooshian, S. et Yapo, P.O. (1998).** Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. *Water Resources Research*, **34**(4), 751-763.
- [86] **Gupta, V.K. et Sorooshian, S. (1983).** Uniqueness and observability of conceptual rainfall-runoff model parameters: the percolation process examined. *Water Resources Research*, **19**(1), 269-276.
- [87] **Gupta, V.K. et Sorooshian, S. (1985).** The relationship between data and the precision of parameter estimates of hydrologic models. *Journal of Hydrology*, **81**, 57-77.
- [88] **Gustard, A., Bullock, A. et Dixon, J.M. (1992).** Low flow estimation in the United Kingdom, *Institute of Hydrology, Wallingford, UK, Report n° 108*.
- [89] **Haan, C.T. (1972).** A water yield model for small watersheds. *Water Resources Research*, **8**(1), 58-69.
- [90] **Hargreaves, G.H. et Samani, Z.A. (1982).** Estimating potential evapotranspiration. *Tech. Note J. Irrig. Drain. Eng.*, **108**(3), 225-230.
- [91] **Hendrickson, J.D., Sorooshian, S. et Brazil, L.E. (1988).** Comparison of Newton-type and direct search algorithms for calibration of conceptual rainfall-runoff models. *Water Resources Research*, **24**(5), 691-700.
- [92] **Hérouin, E. (1998).** Capacité de transport de sédiments dans les cours d'eau en régime non-uniforme. *Thèse de Doctorat, Université Claude Bernard Lyon1 Villeurbanne, Cemagref (Lyon)*, 168 p.
- [93] **Holland, J.H. (1975).** Adaptation in natural and artificial systems. *University of Michigan Press (Ed.), Ann Arbor, Michigan*.
- [94] **Hooke, R. et Jeeves, T.A. (1961).** Direct search solutions of numerical and statistical problems. *J. Ass. Comput. Mach.*, **8**(2), 212-229.
- [95] **Houghton-Carr, H.A. (1999).** Assessment criteria for simple conceptual daily rainfall-runoff models. *Hydrological Sciences Journal*, **44**(2), 237-261.
- [96] **Hsu, K.L., Gupta, H.V. et Sorooshian, S. (1995).** Artificial neural network modeling of the rainfall-runoff processes. *Water Resources Research*, **31**(10), 2517-2530.
- [97] **Hughes, D.A. et Metzler, W. (1998).** Assessment of three monthly rainfall-runoff models for estimating the water resource yield of semiarid catchments in Namibia. *Hydrological Sciences Journal*, **43**(2), 283-298.
- [98] **Ibbitt, R.P. et Hutchinson, P.D. (1984).** Model parameter consistency and fitting criteria. *Proceedings of the 9th IFAC Triennial World Congress, Budapest, Hungary, IFAC Pub.*, 3169-3173.
- [99] **Ibbitt, R.P. et O'Donnell, T. (1971a).** Designing conceptual catchment models for automatic fitting methods. *LAHS Publication n° 101*, 462-475.
- [100] **Ibbitt, R.P. et O'Donnell, T. (1971b).** Fitting methods for conceptual catchment models. *Journal of the Hydraulics Division, ASCE*, **97**(HY9), 1131-1142.

- [101] **Isabel, D. et Villeneuve, J.P. (1986).** Importance of convergence criterion in the automatic calibration of hydrological models. *Water Resources Research*, **16**(6), 1025-1033.
- [102] **Jakeman, A.J. et Hornberger, G.M. (1993).** How much complexity is warranted in a rainfall-runoff model ? *Water Resources Research*, **29**(8), 2637-2649.
- [103] **Jakeman, A.J., Littlewood, I.G. et Whitehead, P.G. (1990).** Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. *Journal of Hydrology*, **117**, 275-300.
- [104] **Jakeman, A.J. et Whitehead, P. (1996).** Unit hydrograph methods and applications. *Hydrological Processes*, **10**(6), 771-893.
- [105] **James, L.D. (1972).** Hydrologic modeling, parameter estimation, and watershed characteristics. *Journal of Hydrology*, **17**, 283-307.
- [106] **Jayawardena, A.W. et Zhou, M.C. (2000).** A modified spatial soil moisture storage capacity distribution curve for the Xinanjiang model. *Journal of Hydrology*, **227**, 93-113.
- [107] **Johansson, B. (1994).** The relationship between catchment characteristics and the parameters of a conceptual runoff model: a study in the south of Sweden. *IAHS Publication n° 221*, 475-482.
- [108] **Johnston, P.R. et Pilgrim, D.H. (1976).** Parameter optimization for watershed models. *Water Resources Research*, **12**(3), 477-486.
- [109] **Juncker, P. (1971).** Toward a global hydrological typology. *Journal of Hydrology (N.Z.)*, **10**(2), 145-153.
- [110] **Kabouya, M. (1990).** Modélisation pluie-débit aux pas de temps mensuel et annuel en Algérie septentrionale. *Thèse de Doctorat, Université Paris Sud Orsay*, 347 p.
- [111] **Kabouya, M. et Michel, C. (1991).** Estimation des ressources en eau superficielle aux pas de temps mensuel et annuel, application à un pays semi-aride. *Revue des Sciences de l'Eau*, **4**(4), 569-587.
- [112] **Karnopp, D.C. (1963).** Random search technique for optimization problems. *Automatica*, **1**, 111-121.
- [113] **Kauark-Leite, L.A. et Nascimento, N.O. (1993).** Développement, utilisation et incertitudes des modèles conceptuels en hydrologie. *Séminaire Eaux - Environnement n° 3, Modélisation du comportement des polluants dans les hydrosystèmes, Paris*, 191-219.
- [114] **Kirkpatrick, S., Gelatt, C.D. et Vecchi, M.P. (1983).** Optimization by simulated annealing. *Science*, **220**(4598), 671-680.
- [115] **Kitanidis, P.K. et Bras, R.L. (1980).** Real-time forecasting with a conceptual hydrologic model. 1. Analysis of uncertainty. *Water Resources Research*, **16**(6), 1025-1033.
- [116] **Kite, G.W. (1975).** Performances of two deterministic hydrological models. *IAHS Publication n° 115*, 136-142.
- [117] **Kite, G.W. (1978).** Development of a hydrologic model for a Canadian watershed. *Can. J. Civ. Eng.*, **5**, 126-134.
- [118] **Klemeš, V. (1982).** Empirical and Causal Models in Hydrology. In: *Studies in Geophysics: Scientific Basis of Water-Resource Management, National Academy Press, Washington*, 95-104.
- [119] **Klemeš, V. (1986a).** Dilettantism in hydrology: transition or destiny ? *Water Resources Research*, **22**(9), 177S-188S.
- [120] **Klemeš, V. (1986b).** Operational testing of hydrological simulation models. *Hydrological Sciences Journal*, **31**(1), 13-24.
- [121] **Klemeš, V. (1988).** A hydrological perspective. *Journal of Hydrology*, **100**, 3-28.
- [122] **Knisel, W.G. et Williams, J.R. (1995).** Hydrology components of CREAMS and GLEAMS models. In: *Computer Models of Watershed Hydrology, Chapter 28, V. P. Singh (Ed.), Water Resources Publications*, 1069-1111.

- [123] **Kuchment, L.S., Demidov, V.N., Naden, P.S., Cooper, D.M. et Broadhurst, P. (1996).** Rainfall-runoff modelling of the Ouse basin, North Yorkshire: an application of a physically based distributed model. *Journal of Hydrology*, **191**, 323-342.
- [124] **Kuczera, G. (1983).** Improved parameter inference in catchment models: 1. Evaluating parameter uncertainty. *Water Resources Research*, **19**(5), 1151-1162.
- [125] **Kuczera, G. et Mroczkowski, M. (1998).** Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. *Water Resources Research*, **34**(6), 1481-1489.
- [126] **Kuczera, G. et Parent, E. (1998).** Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *Journal of Hydrology*, **211**, 69-85.
- [127] **Lamb, R., Beven, K. et Myrabø, S. (1997).** Discharge and water table predictions using a generalized TOPMODEL formulation. *Hydrological Processes*, **11**(9), 1145-1167.
- [128] **Lamb, R., Beven, K. et Myrabø, S. (1998).** Use of spatially distributed water table observations to constrain uncertainty in a rainfall-runoff model. *Advances in Water Resources*, **22**(4), 305-317.
- [129] **Langford, K.J. et O'Shaughnessy, P.J. (1977).** A study of canopy interception in native forests and conifer plantations. *Report n° MMBW-W-0007, Melbourne and Metropolitan Board of Works.*
- [130] **Lavabre, J., Sempere Torres, D. et Cernesson, F. (1993).** Changes in the hydrological response of a small Mediterranean basin a year after a wildfire. *Journal of Hydrology*, **142**, 273-299.
- [131] **Legates, D.R. et McCabe, G.J. (1999).** Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, **35**(1), 233-241.
- [132] **Lek, S., Dimopoulos, I., Derraz, M. et El Ghachtoul, Y. (1996).** Modélisation de la relation pluie-débit à l'aide des réseaux de neurones artificiels. *Revue des Sciences de l'Eau*, **9**(3), 319-331.
- [133] **Levenberg, K. (1944).** A method for the solution of certain nonlinear problems in least squares. *Q. Appl. Math.*, **2**, 164-168.
- [134] **Leviandier, T. (1993).** Scale considerations in propagation of flows through a network of reservoirs. *EGS*, 22 p.
- [135] **Lindström, G., Johansson, B., Persson, M., Gardelin, M. et Bergström, S. (1997).** Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology*, **201**, 272-288.
- [136] **Linsley, R.K. (1982).** Rainfall-runoff models - An overview. In: *Rainfall-runoff relationship, V. P. Singh (Ed.)*, *Water Ressources Publications*, 3-22.
- [137] **Littlewood, I.G., Down, K., Parker, J.R. et Post, D.A. (1997).** The PC version of IHACRES for catchment-scale rainfall-streamflow modelling. Version 1.0. User Guide. *Institute of Hydrology (Ed.)*, 89 p.
- [138] **Loague, K.M. et Freeze, R.A. (1985).** A comparison of rainfall-runoff modeling techniques on small upland catchments. *Water Resources Research*, **21**(2), 229-248.
- [139] **Lørup, J.K., Refsgaard, J.C. et Mazvimavi, D. (1998).** Assessing the effects of land use change on catchment runoff by combined use of statistical tests and hydrological modelling: Case studies from Zimbabwe. *Journal of Hydrology*, **205**, 147-163.
- [140] **Loumagne, C., Chkir, N., Normand, M., Ottlé, C. et Vidal-Madjar, D. (1996).** Introduction of soil/vegetation/atmosphere continuum in a conceptual rainfall-runoff model. *Hydrological Sciences Journal*, **41**(6), 889-902.
- [141] **Loumagne, C., Michel, C., Palagos, B., Baudez, J.C. et Bartoli, F. (1999).** D'une approche globale vers une approche semi-distribuée en modélisation pluie-débit (From a global to a semi-distributed approach in rainfall-runoff modelling). *La Houille Blanche*, **6**, 81-88.

- [142] **Ma, Z.C., Leviandier, T. et Ferry, M. (1990).** Fitting a conceptual hydrologic model accounting for nitrate loss in an agricultural representative basin. *Proceedings of the Conference on Hydrological research basins and the environment, Wageningen*, 319-327.
- [143] **Magette, W.L., Shanholtz, V.O. et Carr, J.C. (1976).** Estimating selected parameters for the Kentucky watershed model from watershed characteristics. *Water Resources Research*, **12**(3), 472-476.
- [144] **Maidment, D.R. (1992).** Hydrology. In: *Applied Hydrology*, D. R. Maidment (Ed.), McGraw Hill, New-York, 1412 p.
- [145] **Makhlouf, Z. (1994).** Compléments sur le modèle pluie-débit GR4J et essai d'estimation de ses paramètres. *Thèse de Doctorat, Université Paris XI Orsay*, 426 p.
- [146] **Manley, R.E. (1975).** A hydrological model with physically realistic parameters. *IAHS Publication n° 115*, 154-161.
- [147] **Marquardt, D.W. (1963).** An algorithm for the least squares estimation of nonlinear parameters. *SIAM J.*, **11**, 431-441.
- [148] **Martinec, J. et Rango, A. (1989).** Merits of statistical criteria for the performance of hydrological models. *Water Resources Bulletin*, **25**(2), 421-432.
- [149] **Masri, S.F., Bekey, G.A. et Safford, F.B. (1976).** An adaptative random search method for identification of large scale nonlinear systems. *Proceedings of the IFAC Symposium on Identification and System Parameter Estimation, Rajbman (Ed.)*, 1645-1654.
- [150] **Mazenc, B., Sanchez, M. et Thiery, D. (1984).** Analyse de l'influence de la physiographie d'un bassin versant sur les paramètres d'un modèle hydrologique global et sur les débits caractéristiques à l'exutoire. *Journal of Hydrology*, **69**, 97-118.
- [151] **McCuen, R.H. (1973).** The role of sensitivity analysis in hydrologic modeling. *Journal of Hydrology*, **18**, 37-53.
- [152] **Mein, R.G. et Brown, B.M. (1978).** Sensitivity of optimized parameters in watershed models. *Water Resources Research*, **14**(2), 299-303.
- [153] **Melching, C.S. (1992).** An improved first-order reliability approach for assessing uncertainties in hydrologic modeling. *Journal of Hydrology*, **132**, 157-177.
- [154] **Melching, C.S. (1995).** Reliability estimation. In: *Computer Models of Watershed Hydrology, Chapter 3*, V. P. Singh (Ed.), *Water Resources Publications*, 69-118.
- [155] **Melo, M.D. et Nascimento, N.O. (1999).** A comparative study on modelling rainfall-runoff processes in large watersheds, using conceptual models of different complexity levels. *Article soumis pour publication*.
- [156] **Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. et Teller, E. (1953).** Equations of state calculations by fast computing machines. *J. Chem. Phys.*, **21**, 1087-1091.
- [157] **Michaud, J. et Sorooshian, S. (1994).** Comparison of simple versus complex distributed runoff models on a midsized semiarid watershed. *Water Resources Research*, **30**(3), 593-605.
- [158] **Michel, C. (1983).** Que peut-on faire en hydrologie avec un modèle conceptuel à un seul paramètre ? *La Houille Blanche*, **1**, 39-44.
- [159] **Michel, C. (1989).** Hydrologie appliquée aux petits bassins versants ruraux, *Cemagref, Antony*.
- [160] **Micovic, Z. et Quick, M.C. (1999).** A rainfall and snowmelt runoff modelling approach to flow estimation at ungauged sites in British Columbia. *Journal of Hydrology*, **226**, 101-120.
- [161] **Moore, I.D. et Mein, R.G. (1975).** An evaluation of three rainfall-runoff models. *Hydrological Symposium, Sydney*, **75/3**, 122-126.
- [162] **Moore, R.J. et Clarke, R.T. (1981).** A distribution function approach to rainfall-runoff modeling. *Water Resources Research*, **17**(5), 1367-1382.
- [163] **Morel-Seytoux, H.J. et Al Hassoun, S. (1989).** The unsaturated component of SWATC: a multiprocess watershed model for runoff generation and routing. In: *Unsaturated flow in hydrologic modeling - Theory and practice*, H. J. Morel-Seytoux (Ed.), *Kluwer Academic Publ.*, 413-433.

- [164] **Morel-Seytoux, H.J. (1999)**. Note descriptive du modèle HMS. *Rapport interne à l'IIBRBS*.
- [165] **Morton, F.I. (1983)**. Operational estimates of actual evapotranspiration and their significance to the science and practice of hydrology. *Journal of Hydrology*, **66**, 1-76.
- [166] **Mouelhi, S. (2000)**. Modélisation pluie-débit aux pas de temps mensuel, annuel et interannuel. *Rapport interne d'avancement, Cemagref*, 58 p.
- [167] **Moussavi, M. et Feyen, J. (1990)**. Comparison of different structures for a monthly water yield model in seasonally snow-covered mountainous watersheds of Iran. *Hydrological Sciences Journal*, **35**(5), 535-546.
- [168] **Mroczkowski, M., Raper, G.P. et Kuczera, G. (1997)**. The quest for more powerful validation of conceptual catchment models. *Water Resources Research*, **33**(10), 2325-2335.
- [169] **Nandakumar, N. et Mein, R.G. (1997)**. Uncertainty in rainfall-runoff model simulations and the implications for predicting the hydrologic effects of land-use change. *Journal of Hydrology*, **192**, 211-232.
- [170] **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, CERGRENE/ENPC, Paris*, 550 p.
- [171] **Nascimento, N.O. et Michel, C. (1992)**. Some epistemological aspects of the development and use of hydrologic conceptual models. *Proceedings of the 4th Junior Scientist Course "Assessment of Modelling Uncertainties and Measurement Error in Hydrology"*, St-Etienne (France), 245-264.
- [172] **Nash, J.E. et Sutcliffe, J.V. (1970)**. River flow forecasting through conceptual models. Part I - A discussion of principles. *Journal of Hydrology*, **27**(3), 282-290.
- [173] **Nathan, R.J. et McMahon, T.A. (1990)**. Identification of homogeneous regions for the purposes of regionalisation. *Journal of Hydrology*, **121**, 217-238.
- [174] **Nathan, R.J. et McMahon, T.A. (1992)**. Estimating low flow characteristics in ungauged catchments. *Water Resources Management*, **6**, 85-100.
- [175] **Ndiritu, J.G. et Daniell, T.M. (1999)**. Assessing model calibration adequacy via global optimisation. *Water SA*, **25**(3), 317-326.
- [176] **Nelder, J.A. et Mead, R. (1965)**. A Simplex method for function minimisation. *The Computer Journal*, **7**(4), 308-313.
- [177] **Newton, D.W. et Vinyard, J.W. (1967)**. Computer-determined Unit Hydrograph from flows. *J. Hydraul. Div. ASCE*, **93**(HY5), 219-235.
- [178] **Nielsen, S.A. et Hansen, E. (1973)**. Numerical simulation of the rainfall-runoff process on a daily basis. *Nordic Hydrology*, **4**, 171-190.
- [179] **O'Connell, P.E., Nash, J.E. et Farrell, J.P. (1970)**. River flow forecasting through conceptual models. Part II - The Brosna catchment at Ferbane. *Journal of Hydrology*, **10**, 317-329.
- [180] **O'Donnell, T. et Canedo, P. (1980)**. The reliability of conceptual basin model calibration. *Proceedings of the Oxford Symposium on Hydrological Forecasting, IAHS Publication n° 129*, 263-269.
- [181] **Oreskes, N., Shrader-Frechette, K. et Belitz, K. (1994)**. Verification, validation and confirmation of numerical models in the Earth Sciences. *Science*, **263**, 644-646.
- [182] **Ouédraogo, M. (1996)**. Caractérisation d'une modification éventuelle de la relation pluie-débit autour des années 1970 en Côte d'Ivoire. *Mémoire de DEA Sciences de l'Eau dans l'Environnement Continental, Université de Montpellier*, 56 p.
- [183] **Ouédraogo, M., Servat, E., Paturel, J.E., Lubès-Niel, H. et Masson, J.M. (1998)**. Caractérisation d'une modification éventuelle de la relation pluie-débit autour des années 1970 en Afrique de l'ouest et centrale non-sahélienne. *Proc. of the Abidjan Conf., Ivory Coast, Water Resources Variability in Africa during the XXth Century, IAHS Publication n° 252*, 315-321.
- [184] **Pardé, M. (1933)**. Fleuves et Rivières, *Paris*, 224 p.

- [185] **Park, J., Kang, I.S. et Singh, V.P. (1999).** Comparison of simple runoff models used in Korea for small watersheds. *Hydrological Processes*, **13**(10), 1527-1540.
- [186] **Paturel, J.E., Servat, E. et Vassiliadis, A. (1995).** Sensitivity of conceptual rainfall-runoff algorithms to errors in input data - case of the GR2M model. *Journal of Hydrology*, **168**, 11-125.
- [187] **Penman, H.L. (1948).** Natural evaporation from open water, bare soil and grass. *Proc. R. Soc. London*, **A193**, 120-145.
- [188] **Perrin, C. (1997).** Comparative assessment of two rainfall-runoff modelling approaches: GR4J and IHACRES. *Mémoire de DEA, ULP/ENGEES/Institute of Hydrology*, 153 p.
- [189] **Perrin, C. et Littlewood, I.G. (2000).** A comparative assessment of two rainfall-runoff modelling approaches: GR4J and IHACRES, *Proceedings of the Liblice Conference (22-24 September 1998)*, V. Elias and I.G. Littlewood (Eds.), *IHP-V, Technical Documents in Hydrology n° 37, UNESCO, Paris*, 191-201.
- [190] **Pickup, G. (1977).** Testing the efficiency of algorithms and strategies for automatic calibration of rainfall-runoff models. *Hydrological Sciences Bulletin*, **XXII**(2), 257-274.
- [191] **Pinault, J.L., Pauwels, H., Fritsche, V. et Cann, C. (1997).** Mechanisms affecting streamflow and stream water quality: a new spectral analysis of time-series approach. *IHP-V, Technical Documents in Hydrology n° 14, UNESCO, Paris*, 77-84.
- [192] **Pirt, J. et Bramley, E.A. (1985).** The application of simple moisture accounting models to ungauged catchments. *Journal of the Institution of Water Engineers and Scientists*, **39**, 169-177.
- [193] **Porter, J.W. et McMahon, T.A. (1971).** A model for the simulation of streamflow data from climatic records. *Journal of Hydrology*, **13**, 297-324.
- [194] **Post, D.A. et Jakeman, A.J. (1996).** Relationships between catchments attributes and hydrological response characteristics in small Australian mountain ash catchments. *Hydrological Processes*, **10**(6), 877-892.
- [195] **Post, D.A. et Jakeman, A.J. (1999).** Predicting the daily streamflow of ungauged catchments in S.E. Australia by regionalising the parameters of a lumped conceptual model. *Ecological Modelling*, **123**, 91-104.
- [196] **Press, W.H., Teukolsky, S.A., Vetterling, W.T. et Flannery, B.P. (1992).** Numerical recipes in Fortran. *Cambridge University Press (Ed.)*, Cambridge, 702 p.
- [197] **Pronzato, L., Walter, E., Venot, A. et Lebruchec, J.F. (1984).** A general-purpose global optimizer: implementation and applications. *Math. Comput. Simul.*, **26**, 412-422.
- [198] **Quesney, A. (1999).** Assimilation de mesures d'humidité de surface dans un modèle hydrologique conceptuel global. Apport de la télédétection radar ERS/SAR. *Thèse de Doctorat, Université de Paris 7 / UFR de Physique*, 179 p.
- [199] **Rakem, Y. (1999).** Analyse critique et reformulation mathématique d'un modèle empirique pluie-débit (GR4J). *Thèse de Doctorat, ENPC*, 263 p.
- [200] **Raper, G.P. et Kuczera, G. (1991).** Groudwater recharge estimation using a lumped parameter catchment process model. *First International Hydrology and Water Resources Symposium, Perth., Inst. Eng. Aust., Canberra, Nat. Conf. Publ.*, 91/22, 563-568.
- [201] **Refsgaard, J.C. et Abbott, M.B. (1996).** The role of distributed hydrological modelling in water resources management - Chapter 1. In: *Distributed Hydrological modelling*, M. B. Abbott et J. C. Refsgaard (Ed.), *Kluwer Academic Publishers, Netherlands*, 1-16.
- [202] **Refsgaard, J.C. et Knudsen, J. (1996).** Operational validation and intercomparison of different types of hydrological models. *Water Resources Research*, **32**(7), 2189-2202.
- [203] **Reimers, W. (1990).** Estimating hydrological parameters from basin characteristics for large semiarid catchments. *IAHS Publication n° 191*, 187-194.

- [204] **Roche, P.A. (1988)**. Les modèles. In: *Gestion des Eaux: Automatisation, Information et Télégestion*, F. Valiron (Ed.), Presses de l'Ecole nationale des Ponts et Chaussées, **III**, 138-166.
- [205] **Rodriguez, J.Y., Sempere-Torres, D. et Obled, C. (1989)**. Nouvelles perspectives de développement dans la modélisation des pluies efficaces par applications de la méthode DPFT. *IAHS Symposium, Baltimore, IAHS Publication n° 181*, 235-244.
- [206] **Rodriguez, J.Y., Sempere-Torres, D. et Obled, C. (1991)**. Extension of lumped operational rainfall-runoff approach models to semilumped modelling: the case of the DPFT-ERHUDIT approach - Chapter 43. *Proceedings of the International Symposium on Hydrological Applications of Weather Radar, Salford, 14-17 August 1989*.
- [207] **Rosenbrock, H.H. (1960)**. An automatic method for finding the greatest or least value of a function. *The Computer Journal*, **3**, 175-184.
- [208] **Salas, J.D. (1993)**. Analysis and modeling of hydrologic time series. Chapter 19. In: *Handbook of Hydrology*, D. R. Maidment (Ed.), McGraw Hill, 1412 p.
- [209] **Saporta, G. (1990)**. Probabilités - Analyse de données et statistiques. *Editions Technip (Ed.)*, 493 p.
- [210] **Sarma, P.B.S., Delleur, J.W. et Rao, A.R. (1973)**. Comparison of rainfall-runoff models for urban area. *Journal of Hydrology*, **18**, 329-347.
- [211] **Scavia, D., Powers, W.F., Canale, R.P. et Moody, J.L. (1981)**. Comparison of first-order error analysis and Monte Carlo simulation in time-dependent lake eutrophication models. *Water Resources Research*, **17**(4), 1051-1059.
- [212] **Schermerhorn, V.P. et Kuehl, D.W. (1968)**. Operational forecasting with the SSARR model. *IAHS Publication n° 80*, 317-328.
- [213] **Schittowski, K. (1985)**. NLQP: A Fortran subroutine solving constrained linear programming problems. *Operations Res.*, **5**, 485-500.
- [214] **Sefe, F.T. et Boughton, W.C. (1982)**. Variation of model parameter values and sensitivity with type of objective function. *Journal of Hydrology*, **21**(2), 117-132.
- [215] **Sefton, C.E.M., Whitehead, P.G., Eatherall, A., Littlewood, I.G. et Jakeman, A.J. (1995)**. Dynamic response characteristics of the Plynlimon catchments and preliminary analysis of relationships to physical descriptors. *Environmetrics*, **6**, 465-472.
- [216] **Seibert, J. (1997)**. Estimation of parameter uncertainty in the HBV model. *Nordic Hydrology*, **28**(4/5), 247-262.
- [217] **Seibert, J. (1999)**. Regionalisation of parameters for a conceptual rainfall-runoff model. *Agricultural and Forest Meteorology*, **98-99**, 279-293.
- [218] **Servat, E. (1986)**. Présentation de trois modèles globaux conceptuels déterministes: CREC 5, MODGLO, MODIBI, ORSTOM, Département F, Unité de Recherche 604.
- [219] **Servat, E. et Dezetter, A. (1991)**. Selection of calibration objective functions in the context of rainfall-runoff modelling in a sudanese savannah area. *Hydrological Sciences Journal*, **36**(4), 307-331.
- [220] **Servat, E. et Dezetter, A. (1992)**. Modélisation de la relation pluie-débit et estimation des apports en eau dans le nord-ouest de la Côte d'Ivoire. *Hydrologie Continentale*, **7**(2), 129-142.
- [221] **Servat, E. et Dezetter, A. (1993)**. Rainfall-runoff modelling and water resources assessment in northwestern Ivory Coast. Tentative extension to ungauged catchments. *Journal of Hydrology*, **148**, 231-248.
- [222] **Servat, E., Paturel, J.E. et Lubès, H. (1996)**. La sécheresse gagne l'Afrique tropicale. *La Recherche*, **290**, 24-25.
- [223] **Servat, E., Paturel, J.E., Lubès, H., Kouamé, B., Ouedraogo, M. et Masson, J.M. (1997)**. Climatic variability in humid Africa along the Gulf of Guinea. Part I: detailed analysis of the phenomenon in Côte d'Ivoire. *Journal of Hydrology*, **191**, 1-15.

- [224] **Shamseldin, A.Y. (1997)**. Application of neural network technique to rainfall-runoff modelling. *Journal of Hydrology*, **199**, 272-294.
- [225] **Shamseldin, A.Y. et O'Connor, K.M. (1999)**. A real-time combination method for the outputs of different rainfall-runoff models. *Hydrological Sciences Journal*, **44**(6), 895-912.
- [226] **Shamseldin, A.Y., O'Connor, K.M. et Liang, G.C. (1997)**. Methods for combining the outputs of different rainfall-runoff models. *Journal of Hydrology*, **197**, 203-229.
- [227] **Sherman, L.K. (1932)**. Streamflow from rainfall by the unit hydrograph method. *Eng. News-Record*, **108**, 501-505.
- [228] **Singh, K.P. (1976)**. Unit hydrographs - a comparative study. *Water Resources Bulletin*, **12**(2), 381-392.
- [229] **Sivapalan, M., Ruprecht, J.K. et Viney, N.R. (1996)**. Water and salt balance modelling to predict the effects of land-use changes in forested catchments. 1. Small catchment water balance model. *Hydrological Processes*, **10**(3), 393-411.
- [230] **Sorooshian, S. (1980)**. Comparison of two direct search algorithms used in calibration of rainfall-runoff models. *Proceedings of the IFAC Conference on Water and Related Resources Systems, Cleveland, Ohio*, 477-485.
- [231] **Sorooshian, S. (1981)**. Parameter estimation of rainfall-runoff models with heteroscedastic streamflow errors - the non-informative data case. *Journal of Hydrology*, **52**, 127-138.
- [232] **Sorooshian, S. (1985)**. Synthesis of hydrologic and system sciences in the development of rainfall-runoff models. *Applied mathematics and Computation*, **17**, 279-298.
- [233] **Sorooshian, S. et Arfi, F. (1982)**. Response surface parameter sensitivity analysis methods for postcalibration studies. *Water Resources Research*, **18**(5), 1531-1538.
- [234] **Sorooshian, S. et Dracup, J.A. (1980)**. Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: correlated and heteroscedastic error cases. *Water Resources Research*, **16**(2), 430-442.
- [235] **Sorooshian, S., Duan, Q. et Gupta, V.K. (1993)**. Calibration of rainfall-runoff models: application of global optimization to the Sacramento soil moisture accounting model. *Water Resources Research*, **29**(3), 1185-1194.
- [236] **Sorooshian, S. et Gupta, V. (1985)**. The analysis of structural identifiability: theory and applications of conceptual rainfall-runoff models. *Water Resources Research*, **21**(4), 487-495.
- [237] **Sorooshian, S. et Gupta, V.K. (1983)**. Automatic calibration of conceptual rainfall-runoff models: the question of parameter observability and uniqueness. *Water Resources Research*, **19**(1), 260-268.
- [238] **Sorooshian, S. et Gupta, V.K. (1995)**. Model calibration. In: *Computer models of watershed hydrology, Chapter 2*, V. P. Singh (Ed.), *Water Resources Publications*, 23-68.
- [239] **Sorooshian, S., Gupta, V.K. et Fulton, J.L. (1983)**. Evaluation of maximum likelihood estimation techniques for conceptual rainfall-runoff models: influence of calibration data variability and length on model credibility. *Water Resources Research*, **19**(1), 251-259.
- [240] **Speers, D.D. (1995)**. SSARR Model. In: *Computer Models of Watershed Hydrology, Chapter 11*, V. P. Singh (Ed.), *Water Resources Publications*, 367-394.
- [241] **Srikanthan, R. et Goodspeed, M.J. (1988)**. Regionalization of conceptual model parameters for meso-scale catchments in the Hunter Valley. *Hydrology and Water Resources Symposium 1988, ANU, Canberra*, 85-90.
- [242] **Steeffel, C.I. et Van Cappellen, P. (1998)**. Reactive transport modeling of natural systems. *Journal of Hydrology*, **209**, 1-7.
- [243] **Sugawara, M. (1979)**. Automatic calibration of the tank model. *Hydrological Sciences Journal*, **24**(3), 375-388.
- [244] **Sugawara, M. (1995)**. Tank model. In: *Computer Models of Watershed Hydrology, Chapter 7*, V. P. Singh (Ed.), *Water Resources Publications*, 165-214.

- [245] **Summer, N.R., Fleming, P.M. et Bates, B.C. (1997).** Calibration of a modified SFB model for twenty-five Australian catchments using simulated annealing. *Journal of Hydrology*, **197**, 166-188.
- [246] **Szöllosi-Nagy, A. (1976).** An adaptative identification and prediction algorithm for the real-time forecasting of hydrological time series. *Hydrological Sciences Bulletin*, **XXI**(1), 163-176.
- [247] **Tan, B.Q. et O'Connor, K.M. (1996).** Application of an empirical infiltration equation in the SMAR conceptual model. *Journal of Hydrology*, **185**, 275-295.
- [248] **Tanakamaru, H. (1995).** Parameter estimation for the Tank Model using global optimisation. *Trans. JSIDRE*, **178**, 103-112.
- [249] **Tangara, M. (1991).** Un ensemble de programmes de traitement de fichiers hydrologiques et de tracé graphique, réalisés en Fortran 77, en vue d'une typologie des petits bassins versants ruraux. *Mémoire d'ingénieur, CNAM/CEMAGREF*, 207 p.
- [250] **Thiery, D. (1982).** Utilisation d'un modèle global pour identifier sur un niveau piézométrique des influences multiples dues à diverses activités humaines. *IAHS Publication n° 136*, 71-77.
- [251] **Thom, R. (1979).** Modélisation et scientificité. *Actes du colloque Elaboration et justification des modèles: applications en biologie, Paris, Maloine Editeur, Tome I*, 21-29.
- [252] **Thomas, H.A. (1981).** Improved methods for rational water assessment. *Report, Contract WR15249270, US Water Resources Council, Washington, DC, USA*.
- [253] **Thorntwaite, C.W. et Mather, J.R. (1955).** The water balance. *Publ. Climatol. Lab. Climatol. Drexel. Inst. Techol.*, **8**(1), 1-104.
- [254] **Thurman, J.L. et Roberts, R.T. (1995).** New strategies for the Water Data Center. *Journal of Soil and Water Conservation*, **50**(5), 530-531.
- [255] **Thyer, M., Kuczera, G. et Bates, B.C. (1999).** Probabilistic optimization for conceptual rainfall-runoff models: a comparison of the shuffled complex evolution and simulated annealing algorithms. *Water Resources Research*, **35**(3), 767-773.
- [256] **Tixeront, J. (1964).** Prévision des apports des cours d'eau. *Symposium sur les Eau de surface, Assemblée générale de l'IUGG, Berkeley*.
- [257] **Todini, E. (1988).** Rainfall-runoff modelling - Past, present and future. *Journal of Hydrology*, **100**, 341-352.
- [258] **Todini, E. (1996).** The ARNO rainfall-runoff model. *Journal of Hydrology*, **175**, 339-382.
- [259] **Troutman, B.M. (1985a).** Errors and parameter estimation in precipitation-runoff modeling. 1. Theory. *Water Resources Research*, **21**(8), 1195-1213.
- [260] **Troutman, B.M. (1985b).** Errors and parameter estimation in precipitation-runoff modeling. 2. Case study. *Water Resources Research*, **21**(8), 1214-1222.
- [261] **Tsykin, E.N. (1985).** Multiple nonlinear statistical models for runoff simulation and prediction. *Journal of Hydrology*, **77**, 209-226.
- [262] **Tung, Y.K., Yeh, K.C. et Yang, J.C. (1997).** Regionalization of unit hydrograph parameters: 1. Comparison of regression analysis techniques. *Stochastic Hydrology and Hydraulics*, **11**(2), 145-171.
- [263] **Uhlenbrook, S., Seibert, J., Leibundgut, C. et Rodhe, A. (1999).** Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. *Hydrological Sciences Journal*, **44**(5), 779-797.
- [264] **Van Straten, G. et Keesman, K.J. (1991).** Uncertainty propagation and speculation in projective forecasts of environmental change: a lake-eutrophication example. *Journal of Forecasting*, **10**, 163-190.
- [265] **Vandewiele, G.L. et Elias, A. (1995).** Monthly water balance of ungauged catchments obtained by geographical regionalization. *Journal of Hydrology*, **170**, 277-291.

- [266] Vandewiele, G.L., Xu, C.Y. et Huybrecht, W. (1991). Regionalisation of physically-based water balance models in Belgium. Application to ungauged catchments. *Water Resources Management*, **5**, 199-208.
- [267] Vandewiele, G.L., Xu, C.Y. et Win, N.L. (1992). Methodology and comparative study of monthly models in Belgium, China and Burma. *Journal of Hydrology*, **134**, 315-347.
- [268] Vemuri, V., Dracup, J.A., Erdman, R.C. et Vemuri, N. (1969). Sensitivity analysis method of system identification and its potential in hydrologic research. *Water Resources Research*, **5**(2), 341-349.
- [269] Wang, Q.J. (1991). The genetic algorithm and its application to calibrating conceptual rainfall-runoff models. *Water Resources Research*, **27**(9), 2467-2471.
- [270] Wang, Q.J. (1997). Using genetic algorithms to optimise model parameters. *Environmental Modelling and Software*, **12**(1), 27-34.
- [271] Warmerdam, P.M.M., Kole, J. et Chormanski, J. (1997). Modelling rainfall-runoff processes in the Hupselse Beek research basin. *Ecohydrological processes in small basins, Proceedings of the Strasbourg Conference (24-26 September 1996), IHP-V, Technical Documents in Hydrology n° 14, UNESCO, Paris*, 155-160.
- [272] Weeks, W.D. et Boughton, W.C. (1987). Tests of ARMA model forms for rainfall-runoff modelling. *Journal of Hydrology*, **91**, 29-47.
- [273] Weeks, W.D. et Hebbert, R.H.B. (1980). A comparison of rainfall-runoff models. *Nordic Hydrology*, **11**, 7-24.
- [274] Weglarczyk, S. (1998). The interdependence and applicability of some statistical quality measures for hydrological models. *Journal of Hydrology*, **206**, 98-103.
- [275] Wheater, H.S., Jakeman, A.J. et Beven, K.J. (1993). Progress and directions in rainfall-runoff modelling - Chapter 5. In: *Modelling Change in Environmental Systems*, A. J. Jakeman, M. B. Beck et M. J. McAleer (Ed.), John Wiley & Sons Ltd, 101-132.
- [276] Wilcox, B.P., Rawls, W.J., Brakensiek, D.L. et Wight, J.R. (1990). Predicting runoff from rangeland catchments: a comparison of two models. *Water Resources Research*, **26**(10), 2401-2410.
- [277] Willmott, C.J. (1981). On the validation of models. *Physical Geography*, **2**(2), 184-194.
- [278] WMO (1975). Intercomparison of conceptual models used in operational hydrological forecasting. *Operational Hydrology Report n° 7, WMO n° 429, World Meteorological Organization, Geneva, Switzerland*.
- [279] WMO (1986). Intercomparison of models of snowmelt runoff. *Operational Hydrology Report n° 23, WMO n° 646, World Meteorological Organization, Geneva, Switzerland*.
- [280] WMO (1992). Simulated real-time intercomparison of hydrological models. *Operational Hydrology Report n° 38, World Meteorological Organization, Geneva, Switzerland*.
- [281] Xu, C.Y. et Vandewiele, G.L. (1995). Parsimonious monthly rainfall-runoff models for humid basins with different input requirements. *Advances in Water Resources*, **18**, 39-48.
- [282] Yan, J. et Haan, C.T. (1991a). Multiobjective parameter estimation for hydrologic models - Multiobjective programming. *Transactions of the ASAE*, **34**(3), 848-856.
- [283] Yan, J. et Haan, C.T. (1991b). Multiobjective parameter estimation for hydrologic models - Weighting of errors. *Transactions of the ASAE*, **34**(1), 135-141.
- [284] Yang, D., Herath, S. et Musiak, K. (2000). Comparison of different distributed hydrological models for characterization of catchment spatial variability. *Hydrological Processes*, **14**(3), 403-416.
- [285] Yang, X. (1993). Mise au point d'une méthode d'utilisation d'un modèle pluie-débit conceptuel pour la prévision des crues en temps réel. *Thèse de Doctorat, ENPC/CERGRENE*, 350 p.
- [286] Yang, X. et Michel, C. (2000). Flood forecasting with a watershed model: a new method of parameter updating. *Hydrological Sciences Journal*, **45**(4), 537-546.

- [287] **Yang, X., Parent, E., Michel, C. et Roche, P.A. (1995).** Comparison of real-time reservoir-operation techniques. *Journal of Water Resources Planning and Management*, **121**(5), 345-351.
- [288] **Yapo, P.O., Gupta, H.V. et Sorooshian, S. (1996).** Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. *Journal of Hydrology*, **181**, 23-48.
- [289] **Yapo, P.O., Gupta, H.V. et Sorooshian, S. (1998).** Multi-objective global optimization for hydrologic models. *Journal of Hydrology*, **204**, 83-97.
- [290] **Ye, W., Bates, B.C., Viney, N.R., Sivapalan, M. et Jakeman, A.J. (1997).** Performance of conceptual rainfall-runoff models in low-yielding ephemeral catchments. *Water Resources Research*, **33**(1), 153-166.
- [291] **Yeh, K.C. et Tung, Y.K. (1993).** Uncertainty and sensitivity analyses of pit-migration model. *Journal of Hydraulic Engineering*, **119**, 262-283.
- [292] **Yu, P.S. et Yang, T.C. (2000).** Using synthetic flow duration curves for rainfall-runoff model calibration at ungauged sites. *Hydrological Processes*, **14**(2), 117-133.
- [293] **Zealand, C.M., Burn, D.H. et Simonovic, S.P. (1999).** Short term streamflow forecasting using artificial neural networks. *Journal of Hydrology*, **214**, 32-48.
- [294] **Zermani, A. (1998).** Apport des SIG à la reconnaissance à moyenne échelle des facteurs d'écoulement et de transfert des nitrates. *Thèse de Doctorat, ENGREF, Cemagref*, 377 p.
- [295] **Zhang, X. et Lindström, G. (1996).** A comparative study of a Swedish and a Chinese hydrological model. *Water Resources Bulletin*, **32**(5), 985-994.
- [296] **Zhao, R.J. et Liu, X.R. (1995).** The Xinanjiang model. In: *Computer models of Watershed Hydrology, Chapter 7, V. P. Singh (Ed.), Water Resources Publications*, 215-232.
- [297] **Zhao, R.J., Zuang, Y.L., Fang, L.R., Liu, X.R. et Zhang, Q.S. (1980).** The Xinanjiang model. *IAHS Publication n° 129*, 351-356.

