



PREVISION DES CRUES AUX CONFLUENCES

Impact de la spatialisation sur les performances des modèles hydrologiques de prévision

Rémi LAMBLIN

**Master hydrologie, hydrochimie, sols et environnement
Spécialité hydrologie**

Examineurs : Denis DARTUS (IMFT)
David LABAT (LMTG)

Maîtres de stage : Yan LACAZE (DIREN Ile-de-France)
Charles PERRIN & Lionel BERTHET (Cemagref Antony)

Septembre 2009

Remerciements

J'ai réalisé mon stage de master au sein du Service de Préviation des Crues Seine Moyenne Yonne Loing, hébergé par la DIREN Ile-de-France à Gentilly, et au Cemagref d'Antony. Avant de présenter les résultats de mes travaux, je souhaite remercier certaines personnes.

Je tiens à remercier tout particulièrement Yan LACAZE, responsable de la cellule "Connaissance et expertise hydrologique" à la DIREN Ile-de-France, Charles PERRIN et Lionel BERTHET, chercheurs au sein de l'équipe hydrologie au Cemagref d'Antony. Sur un plan humain, leur accueil a été fantastique. Leur disponibilité et bonne humeur ont transformé ce travail en un enrichissement permanent très agréable. Sur un plan technique, leurs conseils avisés, leur regard critique et nos "débat hydrologiques" m'ont permis de progresser dans ma démarche. Ce compromis entre autonomie et réflexion sur les orientations prises m'a constamment encouragé et a largement contribué à maintenir intacts ma motivation et mon intérêt pour l'étude effectuée.

Je remercie également toute l'équipe hydrologie du Cemagref, avec une mention spéciale pour Vazken ANDRÉASSIAN qui a notamment contribué à rendre ce partenariat entre la DIREN Ile-de-France et le Cemagref possible. La bonne humeur au sein de cette équipe à Antony a été très agréable, merci donc à Annie, Audrey, Gianluca, Houda, Juan-Alberto, Julien, Laurent, Maria-Helena, Marine, Raji, Thomas-Charles et Vincent.

De très chaleureux remerciements vont à toute l'équipe du SHYRN à la DIREN Ile-de-France, qui m'a extrêmement bien accueilli. Travailler dans ce service a été un réel plaisir. Un grand merci tout particulier à Emmanuel, Jean-Pedro, Marc, Olivier, Sandrine et Sylvain pour leur accueil et leur aide, ainsi qu'à l'équipe "risques", Adèle, Christelle et Nathalie. Merci aussi à toute l'équipe hydrométrie, notamment Arnaud, Christophe, Damien, José, Patrick et Vincent. Une pensée toute particulière va à mes deux collègues stagiaires Hélène et Xavier. Enfin, merci à Fabrice DALY pour son accueil au sein du SHYRN.

Toutes ces personnes ont contribué à rendre ce stage extrêmement enrichissant, intéressant, et des plus agréables. J'aurais d'ailleurs grand plaisir à pouvoir dans le futur travailler avec eux sur des projets d'intérêt commun si l'opportunité se présente.

Résumé

La question de l'intérêt de la distribution spatiale en modélisation hydrologique compte parmi les débats majeurs des dernières décennies. Dans cette étude, on s'intéresse plus particulièrement à la prévision des crues aux confluences, en évaluant l'apport de l'utilisation d'une version semi-distribuée du modèle GR3P modélisant le comportement des deux sous-bassins constitutifs de la confluence par rapport à une modélisation pluie-débit globale.

La méthodologie utilisée est adaptée de celle mise en place par [Andréassian *et al.* \(2004\)](#), mais cette fois-ci en prévision (et non pas en simulation). Deux échantillons de confluences sont utilisés. Un premier, qui constitue le cadre théorique de l'étude, est composé de bassins chimères, confluences virtuelles composées de l'assemblage strict de deux sous-bassins existants mais déconnectés géographiquement. Le second échantillon constitué de confluences réelles permet de confronter les résultats de la partie théorique à un cadre opérationnel de prévision des crues, avec notamment la présence d'un bassin versant intermédiaire entre les stations amont et aval comme élément supplémentaire déterminant par rapport aux confluences idéalisées.

Un net gain de performance induit par l'utilisation de la version semi-distribuée du modèle est observé pour les confluences chimères. En revanche, ce gain demeure très faible lorsque la confluence chimère est composée de bassins versants confluant réellement. Sans que les résultats ne le montrent de manière nette, il semble que le gain à attendre d'une version semi-distribuée du modèle augmente à mesure que la différence de comportement entre les deux sous-bassins constitutifs de la confluence s'accroît.

En ce qui concerne le cadre opérationnel de la prévision des crues aux confluences, un outil de propagation simple est utilisé dans le modèle semi-distribué pour tenir compte de l'effet du bassin intermédiaire. Un avantage se dégage pour la version semi-distribuée du modèle pour des horizons de prévision courts, mais diminue voire s'annule pour les horizons plus lointains. Ces résultats se sont cependant montrés assez sensibles aux modalités de calage du module de routage, en particulier l'horizon de prévision ciblé.

Mots clés : Modélisation hydrologique globale et semi-distribuée, prévision des crues, confluences, chimères, bassin intermédiaire, hétérogénéité de bassins

Abstract

The issue of spatial distribution in hydrological modelling remains one of the major debates over the last decades. In this study, we focused on flood forecasting at river confluences. We evaluated the performance gain that could be obtained by using a semi-distributed version of the GR3P model by modelling each of the upstream subcatchment separately in comparison with a lumped model version.

The methodology derives from the one proposed by [Andréassian *et al.* \(2004\)](#), but here test were made in forecasting mode instead of simulation mode. Two sets of confluences were built. The first one constitutes the theoretical side of the study and uses chimera confluences composed of two remote subcatchments having no actual connections. The second set is composed of real confluences, which include an intermediary catchment between the upstream and downstream stations. This second set gives the opportunity to confront the results obtained in a theoretical framework to those that could be obtained in an operational context.

For the chimera confluences, results were significantly improved by using the semi-distributed version. When the chimera is composed of subcatchments that actually constitute a confluence, the improvements are much more limited. The improvement that can be expected from the semi-distributed approach seems to increase with the heterogeneity of the two subcatchments, but this trend is quite weak.

On the set of actual confluences, a simple routing module was used in the semi-distributed version to account for the impact of the intermediary catchment. The semi-distributed version performs better than the lumped one for short lead times, but the improvement decreases when the lead-time increases. However, it was noticed that results were quite sensitive to the way the routing module was optimized, especially to the targeted lead-time.

Keywords : Lumped and semi-distributed hydrologic modeling, chimera watersheds, flood forecasting

SOMMAIRE

1. Introduction	1
2. Contexte, enjeux et objectifs	2
2.1. La prévision des crues	2
2.1.1. Missions	2
2.1.2. Moyens et outils utilisés	3
2.2. Les modèles hydrologiques	3
2.3. Différence prévision – simulation	4
2.4. La question de la distribution spatiale	5
2.5. Objectifs	6
3. Méthodologie	7
3.1. Modèle pluie-débit utilisé : le modèle GR3P	7
3.2. Différentes configurations de modélisation	9
3.3. Calage / contrôle du modèle	11
3.4. Critère de qualité des prévisions, le C2MP	11
3.5. Horizons de prévision	12
4. Echantillons d'application	13
4.1. Echantillon 1 : les confluences sans bassin intermédiaire	13
4.1.1. Sous-échantillon 1 : les bassins chimères	13
4.1.2. Sous-échantillon 2 : les confluences idéalisées	14
4.2. Echantillon 2 : les confluences réelles	14
4.2.1. Sélection des confluences	15
4.2.2. Analyse de l'influence du bassin intermédiaire	16
5. Résultats sur les confluences sans bassin intermédiaire	19
5.1. Résultats sur les bassins chimères	19
5.2. Résultats sur les confluences idéalisées	20
5.3. Analyse et exploitation	21
5.3.1. Influence du rapport de taille des sous-bassins	21
5.3.2. Rôle de l'hétérogénéité des sous-bassins	22
6. Résultats sur les confluences réelles	27
6.1. Résultats	27
6.2. Analyse de sensibilité	27
7. Conclusion et perspectives	29
8. Bibliographie	31

LISTE DES ILLUSTRATIONS

Figure 1 : Zone d'action du SPC SMYL.....	1
Figure 2 : Différentes approches de discrétisation spatiale	4
Figure 3 : Différence entre simulation (à gauche) et prévision (à droite).....	5
Figure 4 : Confluence théorique – Confluence réelle	7
Figure 5 : Schéma structurel du modèle GR3P avec représentation de la procédure d'exploitation des débits observés	8
Figure 6 : Tableau des différentes configurations du modèle testées dans cette étude	10
Figure 7 : Schéma de principe de la procédure de calage / contrôle utilisée	11
Figure 8 : Confluence chimère	13
Figure 9 : Localisation des 1070 stations hydrométriques disponibles dans la base de données du Cemagref (Le Moine, 2008).....	13
Figure 10 : Cas d'étude opérationnel par rapport au cas d'étude théorique	14
Figure 11 : Exemples d'hydrogrammes influencés (stations supprimées de l'échantillon)	15
Figure 12 : Localisation des confluences réelles étudiées	16
Figure 13 : Contribution du bassin intermédiaire dans la confluence étudiée.....	16
Figure 14 : Qualité de la simulation du débit en fonction de l'importance du bassin intermédiaire.....	17
Figure 15 : Qualité de la simulation à Neublans à partir des débits à Rochefort et Champagne (confluence Loue - Doubs).....	18
Figure 16 : Qualité de la simulation à Louviers à partir des débits à Cailly et Normanville (confluence Iton - Eure)	18
Figure 17 : Comparaison de la qualité (critère de Nash) de la simulation du débit à la station aval à partir des débits observés aux deux stations amont	18
Figure 18 : Performance des différentes configurations en fonction de l'horizon de prévision (calage à 12h)	19
Figure 19 : Distribution cumulée des résultats pour les différentes configurations (prévision à 12h).....	19
Figure 20 : Performances (en critère de persistance) du modèle GR3P en configuration semi-distribuée par rapport au même modèle en configuration globale.	20
Figure 21 : Comparaison du critère de performance entre les différentes configurations en fonction de l'horizon de prévision.....	20
Figure 22 : Distribution cumulée des résultats entre les différentes configurations (prévision à 12h).....	20
Figure 23 : Performances (en critère de persistance) du modèle GR3P en configuration semi-distribuée par rapport au même modèle en configuration globale – Délai de prévision : 12 heures.....	21
Figure 24 : Amélioration induite par l'utilisation du modèle semi-distribué en fonction des classes de rapport des débits moyens des deux sous-bassins constitutifs de la chimère (de 0 [bassins très différents] à 10 [bassins de même taille]).....	21
Figure 25 : Amélioration induite par l'utilisation du modèle semi-distribué en fonction du rapport des débits moyens des deux sous-bassins constitutifs du bassin confluence (rapport compris entre 0 [bassins très différents] à 1 [bassins de même taille]).....	22
Figure 26 : Graphes comparés de distribution des différents indicateurs d'hétérogénéité entre l'échantillon chimère et l'échantillon de confluences réelles	24

Figure 27 : Graphes représentant l'amélioration induite par l'utilisation d'un modèle semi-distribué en fonction des différents indicateurs d'hétérogénéité pour les deux échantillons de confluences (chimères et réels).....	25
Figure 28 : Graphes représentant l'amélioration induite par l'utilisation d'un modèle semi-distribué en fonction des différents indicateurs d'hétérogénéité – bassins chimères réunis en classes d'effectifs homogènes	26
Figure 29 : Performance des différentes configurations de modélisation en fonction de l'horizon de prévision	27
Figure 30 : Comparaison des temps de propagation obtenus en utilisant une célérité commune aux deux stations amont dans le module hydraulique par rapport à l'utilisation d'une célérité propre à chaque station amont.....	28
Figure 31 : Comparaison entre les temps de propagation obtenus avec le module hydraulique et ceux estimés par les prévisionnistes du SPC SMYL	28

LISTE DES TABLEAUX

Tableau 1 : Niveau de vigilance sur un tronçon de rivière (Source : site Internet vigicrue).....	2
Tableau 2 : Sélection d'études menées sur la question de l'apport de la distribution spatiale en modélisation hydrologique.....	6
Tableau 3 : Critère de Nash obtenu en modélisant le débit aval par "Q amont1+Q amont2" – Confluences à deux sous-bassins	17
Tableau 4 : Description des indicateurs d'hétérogénéité du bassin confluence	23

1. Introduction

En France, la prévision des crues est une mission assurée par l'État sur les cours d'eau à enjeux majeurs. Les outils dont disposent les Services de Prévision des Crues (SPC) évoluent très rapidement depuis quelques années. Toutefois, malgré les efforts de nombreuses équipes de météorologues, hydrauliciens et hydrologues, le stade d'une utilisation systématique d'un ou de plusieurs modèles fiables et robustes est encore loin d'être atteint. L'enjeu et l'objectif vers lequel tendre sont donc de les rendre de plus en plus compétitifs dans un cadre opérationnel en vue de convaincre les utilisateurs.

L'évolution constante des données disponibles, de plus en plus fines et spatialisées, laisse penser que les modélisations (semi-)distribuées, qui peuvent intégrer ces données comme entrées, ont a priori un potentiel plus important pour retranscrire avec plus de fiabilité la variabilité de comportement des bassins. Ce postulat mérite cependant d'être vérifié, les modèles distribués étant généralement plus complexes et par conséquent moins robustes.

L'enjeu du travail retranscrit ici est d'évaluer si la distribution d'un modèle présente un véritable intérêt, permettant d'entrevoir de meilleures performances, au niveau le plus basique de distribution, à savoir le découpage en deux sous-bassins. L'application concrète d'une telle configuration est la prévision des crues aux confluences, qui offrent un cadre opérationnel privilégié pour tester l'intérêt de la mise en place de cette configuration de semi-distribution de modèle.

Pour ce faire, deux échantillons distincts de confluences seront utilisés. D'une part, des confluences chimères, bassins virtuels composés de deux sous-bassins existants mais géographiquement déconnectés, (d'après la méthodologie d'Andréassian *et al.*, 2004) constitueront un cadre théorique idéalisé de prévision des crues aux confluences, permettant d'évaluer l'intérêt d'une modélisation semi-distribuée. D'autre part, un échantillon de confluences réelles permettra d'évaluer si les résultats théoriques se confirment dans un cadre opérationnel imposant de s'intéresser à la problématique du bassin intermédiaire, qui n'est pas présent dans le cadre théorique de confluences idéalisées. Pour cela, un modèle hydrologique global appliqué à chacun des bassins (ou sous-bassins) sera utilisé.

La zone de compétence du Service de Prévision des crues Seine Moyenne Yonne Loing (noté SPC SMYL par la suite) constituera un cas d'étude privilégié dans ce travail. Créé en 2006 et hébergé par la DIREN Ile-de-France à Gentilly, le SPC SMYL est issu d'un Service d'Annonce de Crue historique. Son territoire d'action (Fig. 1) compte parmi les plus étendus de France, comprenant toute la partie moyenne de la Seine, depuis la confluence avec l'Aube jusqu'à peu après la confluence avec l'Oise (avant la partie maritime de la Seine). Il englobe les bassins de l'Yonne et du Loing, principaux affluents de la Seine sur ce territoire, ainsi que la partie aval de la Marne.

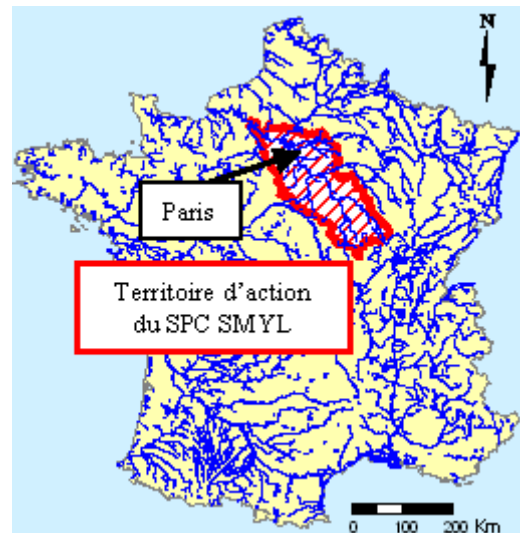


Figure 1 : Zone d'action du SPC SMYL

Dans la suite, une synthèse de la problématique et de l'état de l'art sur la question de la modélisation semi-distribuée sera proposée. La méthodologie employée, le modèle et les données utilisées seront ensuite présentés. Enfin, les résultats obtenus ainsi que les principales conclusions à tirer de ce travail seront détaillés.

2. Contexte, enjeux et objectifs

2.1. La prévision des crues

Dans le domaine des risques liés aux inondations, une des priorités est de prévenir de façon convenable du danger encouru. L'objectif est à la fois d'informer les citoyens, mais également et surtout les autorités locales (préfectures) qui gèrent la mise en œuvre des plans d'alerte et des plans de secours. Cette mission est assurée depuis 2003¹ par les Services de Prévisions des Crues (SPC) sur les cours d'eau surveillés par l'Etat. Au nombre de 22, ils sont hébergés par des services ministériels déconcentrés (DDEA, Services de Navigation, DIREN) ou des établissements publics (Météo France). En outre, un service technique central a été créé. Dénommé SCHAPI (Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations), il assure plusieurs missions : coordination des actions des SPC, appui technique, aide à la décision en cas de crise, harmonisation des outils pour l'ensemble des SPC et facilitation des relations avec les services de Météo France.

2.1.1. Missions

Les missions des SPC sont axées sur la production et la diffusion de l'information aux services en charge de la gestion de crise, comme au grand public. Les cours d'eau surveillés sont définis par le Schéma Directeur de Prévision des Crues, et sont divisés en tronçons hydrologiquement cohérents qui constituent l'entité de base du travail des prévisionnistes. Une ou plusieurs stations représentatives sont définies pour chaque tronçon. Les objectifs principaux sont :

- la vigilance de crue : elle consiste, pour chaque tronçon, à qualifier le risque de crues pour les 24 heures à venir. L'objectif est de le faire apparaître de la couleur correspondant au seuil qui sera franchi avec un jour d'anticipation. Le niveau de vigilance dépend du niveau (débit ou hauteur) atteint sur une ou plusieurs stations représentatives du risque sur le territoire concerné. L'information est ensuite portée sur une « carte de vigilance » nationale, mise à disposition du grand public sur le site : www.vigicrues.ecologie.gouv.fr. Lorsqu'aucune vigilance particulière n'est requise, la couleur verte est utilisée. Ensuite, la gradation de jaune à rouge est liée à l'importance des dommages potentiels sur le tronçon concerné (Tableau 1) ;





	Rouge : Risque de crue majeure. Menace directe et généralisée de la sécurité des personnes et des biens.
	Orange : Risque de crue génératrice de débordements importants susceptibles d'avoir un impact significatif sur la vie collective et la sécurité des biens et des personnes.
	Jaune : Risque de crue ou de montée rapide des eaux n'entraînant pas de dommages significatifs, mais nécessitant une vigilance particulière dans le cas d'activités saisonnières et/ou exposées.
	Vert : Pas de vigilance particulière requise.

Tableau 1 : Niveau de vigilance sur un tronçon de rivière (Source : site Internet vigicrue)

- la prévision de crue : en cas d'épisode de crue avéré, elle consiste à prévoir les niveaux qui vont être atteints, en vue d'une qualification plus précise de l'événement et d'une amélioration de l'anticipation. Cela permet notamment aux préfetures d'apprécier plus finement l'événement en cours et de mettre en œuvre les dispositifs de secours adéquats ;

¹ De 1984 à 2003, cette mission était prise en charge par 55 Services d'Annonce de Crues (SAC) [arrêtés du 27/02/1984]. Leurs rôles consistaient à surveiller les niveaux des cours d'eau et à informer les préfetures en cas de dépassements de seuils d'alerte. Celles-ci déclenchaient alors les procédures d'alerte et transmettaient l'information aux maires des communes concernées par le risque de crue. Les SPC, qui sont pour la plupart d'anciens SAC, sont nés d'une profonde réorganisation de l'annonce de crue en France [définie par les articles 41 et 72 de la loi « risques » du 30 juillet 2003].

- la mise à disposition en temps réel d'informations et de données hydrométriques sur serveur Internet ;
- un appui aux collectivités locales qui en formulent la demande, en liaison avec les préfetures, pour la mise en place d'un système de prévision/surveillance/alerte sur des cours d'eau non surveillés par les SPC.

2.1.2. Moyens et outils utilisés

Les SPC élaborent des formules empiriques et des abaques sur chaque station clé des tronçons pour remplir à bien les missions de vigilance et de prévision. Ils prennent en compte les pluies et débits observés ainsi que les précipitations prévues¹, calculant ensuite une valeur de débit (ou de hauteur) maximale qui sera atteinte durant les 24 prochaines heures. Si cette valeur dépasse les seuils fixés pour la station, le SPC passe en vigilance jaune, orange ou rouge. Bien entendu, les résultats obtenus avec ces formules et abaques sont pondérés par l'expérience des prévisionnistes, leur connaissance des bassins et l'appréciation du risque général que représente l'événement.

Les SPC disposent également de modèles hydrologiques et/ou hydrauliques. Lorsque le débit d'un cours d'eau surveillé atteint des valeurs approchant les seuils de vigilance (a fortiori lorsqu'elles les dépassent), les prévisionnistes utilisent ces modèles qui servent d'indication pour fournir des prévisions de débits (ou de hauteurs) sur le tronçon. Plus le modèle est fiable, plus le prévisionniste se fiera aux valeurs calculées, la confiance dans un modèle étant donc primordiale. Plusieurs types de modèles sont utilisés :

- des modèles débit-débit, utilisant les valeurs de débits mesurées aux stations en amont ;
- des modèles pluie-débit se basant sur les données des stations pluviométriques alentour ;
- des combinaisons de ces deux types de modèles.

2.2. Les modèles hydrologiques

L'étude présentée ici s'est restreinte à l'approche pluie-débit. Complémentaires des modèles débit-débit, les modélisations pluie-débit sont très couramment employées pour plusieurs raisons :

- elles sont utilisables en tête de bassin, pour les stations pour lesquelles aucune donnée hydrométrique amont n'est disponible ;
- elles peuvent permettre de modéliser les apports ponctuels ou diffus injectés dans un modèle hydraulique ;
- elles utilisent des entrées de pluie, pouvant pallier certaines difficultés d'ordre opérationnel rencontrées par les modèles débit-débit : si une station hydrométrique amont ne fonctionne pas, ou est momentanément en panne, le modèle devient quasiment inutilisable ;
- rarement meilleurs que les modèles débit-débit sur des délais de prévision courts (inférieurs au temps de propagation entre une station amont et une station aval), les modèles pluie-débit permettent de fournir des prévisions à des horizons bien plus importants, jusqu'au temps de réaction du bassin et même au delà dès lors qu'ils sont alimentés avec des prévisions de pluies futures fiables.

¹ Les prévisions de pluies sont fournies deux fois par jour par Météo France, sous forme de bulletins de précipitations.

Dans la pratique, un grand nombre de modèles pluie-débit existent (Singh *et al.*, 2002). Une personne désireuse de mettre en place ce type de modèles sur un bassin versant peut donc rapidement se retrouver perdue. De la même manière, les organismes pour lesquels la connaissance des débits en rivière est primordiale (prévision des crues, gestion des étiages, services de navigations etc..) dispose d'un choix très, voire trop large d'outils. Les différents modèles créés par de multiples équipes d'hydrologues à travers le monde peuvent être classés en plusieurs catégories, notamment en fonction de leur degré de distribution. Quelle que soit leur structure de base, les modèles peuvent être soit (Fig. 2) :

- globaux : traitant le bassin de manière globale, comme une seule entité, sans prise en compte explicite des hétérogénéités ;
- distribués : discrétisant le bassin en éléments de taille réduite, le plus souvent via un maillage régulier ;
- semi-distribués : état "intermédiaire", divisant le bassin en sous-bassins plus ou moins importants.

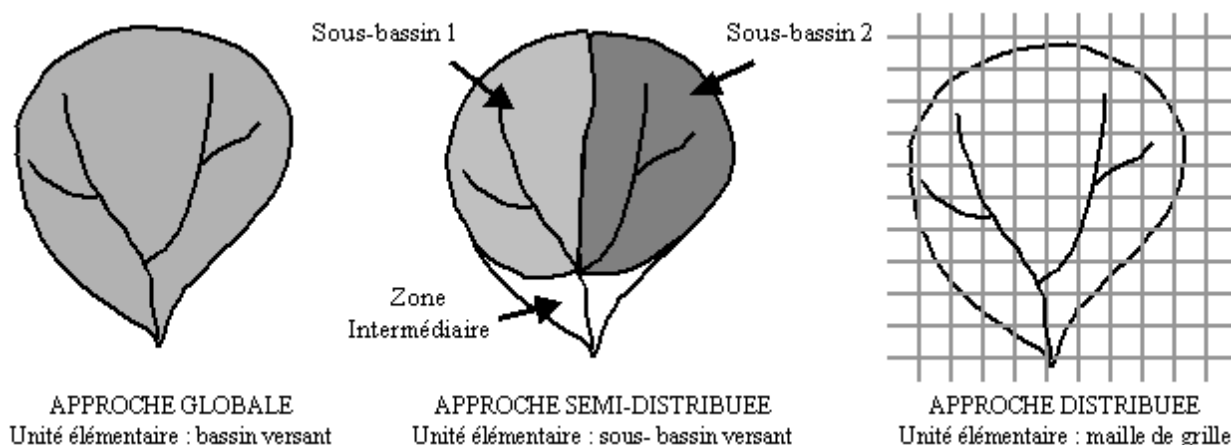


Figure 2 : Différentes approches de discrétisation spatiale

Chacune de ces approches comporte ses avantages et ses points faibles, et donc ses défenseurs et détracteurs. Les modèles distribués sont censés pouvoir mettre à profit les informations diverses notamment issues des radars (pluies) et des cartes représentant l'hétérogénéité des bassins. A l'inverse, les modèles globaux ne sont pas conçus pour exploiter ces informations mais sont moins gourmands en paramètres à caler, gagnant ainsi en robustesse et demeurant moins sensibles aux erreurs liées à l'actuelle imprécision des données spatialisées. La question de l'éventuel apport de la distribution et de la complexification des modèles hydrologiques compte donc depuis maintenant plus d'un demi-siècle parmi les débats majeurs en hydrologie. Le problème posé est d'évaluer le degré de distribution spatiale le plus adapté aux objectifs de la modélisation, à savoir simuler ou prévoir au mieux la réponse des bassins, et d'évaluer la sensibilité aux hétérogénéités de ceux-ci.

2.3. Différence entre prévision et simulation

La prévision des crues est un domaine qui, bien qu'intimement lié à celui de la simulation hydrologique, s'en distingue sur plusieurs points. La différence majeure est le fait de disposer en temps réel de données observées de débit au moment où s'effectue la prévision (Fig. 3). Cette observation est mise à profit pour corriger (et donc améliorer) les performances des modèles en temps réel. Cela passe par un processus appelé mise à jour par assimilation de données, *updating* en anglais (Yang *et al.*, 2000). Il permet d'établir un lien entre une information externe (une variable observée) et les variables calculées par le modèle afin de réduire l'erreur du modèle au cours de la prévision (Tangara, 2005). Il faut bien noter que simulation et prévision sont deux applications différentes et ne sont par conséquent que difficilement comparables. Un modèle de prévision est souvent issu de la fusion d'un modèle de simulation et d'un module d'assimilation de données. Cette assimilation peut se faire de plusieurs manières, en agissant sur les entrées du modèle, sur ses paramètres, sur ses états internes ou encore sur les sorties (Refsgaard *et al.*, 1997).

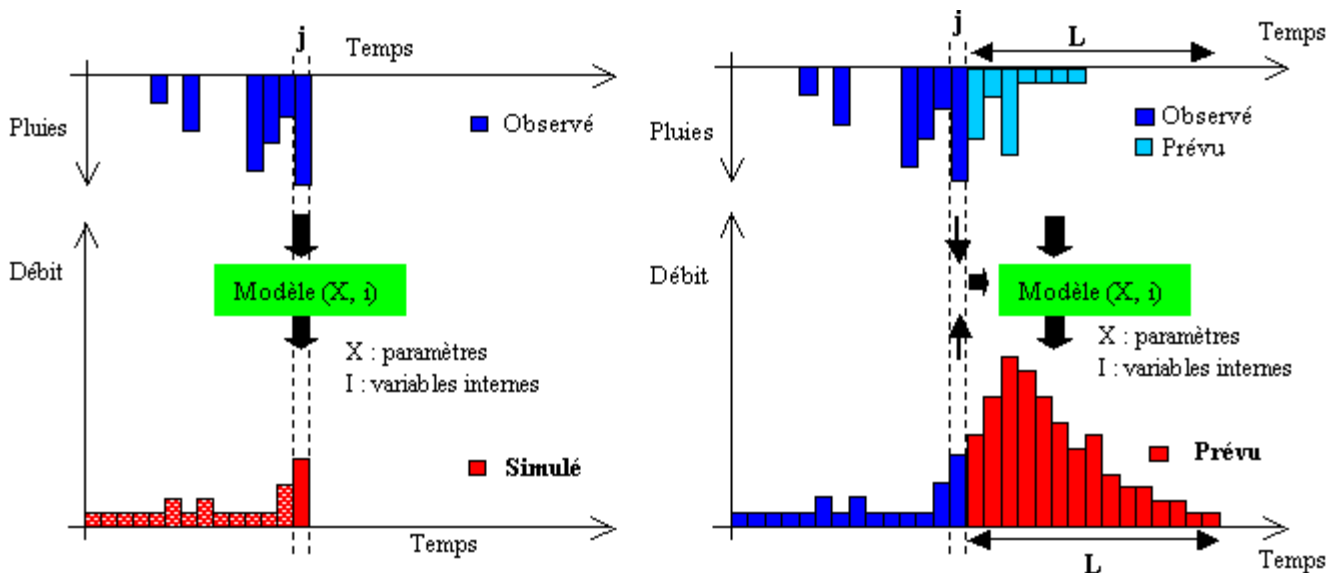


Figure 3 : Différence entre simulation (à gauche) et prévision (à droite)

2.4. La question de la distribution spatiale

De nombreuses études se sont, depuis plusieurs décennies, attachées à évaluer les évolutions de performance induites par l'utilisation d'informations spatialisées dans les modèles pluie-débit. Andréassian *et al.* (2004) ont notamment mené une analyse comparative sur des chimères, bassins versants virtuels constitués de deux bassins réels mais géographiquement déconnectés. L'objectif de ce procédé était de forcer le caractère hétérogène des bassins étudiés, l'amélioration à attendre de la distribution étant en effet a priori plus forte dans le cas de sous-bassins très hétérogènes. Un état des lieux des différentes études antérieures à 2003 portant sur la question de l'apport de la spatialisation y est également effectué. Ces études, menées dans un contexte de simulation, ont montré que l'amélioration apportée par la discrétisation spatiale n'est pas aussi importante qu'attendue. En outre, la spatialisation des pluies semble majoritairement être à l'origine des améliorations de la qualité des simulations, l'impact de la distribution des paramètres sur le bassin restant marginal (Boyle *et al.*, 2001).

Depuis 2003, le débat quant à l'intérêt et l'apport de la spatialisation / distribution des modèles hydrologiques se poursuit avec tout autant de vigueur. Le fait que les outils et moyens disponibles se développent et deviennent de plus en plus accessibles a renforcé l'engouement pour les modèles distribués, souvent gourmands en données d'entrée. C'est une composante supplémentaire qui s'est ajoutée au débat. En outre, la puissance des ordinateurs a augmenté, rendant possible la multiplication des calculs. Le temps de calcul pour des maillages très fins ou pour des jeux de paramètres conséquents à caler n'est plus autant pénalisant. De même, les données issues de radars ou de géosatellites se multiplient, se fiablent, et deviennent accessibles au format numérique (exploitable rapidement). Ainsi, de plus en plus de données (pas toujours validées) sont disponibles pour caractériser l'hétérogénéité des bassins versants (température, évapotranspiration, humidité du sol, occupation des sols etc.) à des pas d'espace relativement fins. Les Modèles Numériques de Terrain (MNT) deviennent également plus précis. De nouvelles études comparatives entre modèles globaux, semi-distribués et totalement distribués se sont donc de nouveau multipliées pour montrer l'intérêt de l'utilisation de modèles distribués. Le tableau 2 en recense certaines de manière non exhaustive avec leurs conclusions majeures.

Le revers de la médaille de l'utilisation de modèles distribués vient souvent du fait que les données spatialisées sur lesquelles se basent ces modèles ne sont pas toujours fiables. Certaines études ont montré que la qualité de ces modèles spatialisés s'en trouve parfois dégradée (Tableau 2). Certaines équipes de chercheurs se sont plus particulièrement attachées à montrer que l'incertitude sur les pluies spatialisées rend les modèles distribués moins robustes. Malgré tout, tous concluent que la prise en compte de la variabilité des pluies est fondamentale et source d'amélioration de la qualité des simulations d'hydrogramme à l'exutoire (pour les modèles (semi-)distribués au niveau des pluies).

Auteurs	Modèles Utilisés	Configuration	Conclusions
MODELISATIONS DISTRIBUEES PRESENTENT DE MEILLEURES PERFORMANCES			
Ajami et al. 2004	SAC-SMA Sacramento Soil Moisture Accounting	Globale, Semi-distribuée Distribuée	Performances supérieures pour la version distribuée Utiliser des jeux de paramètres différents sur les sous-bassins n'a pas été source d'améliorations significatives
El-Nasr et al. 2005	MIKE-SHE SWAT	Distribué Semi-distribué	Performances quasi-identiques pour les deux modèles Avantage à MIKE-SHE dans la retranscription des variations de débit
Carpenter et al. 2006	Adaptation de modèles utilisés par le US National Weather Service	Globale Distribuée	La version distribuée surpasse la version globale du modèle L'incertitude sur les prévisions de pluie peut rendre la version distribuée moins robuste
Khakbaz et al. 2009	SAC-SMA	Globale Semi-distribuée	L'utilisation d'informations de débits à l'intérieur du bassin dans un modèle semi-distribué entraîne un gain de performance
PAS D'AMELIORATION OBTENUE AVEC LES CONFIGURATIONS DISTRIBUEES			
Das et al. 2008	HBV	Globale, Semi-distribuée Distribuée	La version globale du modèle surpasse la version distribuée dans la majorité des cas La spatialisation des pluies, obtenue par krigeage, est une source d'erreur importante et pénalise le modèle distribué
Breuer et al. 2008	10 modèles testés dont MIKE-SHE, IHACRES, HBV	Globale (1) Semi-distribuée (5) Distribuée (4)	Performances très similaires pour tous les modèles Versions globales et semi-distribuées parfois supérieures, du fait des incertitudes de calage pour les modèles distribués
Kling et al. 2009	Water balance model	Distribué 1*1 km raster Semi-distribué	Pas de gain de performance pour la version totalement distribuée
McIntyre et al. 2009	10 modèles testés dont IHACRES	Globale Semi-distribuée	La prise en compte de la variabilité des pluies augmente les performances des modèles Les modèles avec le plus petit nombre de paramètres à caler sont plus robustes
VARIABILITE DES PLUIES JOUANT UN ROLE PREPONDERANT			
Lopez et al., 1996 Koren et al., 1998	SAC-SMA, SWB "Run-off erosion model"	Globale, Distribuée	L'incertitude inhérente aux pluies spatialisées limite le degré de spatialisation des modèles hydrologiques
Chaubey et al. 1999	H/WQ	Globale, Semi-distribuée	Prise en compte variabilité des pluies améliore la qualité des simulations Erreur sur les pluies pénalise le modèle semi-distribué
Smith et al. 2004	Modèles utilisés lors du "DMIP"	Globale, Distribué	Variabilité des pluies implique que les modèles avec un plus grand degré de spatialisation soient meilleurs que les versions globales

Tableau 2 : Sélection d'études menées sur la question de l'apport de la distribution spatiale en modélisation hydrologique

La vaste opération mise sur pied aux Etats-Unis par le National Weather Service pour comparer différents modèles distribués est un témoin fort de l'intérêt porté à la question de la spatialisation des modèles hydrologiques par la communauté scientifique. Nommée Distributed Models Intercomparison Project (DMIP), elle a confronté 12 modèles distribués sur différents bassins / chroniques / événements de crue (identiques pour tous les modèles). Un modèle global (inspiré du modèle utilisé par les National Weather Service River Forecast Centers) a également été testé pour servir de base de comparaison aux différentes performances et pour estimer l'amélioration apportée par l'utilisation d'un modèle distribué par rapport à un modèle global. L'analyse des différents résultats retranscrite par [Reed et al. \(2004\)](#), montre que le modèle global n'a été surpassé par les autres modèles que sur un faible nombre de bassins versants. Pour un petit nombre de bassins, de grosses améliorations ont cependant été constatées avec l'utilisation des modèles distribués. Des études supplémentaires sont nécessaires pour pouvoir étendre ces résultats à d'autres types de bassins, ce qui a fait l'objet d'un second projet (DMIP 2) dont les conclusions ne sont pas encore parues à l'heure actuelle. Cette étude a été menée en simulation, il reste donc à savoir si les conclusions sont les mêmes en prévision.

2.5. Objectifs

Le but de notre étude est de reprendre les travaux cités auparavant dans un contexte de prévision. En simulation, l'amélioration induite par la semi-distribution s'est avérée relativement faible, et surtout étroitement liée à la spatialisation des pluies (et non à la prise en considération des hétérogénéités de comportement des sous-bassins). La possibilité d'assimiler des données de débits en temps réel dans le modèle laisse penser que cette discrétisation spatiale pourra avoir un impact plus important en prévision. C'est ce que l'on veut tester et quantifier dans cette étude. L'idée est d'adapter la méthodologie générale mise en place par [Andréassian et al. \(2004\)](#) dans un contexte de prévision, des bassins chimères étant également utilisés pour forcer le caractère hétérogène des bassins.

En outre, la composante supplémentaire par rapport à l'étude de 2004 est de réaliser des tests sur des confluences réelles dans le prolongement de l'étude réalisée en simulation par Baudez *et al.* (1999), d'une part pour estimer si les conclusions tirées sont généralisables à un contexte opérationnel, d'autre part pour exploiter les conclusions en vue d'améliorer la qualité de la prévision des crues au niveau des confluences en utilisant un modèle pluie-débit semi-distribué.

Les tests sont donc réalisés sur de larges échantillons de bassins versants. Le système d'étude choisi est le cas particulier d'un bassin versant résultant d'une confluence de deux sous-bassins amont. Il sera analysé :

- dans un cadre théorique où le bassin aval est la stricte somme des deux sous-bassins amont, n'incluant donc pas le bassin intermédiaire, ces confluences idéalisées correspondant soit à des sous-bassins confluents réels, soit à des chimères telles qu'elles ont été définies par Andréassian *et al.* (2004), c'est-à-dire à la réunion de bassins éloignés géographiquement ;
- dans le cas de confluences réelles, comprenant un bassin intermédiaire entre les stations amont et la station aval, une attention particulière étant portée aux confluences situées sur la zone de compétence du SPC SMYL, au sein duquel a été réalisée cette étude. Un des objectifs de l'étude est l'amélioration des prévisions des crues aux confluences. Ainsi, travailler au SPC SMYL permet de profiter de l'expérience des prévisionnistes et de leur connaissance des bassins pour analyser au mieux les résultats, en ayant un niveau d'expertise supplémentaire par rapport aux autres confluences.

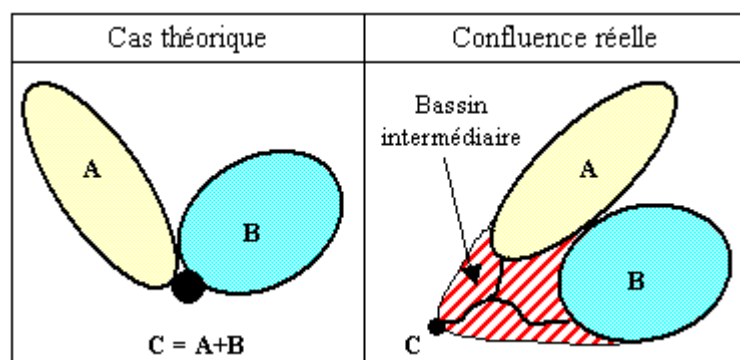


Figure 4 : Confluence théorique – Confluence réelle

3. Méthodologie

3.1. Modèle pluie-débit utilisé : le modèle GR3P

Le modèle hydrologique (pluie-débit) GR3P (pour Génie Rural à 3 paramètres pour la Prévision) a été choisi pour cette étude. Développé par le Cemagref, il a initialement été proposé par Tangara (2005) et est spécifiquement développé pour la prévision des crues. Utilisé dans plusieurs SPC français, c'est un modèle :

- continu (par opposition à un modèle événementiel) ;
- global (par opposition à un modèle spatialement distribué) ;
- ayant une structure à réservoirs ;
- construit pour une exploitation directe des débits observés ;
- dépendant de trois paramètres optimisables.

Il intègre un réservoir de production (de capacité fixe) fonctionnant en continu. Il génère une pluie efficace, multipliée par un coefficient d'ajustement des volumes - X_1 (adimensionnel)-. Le routage s'effectue via un hydrogramme unitaire -de temps de base X_3 (heures) - et un réservoir de routage quadratique - de capacité X_2 (millimètres) - (Fig. 5).

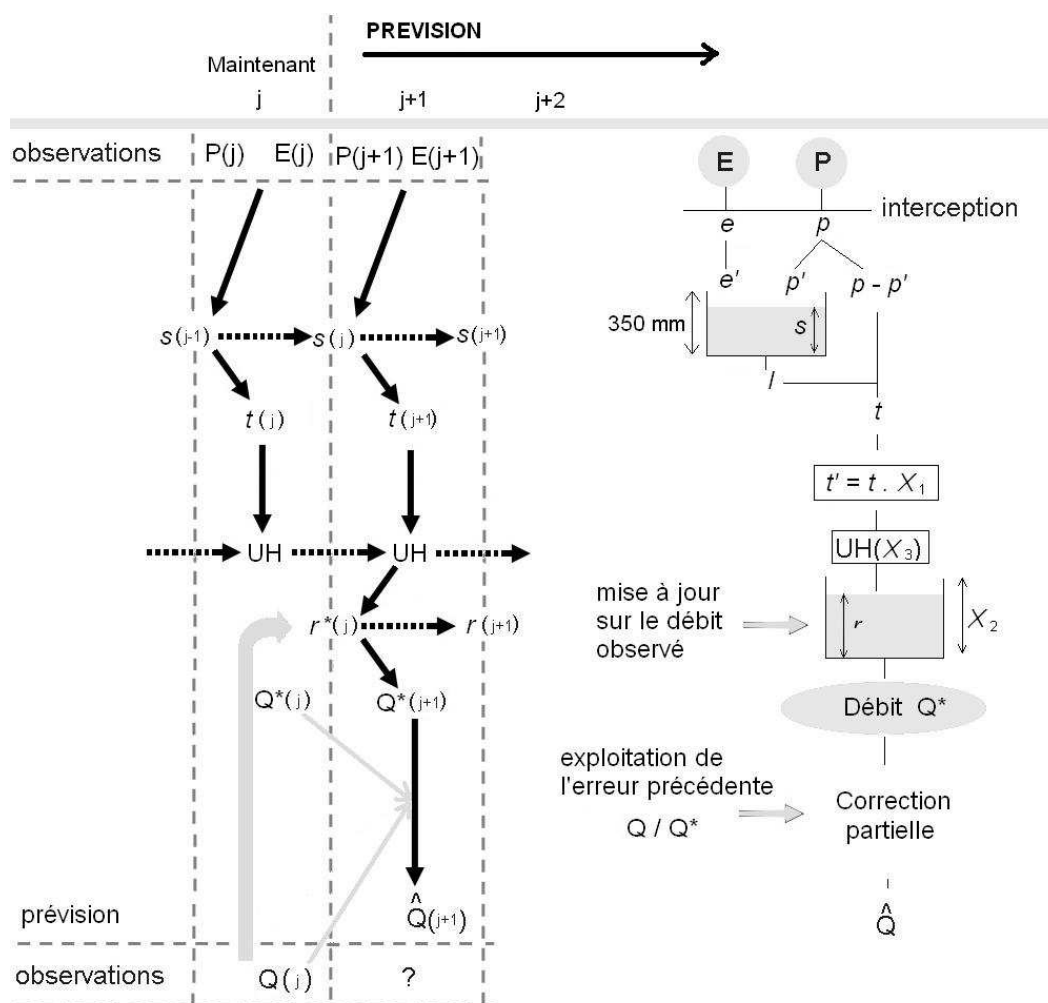


Figure 5 : Schéma structurel du modèle GR3P avec représentation de la procédure d'exploitation des débits observés (E : évapotranspiration potentielle, P : pluie, X : paramètre)

La phase d'exploitation de l'information apportée par le dernier débit observé (assimilation de données) se déroule en deux temps :

- une mise à jour directe du réservoir de routage du modèle, dont le niveau est recalculé de manière à pouvoir générer un débit exactement égal au dernier débit observé ;
- une exploitation de l'erreur de prévision faite par le modèle au pas de temps précédent. On considère en fait que l'erreur du modèle présente une certaine autocorrélation et on cherche alors, par un procédé simple, à exploiter cette propriété.

Dans cette étude, le modèle GR3P est utilisé au pas de temps horaire (Berthet et al., 2009). La pluie horaire moyenne de bassin jusqu'à l'horizon de prévision est nécessaire, calculée par moyenne arithmétique ou pondérée sur la base des postes pluviométriques disponibles en temps réel sur le bassin. La deuxième entrée nécessaire est le dernier débit observé à la station hydrométrique cible. Des données d'évapotranspiration potentielle sont également utilisées pour le réservoir de production.

Ce modèle de prévision requiert l'utilisation de scénarii de pluies futures. En effet, les prévisions quantitatives de pluie influencent notablement les prévisions hydrologiques, dès lors qu'on se situe au-delà du temps de réaction du bassin. Nous ne disposons pas pour cette étude d'historiques de prévisions de pluies exploitables. Le choix s'est donc porté sur deux types de scénarios simples de pluies futures, à savoir :

- des pluies futures nulles, scénario défavorable en particulier pour les horizons de prévisions supérieurs au temps de réaction des bassins. Un tel scénario est cependant utilisé par défaut en conditions opérationnelles en absence de prévision de pluie ;
- des pluies futures dites "parfaites" (égales aux observations de pluies à venir), scénario de test idéal. Un tel scénario est irréaliste en conditions opérationnelles, mais permet d'avoir une évaluation des performances du modèle sans les bruyers par la méconnaissance des pluies futures.

3.2. Différentes configurations de modélisation

L'objectif de l'étude est d'évaluer l'éventuelle plus-value induite par l'utilisation d'un modèle semi-distribué, et d'améliorer les prévisions dans le cas des confluences. Pour ce faire, la méthode adoptée comporte deux étapes. La première consiste à utiliser des confluences idéalisées, c'est-à-dire sans bassin intermédiaire, et de comparer les performances obtenues avec les versions globale, semi-globale et semi-distribuée du modèle pluie-débit. La seconde consiste à étendre cette comparaison à des cas de confluences réelles, ce qui requiert une prise en compte explicite de la partie intermédiaire.

La configuration globale du modèle GR3P (Fig. 6, configuration 1) sert donc de base de comparaison pour estimer les éventuelles améliorations apportées par la prévision aux confluences. Les prévisions sur le bassin s'effectuent avec une pluie de bassin (globale) et un jeu de paramètres unique pour le bassin. Les configurations suivantes sont ensuite utilisées pour évaluer l'intérêt de différents degrés de distribution du modèle :

- l'option semi-globale (Fig. 6, configuration 2) : un jeu de paramètre unique est utilisé pour les deux sous-bassins mais chacun d'entre eux reçoit une pluie de bassin qui lui est propre ;
- l'option semi-distribuée (Fig. 6, configuration 3a) : chacun des deux sous-bassins possède son propre jeu de paramètres calés, et reçoit une pluie de bassin qui lui est propre. C'est le niveau maximal de spatialisation possible en utilisant deux sous-bassins pour modéliser le comportement d'un bassin en aval d'une confluence ;
- la configuration traitant chacun des sous-bassins avec une pluie moyenne globale (Fig. 6, case vide) n'a pas été testée car elle conduit à des résultats médiocres pour les bassins de type chimère (Andréassian *et al.*, 2004) ;

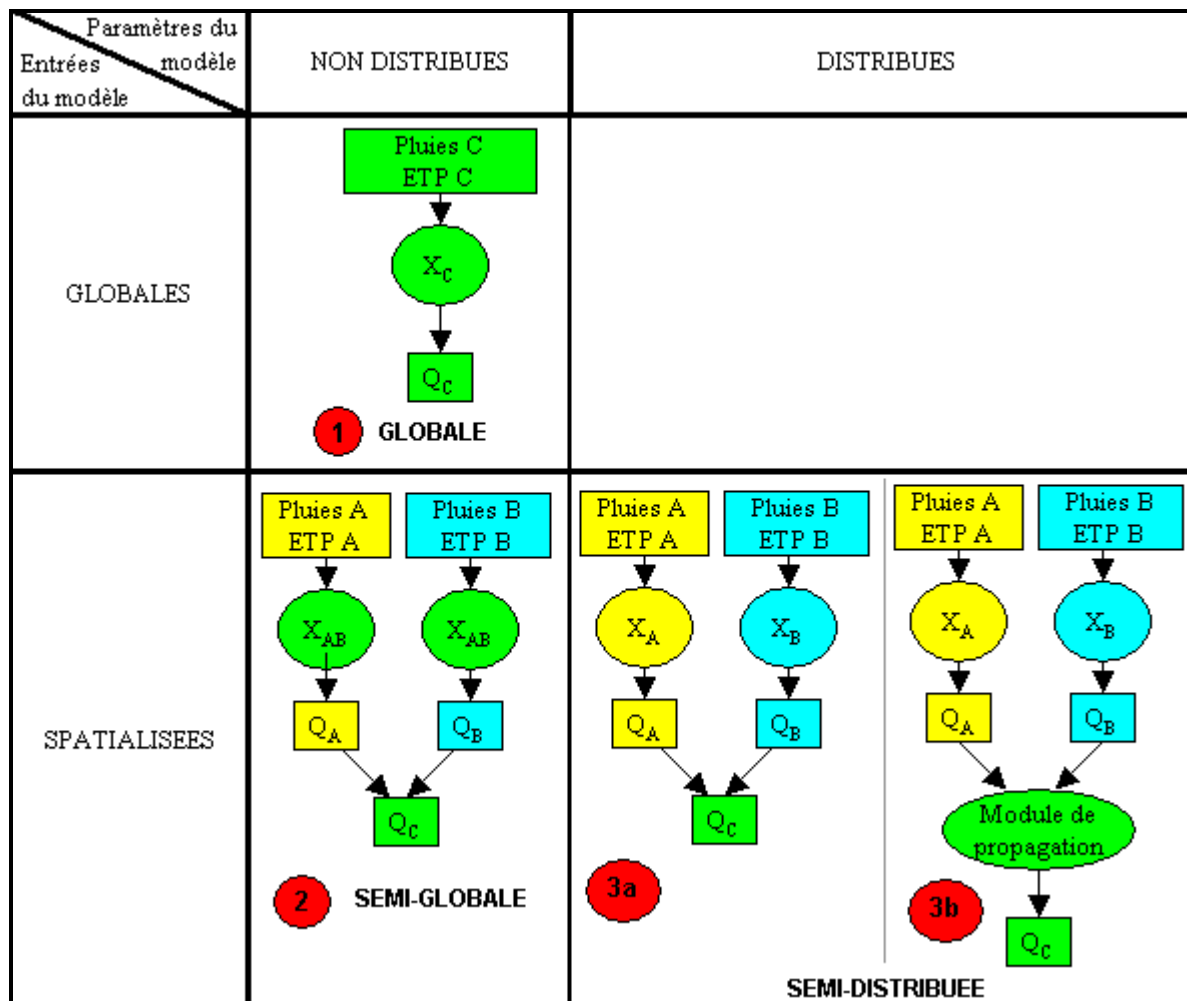


Figure 6 : Tableau des différentes configurations du modèle testées dans cette étude

- l'option semi-distribuée avec routage (Fig. 6, configuration 3b) : Cette configuration du modèle s'appuie sur la version semi-distribuée que l'on vient d'évoquer. Dans le cas de confluences réelles, il existe un bassin intermédiaire entre les stations amont et la station en aval de la confluence. Un module de propagation est donc ajouté, censé limiter les erreurs dues à la non-prise en compte du bassin intermédiaire modélisant le comportement de celui-ci. L'idée n'est pas ici de développer une modélisation complète de la zone intermédiaire, comme cela a été fait par Lerat (2009) mais d'utiliser un modèle simple, pour rester dans une démarche d'amélioration à moindre coût, et avec une complexification des modèles pas à pas. Le choix s'est donc porté sur un module comportant :
 - un hydrogramme unitaire appliqué à partir de chacune des deux stations amont. La distance hydraulique entre chacune des stations amont et la station post-confluence a été calculée (à partir d'un réseau de drainage créé via un Modèle Numérique de Terrain [de mailles carrées de 75 mètres de coté] sur la base des travaux de Lemoine (2008) et Lerat (2009)). Le décalage temporel est obtenu via cette distance hydraulique et une estimation de la célérité hydraulique propre à la confluence. Chaque temps de propagation est ainsi fonction de la distance hydraulique entre la station amont et la station aval et de cette célérité, qui constitue le premier paramètre à caler ;
 - un coefficient d'ajustement des volumes. Après addition des débits amont, ceux-ci sont multipliés par un second paramètre calé, de manière à prendre en compte les éventuels apports ou pertes au niveau du bassin intermédiaire.

Le modèle GR3P est donc calé pour ces différentes configurations, sur deux échantillons de bassins (les confluences idéalisées, sans bassin intermédiaire, et les confluences réelles) qui seront détaillés par la suite, et en appliquant la procédure explicitée ci-dessous.

3.3. Calage / contrôle du modèle

Un algorithme local de recherche d'optimum dit algorithme "pas à pas", développé par le Cemagref, a été utilisé¹. Sa robustesse a été démontrée lorsque le nombre de paramètres à caler reste réduit (inférieur à dix), ce qui est le cas dans cette étude (3 paramètres à caler simultanément au maximum). Le critère mathématique choisi comme fonction objectif est le C2MP (transformation de la persistance), explicité dans le prochain paragraphe.

En temps réel, le modèle de prévision est par définition utilisé sur des événements non encore rencontrés. Il faut donc mettre le modèle dans une situation d'évaluation similaire. C'est pour cette raison qu'une procédure de calages / contrôles croisés sur sous-périodes indépendantes (Klemeš, 1986) a été utilisée (Fig. 7). Elle consiste à caler alternativement le modèle sur l'une des sous-périodes et à le contrôler sur l'autre, l'évaluant donc en contrôle sur l'intégralité des données disponibles. Les performances sont analysées en phase de contrôle, puisque c'est dans cette phase (en dehors du calage) que fonctionnent les modèles en opérationnel, les performances en calage donnant en effet une vision trop optimiste des capacités des modèles. La chronique est divisée en deux parties de même longueur. La première année de chaque période est réservée pour l'initialisation du modèle (mise en route), ce qui signifie que les critères de performances n'intègrent pas les erreurs commises au cours de celle-ci. Le résultat final est la moyenne des deux critères de qualité obtenus en contrôle.

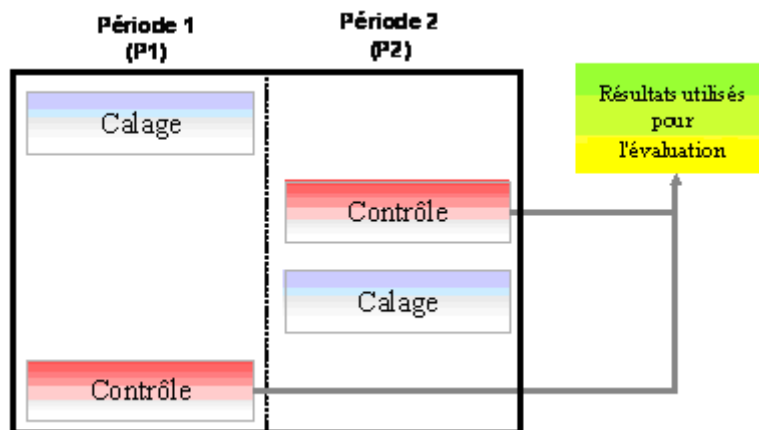


Figure 7 : Schéma de principe de la procédure de calage / contrôle utilisée

3.4. Critère de qualité des prévisions, le C2MP

L'objectif du calage du modèle est qu'il soit le plus robuste et le plus fiable possible. Cependant, ces notions de qualité sont délicates à quantifier avec un critère mathématique. Pour l'évaluation comparative des modèles sur un grand échantillon de bassins versants, seuls les critères quantitatifs permettent une évaluation et une comparaison raisonnablement objective des performances des modèles. Le critère doit donc être relatif, c'est-à-dire non absolu, car il est très subjectif de dire qu'un modèle est bon ou mauvais de manière absolue. En revanche, dans un cadre comparatif, il est beaucoup plus objectif de dire qu'un modèle est meilleur ou moins bon qu'un autre modèle.

¹ Descriptif du fonctionnement de cet algorithme de calage disponible dans la thèse de Mathevet (2005)

Le critère de [Nash-Sutcliffe \(1970\)](#) classiquement utilisé pour l'évaluation des modèles de simulation n'est pas adapté aux modèles de prévision. En effet, il compare l'erreur commise par le modèle à celle d'un modèle naïf donnant comme débit prévu le débit moyen sur la période. Étant beaucoup trop grossier dans un exercice de prévision, il conduit à des valeurs du critère de Nash-Sutcliffe très proches de 1, donnant une vision trop optimiste des performances du modèle de prévision. Il convient donc de choisir un modèle de référence plus pertinent pour l'exercice de prévision. En prévision à un horizon h , on prend comme modèle de référence un modèle naïf où le débit prévu est égal au débit observé à l'instant de prévision (modèle de persistance). On a alors :

$$Q_{\text{prevu}}(t+h) = Q_{\text{obs}}(t) \quad \text{Équation 3-1}$$

avec $Q_{\text{obs}}(t)$ et $Q_{\text{prevu}}(t)$ débits respectivement observé et prévu à l'instant t , et h le délai de prévision.

En prévision à l'horizon h , [Kitadinis et Bras \(1980\)](#) proposent d'évaluer l'efficacité de la prévision du modèle avec le coefficient de persistance défini par :

$$\text{Pers} = 1 - \frac{\sum (Q_{\text{prevu}}(t+h) - Q_{\text{obs}}(t+h))^2}{\sum_t (Q_{\text{obs}}(t) - Q_{\text{obs}}(t+h))^2} \quad \text{Équation 3-2}$$

avec n le nombre de pas de temps.

Ce critère vaut 1 pour un modèle parfait, 0 pour un modèle commettant des erreurs équivalentes à celles du modèle de persistance. Il prend des valeurs négatives quand le modèle est moins bon que celui de persistance. Lorsque l'on s'intéresse aux performances moyennes sur un large échantillon de bassins, d'éventuelles valeurs fortement négatives du critère peuvent venir biaiser la moyenne. Il est donc intéressant de se reporter à un critère borné, variant dans $]-1 ; 1]$ (plutôt que dans $]-\infty ; 1]$) comme cela a été proposé par [Mathevet et al. \(2006\)](#). Le critère noté C2MP sera donc utilisé, transformée du critère de persistance, et donné par :

$$\text{C2MP} = \frac{\text{Pers}}{2 - \text{Pers}} \quad \text{Équation 3-3}$$

La valeur moyenne des C2MP sur tous les bassins traités sera utilisée comme critère final d'appréciation des performances des différentes options de modélisation testées. D'autres critères ont par ailleurs été utilisés pour contrôler les résultats mais ne seront pas exposés ici par soucis de concision. Ils apportent une information supplémentaire indispensable, car un modèle de prévision de bonne qualité peut présenter un C2MP très faible pour l'horizon "h = 1 heure", le modèle naïf étant parfois très bon pour cet horizon de prévision dans les bassins à réaction lente.

3.5. Horizons de prévision

Le panel de bassins versants utilisés est très hétérogène. Le plus gros bassin (61820 km²) étudié par la suite est celui de la Seine à Poissy, le plus petit celui de la Cadanne à Pondaurat, drainant une superficie de 10 km². Les temps de réactions de ces différents bassins sont de fait tout aussi hétérogènes. Les horizons au delà de 12 heures n'ont pas d'intérêt à être étudié pour les plus petits. A l'inverse, étudier les prévisions de crue à une heure pour le bassin de la Seine à Poissy n'a pas non plus d'intérêt. Cette hétérogénéité forte impose donc de traiter une large gamme d'horizons de prévisions. Par conséquent, dix horizons de prévisions sont testés : 1, 3, 6, 9, 12, 18, 24, 36, 48 et 72 heures

4. Echantillons d'application

4.1. Echantillon 1 : les confluences sans bassin intermédiaire

4.1.1. Sous-échantillon 1 : les bassins chimères

Une confluence chimère est un bassin, virtuel créé à partir de la combinaison de deux sous-bassins existant réellement mais n'ayant aucun lien géographique (Fig. 8). Le débit de la confluence chimère est la somme directe des deux débits amont.

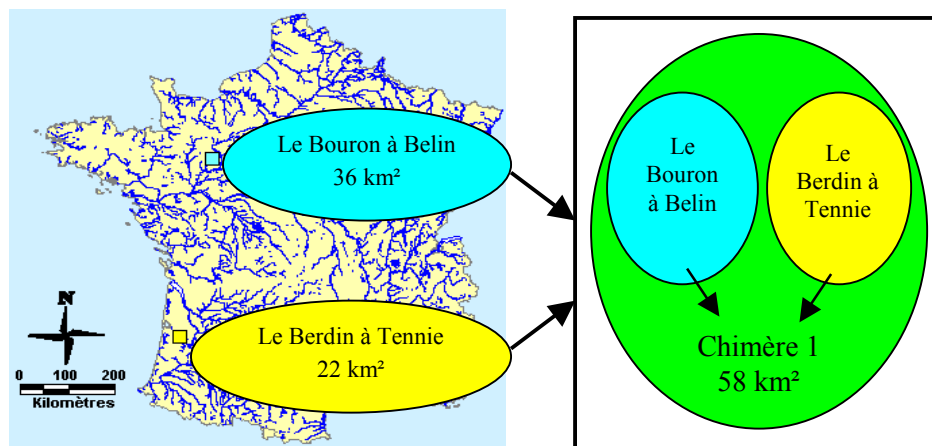


Figure 8 : Confluence chimère

Pour mener à bien cette étude, une base de données de 1070 bassins versants français a été utilisée (Fig. 9). Elle a été créée par [Le Moine \(2006\)](#) et est composée de bassins français disposant de données au pas de temps horaire sur la période 1995-2005 et drainant une superficie comprise entre 10 et 10000 km². Les stations influencées "hydrologiquement" (barrages etc..) n'ont pas été conservées. La méthodologie adoptée pour la création de cette base est détaillée dans la thèse de [Le Moine \(2008\)](#). En ce qui concerne les données utilisées :

- les débits sont issus de la banque Hydro ;
- les pluies, fournies par Météo France, sont issues du système SAFRAN (Système d'Analyse Fournissant des Renseignements Atmosphériques pour la Nivologie) puis désagrégées en se basant sur un réseau de 580 pluviographes au sol, la méthode des polygones de Thiessen ayant été utilisée pour définir leurs zones d'influence sur les mailles SAFRAN (mailles de 8 kilomètres de côté) ;
- les évapotranspirations potentielles (ETP) ont été calculées via la formulation d'[Oudin \(2004\)](#), à partir des chroniques de température également issues de l'archive SAFRAN. [Oudin \(2004\)](#) a montré que cette formulation à partir des données de température conduit à de meilleurs résultats en terme de simulation des débits qu'une série d'ETP de Penman-Monteith non datées.

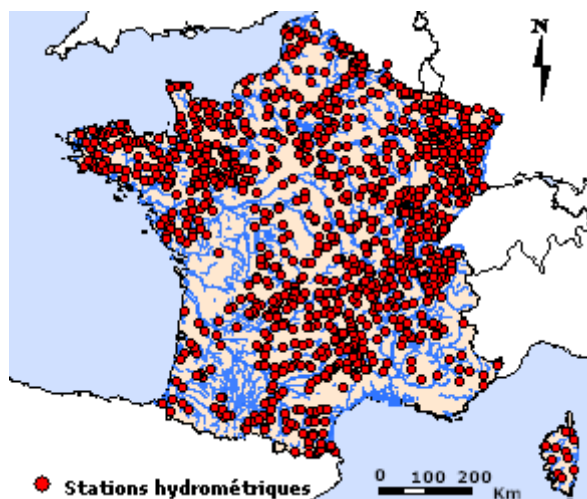


Figure 9 : Localisation des 1070 stations hydrométriques disponibles dans la base de données du Cemagref ([Le Moine, 2008](#))

A partir de ces 1070 bassins, la méthodologie adoptée pour créer les chimères a été relativement simple. La contrainte principale est le temps de calcul nécessaire à faire tourner toutes les configurations de modèles sur les échantillons de bassin. Potentiellement, en disposant de 1070 bassins, on peut obtenir $\frac{1}{2} \cdot [1070 \times 1069] = 571915$ bassins chimères. Il n'a donc pas été envisageable pour les différents scénarios (4 configurations à tester, sur plus de trois horizons de prévisions etc.) de caler le modèle GR3P sur ce demi-million de chimères.

La méthode choisie a donc consisté à créer 535 chimères, paires de deux sous-bassins extraites via un tirage aléatoire sans remise dans le panel des 1070 bassins. Cette manière de procéder permet d'utiliser un nombre maximal de bassins tout en évitant d'influencer la création des chimères avec des critères subjectifs. Les confluences constituées de deux sous-bassins situés dans la même zone hydrographique (première lettre du code Hydro de la station exutoire) ont été supprimées de l'échantillon (de manière à forcer le caractère hétérogène des chimères), pour finalement obtenir un échantillon de 494 chimères.

4.1.2. Sous-échantillon 2 : les confluences idéalisées

Par ailleurs, 41 confluences réelles (explicitées dans la prochaine section) sont étudiées en configuration idéalisée. Dans la pratique, cela revient à considérer que les deux stations amont se trouvent juste avant la confluence, et que le bassin intermédiaire est nul. La somme des deux débits amont est utilisée comme station de contrôle du bassin versant, et non pas la station aval. Cela constitue donc un sous-échantillon de confluences sans bassin intermédiaire, précieux pour observer les différences de comportement entre les chimères issues du regroupement de deux bassins éloignés géographiquement et sur des confluences issues de sous-bassins côte à côte dans la réalité.

4.2. Echantillon 2 : les confluences réelles

Si l'échantillon 1 correspond à des cas "théoriques" de confluences, il est intéressant de tester l'apport de la spatialisation des modèles dans un contexte appliqué, opérationnel, sur des bassins versants cette fois-ci réels, c'est-à-dire comprenant un bassin intermédiaire. En effet, dans la réalité, les stations hydrométriques ne sont pas placées juste en amont et aval de la confluence. Un linéaire de rivière non négligeable qui draine tout un bassin intermédiaire existe donc toujours (Fig. 10).

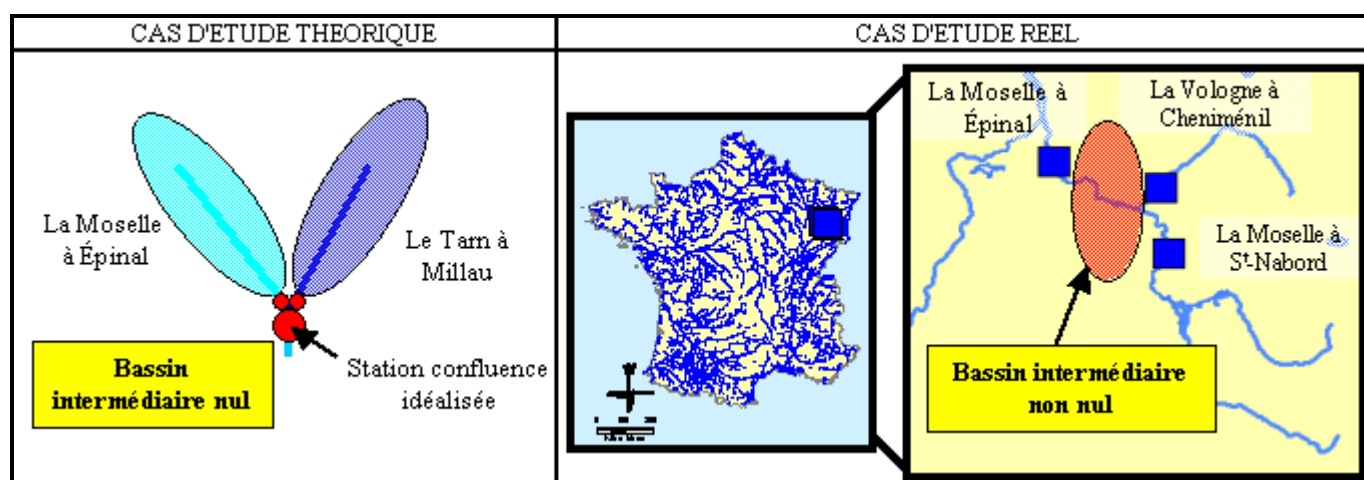


Figure 10 : Cas d'étude opérationnel par rapport au cas d'étude théorique

Dans leurs travaux de comparaison entre l'approche globale et distribuée en fonction de l'hétérogénéité des bassins versants, *Baudez et al. (1999)* ont sélectionné des bassins tels que la superficie du bassin intermédiaire ne dépasse pas 20% de la superficie totale du bassin étudié. Ainsi, seulement quinze bassins ont été utilisés à l'époque. Ceci met bien en lumière la difficulté de trouver de telles configurations de mesure avec le réseau hydrométrique existant.

4.2.1. Sélection des confluences

Pour ce qui est des confluences issues de la base de bassins du Cemagref, une analyse cartographique de la localisation des 1070 bassins a été effectuée pour identifier tous les cas de confluences réelles. Les configurations présentant une station en aval précédée de deux stations amont ont été au préalable toutes retenues.

Un travail complémentaire d'analyse du territoire et des données du SPC SMYL a permis de retenir un échantillon supplémentaire de confluences. Les stations hydrométriques (et par conséquent les confluences impliquées) présentant une chronique de données trop restreinte¹ (inférieure à 6 années de données) ont été écartées. De plus, les stations influencées ont été retirées de l'échantillon, et par la même occasion toute confluence dans laquelle cette station intervenait. Lorsqu'un cours d'eau est trop influencé (barrages de navigation ou autres barrages plus importants [écrêtement de crue, soutien d'étiage]), il n'est pas possible de mettre en place un modèle pluie-débit cohérent sachant que la réponse en débit à l'exutoire est majoritairement dépendante de la gestion du barrage, et dans une moindre mesure du régime des pluies. Un exemple concret est illustré ci-dessous (Fig. 11), où l'on peut voir que la réaction brutale en débit n'est pas issue d'un phénomène météorologique, les stations en amont de la station analysée ne présentant aucune réaction.

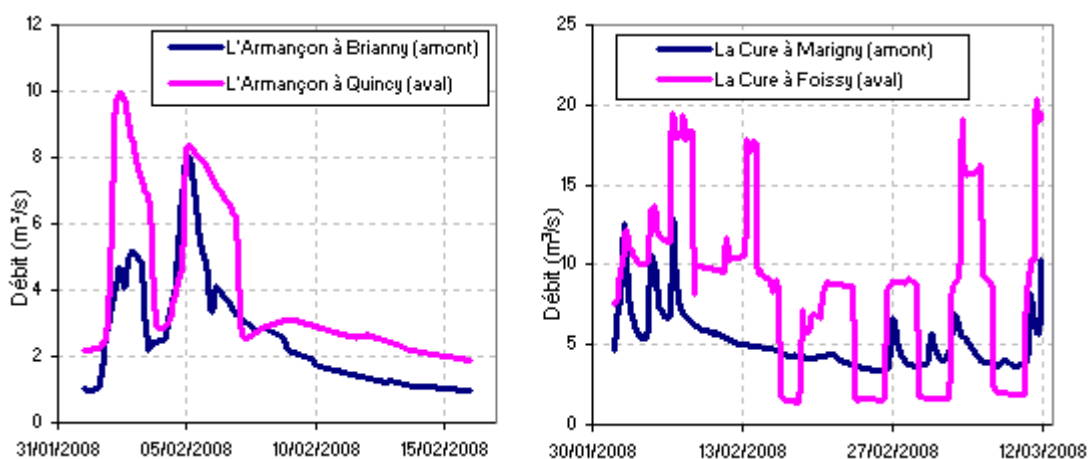


Figure 11 : Exemples d'hydrogrammes influencés (stations supprimées de l'échantillon)

Deux filtres ont ensuite été appliqués à cet échantillon "brut" de confluences issues de la base de bassins du Cemagref et de la zone d'action du SPC SMYL :

- l'échantillon a été réduit à des confluences pour lesquelles le bassin intermédiaire est le plus réduit possible, pour ne pas avoir à développer de modélisation complète du bassin intermédiaire comme l'a fait par exemple [Lerat \(2009\)](#). Ainsi, les confluences dont le bassin intermédiaire est supérieur en superficie à 50% du bassin global ont été écartées ;
- le second critère de sélection a porté sur la différence de taille entre les deux (ou trois) sous-bassins amont. Lorsqu'un des deux bassins amont est trop important par rapport au second, l'influence du sous-bassin le plus petit n'est même pas visible sur les hydrogrammes au niveau de la station aval. Cela correspond souvent à des cas où la taille du bassin intermédiaire est supérieure à celle du second sous-bassin. Le cas extrême qui illustre bien ce fait est la confluence entre la Seine et la Mauldre en amont de Poissy, le débit moyen au niveau de la Seine étant plus de 150 fois plus important que celui de la Mauldre. Dans un tel cas, le débit à l'aval sera très peu sensible aux variations de débit du petit bassin amont et il paraît inutile de mettre en place un modèle semi-distribué avec une telle configuration. Nous avons fixé la limite à un facteur 5 entre les superficies deux sous-bassins.

¹ Les débits utilisés pour les stations du SPC SMYL sont tirés de la banque Hydro au pas de temps horaire. Les données pluviométriques utilisées sont issues d'un réseau de pluviomètres au sol et mises à disposition par différents réseaux, Météo France, DIREN etc. Enfin, les données d'évapotranspiration utilisées sont calculées de la même manière que pour la base des 1070 bassins.

Finalement, l'échantillon de confluences réelles obtenu se compose de 41 éléments, dont 11 situées sur le territoire d'action du SPC SMYL, et 5 étant des confluences triples (trois sous-bassins constitutifs). La figure 12 illustre leur localisation.

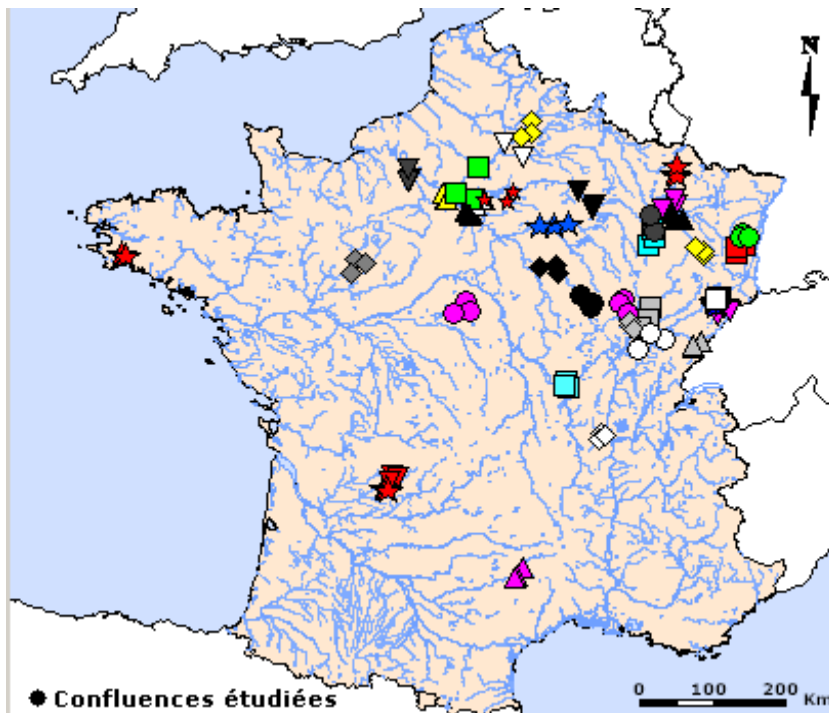


Figure 12 : Localisation des confluences réelles étudiées

4.2.2. Analyse de l'influence du bassin intermédiaire

La figure 13 illustre la contribution du bassin intermédiaire au bassin total étudié pour les 41 confluences réelles. Elle confirme le fait que le bassin intermédiaire demeure non-négligeable dans la quasi-majorité des cas. En moyenne, le bassin intermédiaire contribue à hauteur de 10% en termes de débit moyen.

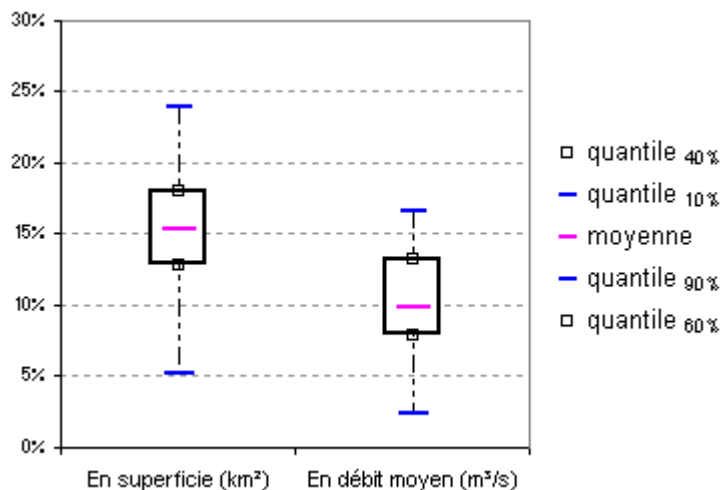


Figure 13 : Contribution du bassin intermédiaire dans la confluence étudiée

Un travail s'impose à ce stade. L'objectif est d'évaluer si le débit aval peut être reconstitué à moindre coût sur la base des débits amont observés. Il s'agit de voir si une connaissance hypothétique parfaite des deux débits amont permet de déterminer de manière convenable le débit à la station de contrôle de la confluence. Pour résumer, la question est de savoir s'il est possible de trouver des coefficients α et L_i tels que :

$$Q_{\text{aval}}(t) = \alpha \cdot [Q_{\text{amont}_1}(t-L_1) + Q_{\text{amont}_2}(t-L_2)] \quad \text{Équation 4-1}$$

Dans un premier temps, il est possible d'utiliser la modélisation la plus simple possible, en prenant L_1 et L_2 nuls et α égal à 1. Cela revient à quantifier l'erreur commise utilisant :

$$Q_{\text{aval}}(t) = Q_{\text{amont}_1}(t) + Q_{\text{amont}_2}(t) \quad \text{Équation 4-2}$$

Les résultats obtenus sont présentés dans le [tableau 3](#). On est ici dans un contexte de simulation, le critère classiquement utilisé est donc celui de [Nash-Sutcliffe \(1970\)](#). Certains bassins présentent de très bons résultats avec cette approximation, alors que le comportement de certains autres ne peut manifestement pas être retranscrit de manière correcte avec cette approximation.

Moyenne	0.887	Confluence	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Ecart-type	0.111	Nash	0.96	0.94	0.90	0.97	0.97	0.94	0.84	0.86	0.64	0.75	0.82	0.96	0.95	0.77	0.98	0.83	0.83	0.97
Min	0.461	Confluence	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Max	0.980	Nash	0.67	0.92	0.95	0.96	0.95	0.98	0.90	0.94	0.90	0.46	0.91	0.97	0.93	0.88	-0.27	0.90	0.93	0.90

Tableau 3 : Critère de Nash obtenu en modélisant le débit aval par " $Q_{\text{amont}_1} + Q_{\text{amont}_2}$ " – Confluences à deux sous-bassins

Il est d'ailleurs très intéressant de confronter ces performances à l'importance du bassin versant intermédiaire par rapport au bassin total ([Fig. 14](#)). Il n'est pas possible, au vu des résultats, de conclure de manière claire quant à une hypothétique corrélation entre ces performances et la proportion du bassin intermédiaire.

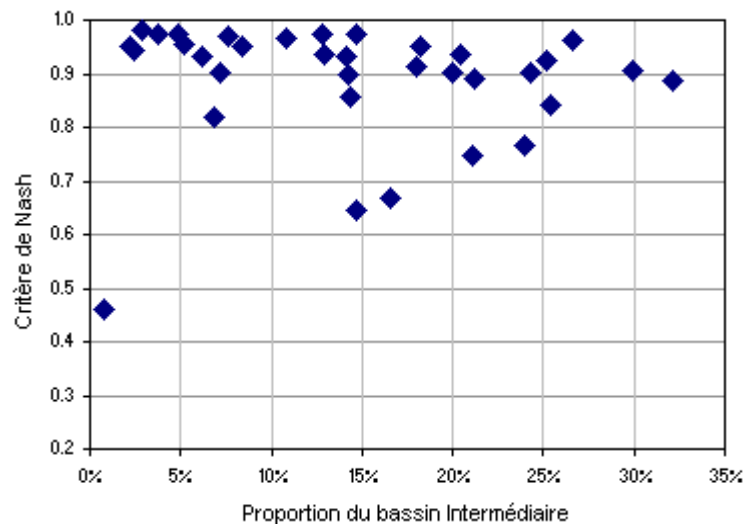


Figure 14 : Qualité de la simulation du débit en fonction de l'importance du bassin intermédiaire

Pour contrôler les délais de propagations calés (supposés optimaux), les "cartes" de valeur de Nash en fonction du décalage ont été tracées. Dans certains cas ([Fig. 15](#)), un optimum clair de solution est observé. Cette configuration donne du poids à la valeur obtenue et donc un certain niveau de confiance. En revanche, pour certains bassins, cette "carte" donne des résultats ([Fig. 16](#)) qui amènent à conclure que, bien qu'on puisse toujours trouver une valeur de Nash maximale, elle ne correspond pas forcément à une réalité physique ayant pour correspondance un délai de propagation entre amont et aval, de nombreuses solutions étant équivalentes.

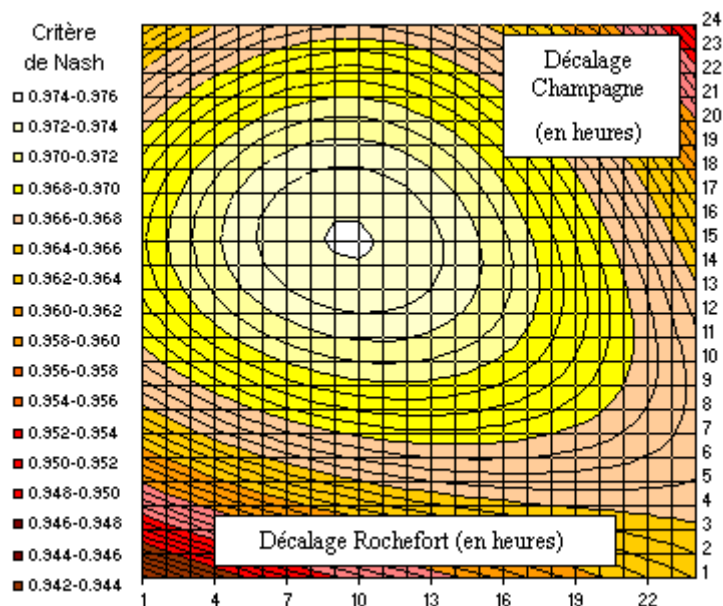


Figure 15 : Qualité de la simulation à Neublans à partir des débits à Rochefort et Champagne (confluence Loue - Doubs)

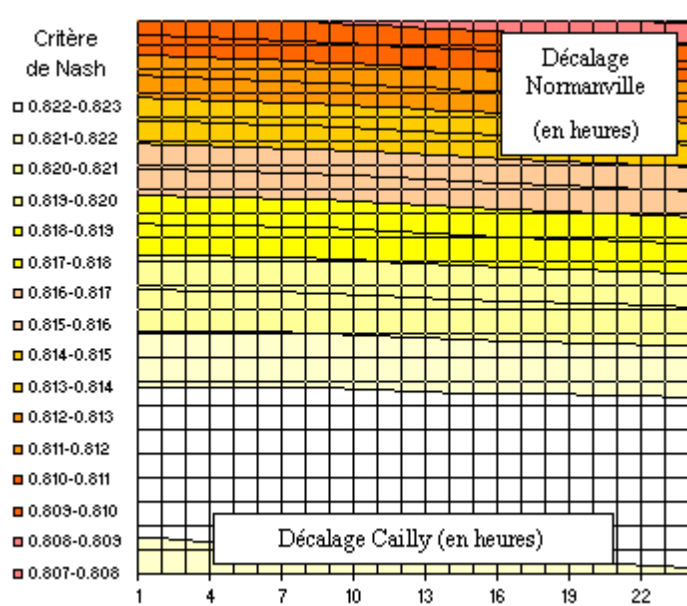


Figure 16 : Qualité de la simulation à Louviers à partir des débits à Cailly et Normanville (confluence Iton - Eure)

La figure 17 présente les résultats obtenus en calant les trois paramètres α , L_1 et L_2 , en bornant L_1 et L_2 à 72 heures. Les niveaux de performance obtenus ici en simulation de manière optimale pour les différentes confluences sont satisfaisants pour la grande majorité des bassins, mais ne préjugent cependant pas des performances en prévision. Cette analyse simple des liens de dépendance amont-aval permettra cependant d'avoir un regard critique sur les résultats des modélisations réalisées sur ces bassins. D'une part, elle permet de disposer d'indicateurs pouvant aider à caractériser le bassin intermédiaire. D'autre part, elle montre qu'il est possible de reconstituer l'hydrogramme à l'aval de la confluence à partir des hydrogrammes des deux sous-bassins amont par le biais de transformations simples du type de celles mises en place dans le module de routage. Enfin, ce travail sur les données a permis de mettre en lumière la qualité douteuse des données de débit à Poissy (aval de la confluence Oise-Seine) au pas de temps horaire (confluence grisée dans le tableau 3). Une analyse des données hydrométriques brutes (hauteurs, courbes de tarage, jaugeages) a permis de confirmer ces soupçons.

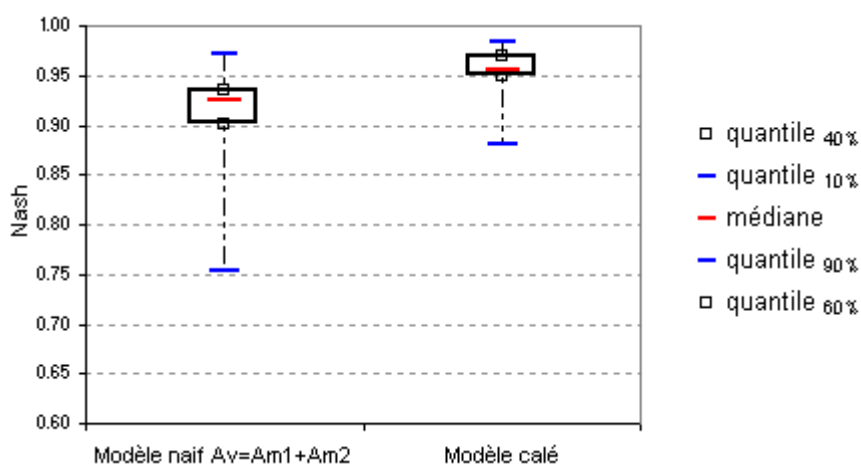


Figure 17 : Comparaison de la qualité (critère de Nash) de la simulation du débit à la station aval à partir des débits observés aux deux stations amont

5. Résultats sur les confluences sans bassin intermédiaire

5.1. Résultats sur les bassins chimères

La méthodologie détaillée en partie 3 a été appliquée aux deux échantillons de bassins versants que l'on vient de décrire. Les graphes (Fig. 18 et 19) présentent les performances du GR3P en configuration globale, semi-distribuée et semi-globale pour les confluences chimères.

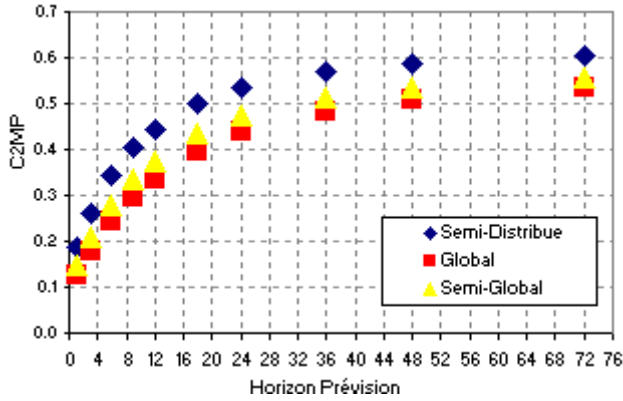


Figure 18 : Performance des différentes configurations en fonction de l'horizon de prévision (calage à 12h)

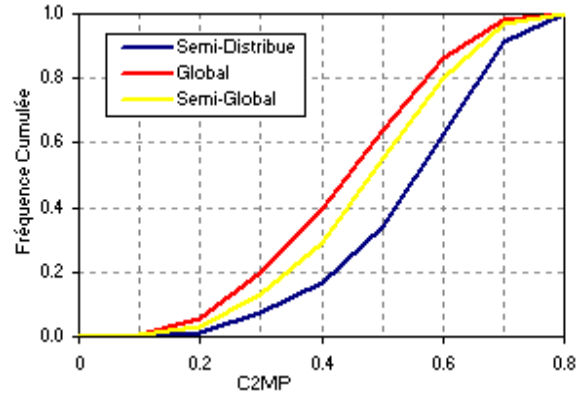


Figure 19 : Distribution cumulée des résultats pour les différentes configurations (prévision à 12h)

La tendance ne varie pas selon l'horizon de prévision auquel on se place : le modèle GR3P en configuration semi-distribuée surpasse les deux autres configurations, globale et semi-globale. De plus, la configuration semi-globale du modèle est plus proche en terme de performance du modèle global que de la version semi-distribuée. Par conséquent, la spatialisation des pluies ne semble pas être l'unique explication de l'amélioration des résultats, la modélisation du comportement des deux sous-bassins ayant également un impact fort sur cette amélioration dans la configuration semi-distribuée.

L'objectif est ensuite de comparer plus finement les performances des versions globale et semi-distribuée du modèle GR3P. A ce stade, il convient donc de définir un critère permettant de quantifier cette variation de performance. L'idée retenue est de se baser sur un critère lié au rapport des erreurs quadratiques commises dans chacune des deux configurations. Pour faciliter la compréhension des graphes, on retranchera systématiquement 1 à ce rapport. A partir des équations détaillées dans la partie consacrée à la méthodologie, l'amélioration du modèle semi-distribué par rapport au global ($AM_{SD/GL}$) peut s'écrire :

$$AM_{SD/GL} = \frac{MSE_{GL}}{MSE_{SD}} - 1 = \frac{\sum_t (Q_{obs}(t+h) - Q_{prev}(t+h))^2_{GLOBAL}}{\sum_t (Q_{obs}(t+h) - Q_{prev}(t+h))^2_{SEMI-DIST}} - 1 \quad \text{Équation 5-1}$$

De fait, lorsque $AM_{SD/GL}$ est positif, le modèle semi-distribué surpasse le modèle global et inversement. Quand " $AM_{SD/GL} = 1$ ", l'erreur quadratique commise avec la configuration semi-distribuée du modèle GR3P est deux fois moins importante que celle commise avec la configuration globale du modèle.

En étudiant les performances comparées pour toutes les chimères (Fig. 20), un premier résultat majeur apparaît : les chimères pour lesquelles une diminution de performance avec la version semi-distribuée est observée sont très marginales (moins de 5% de l'échantillon). En fixant un seuil d'amélioration (relative) significative à 0.2 (lorsque $AM_{SD/GL}$ est supérieur à 0.20 en valeur absolue, un des deux modèles commet une erreur quadratique réduite d'au moins 20%), une "amélioration significative" des performances avec la configuration semi-distribuée est observée dans plus de 40% des cas.

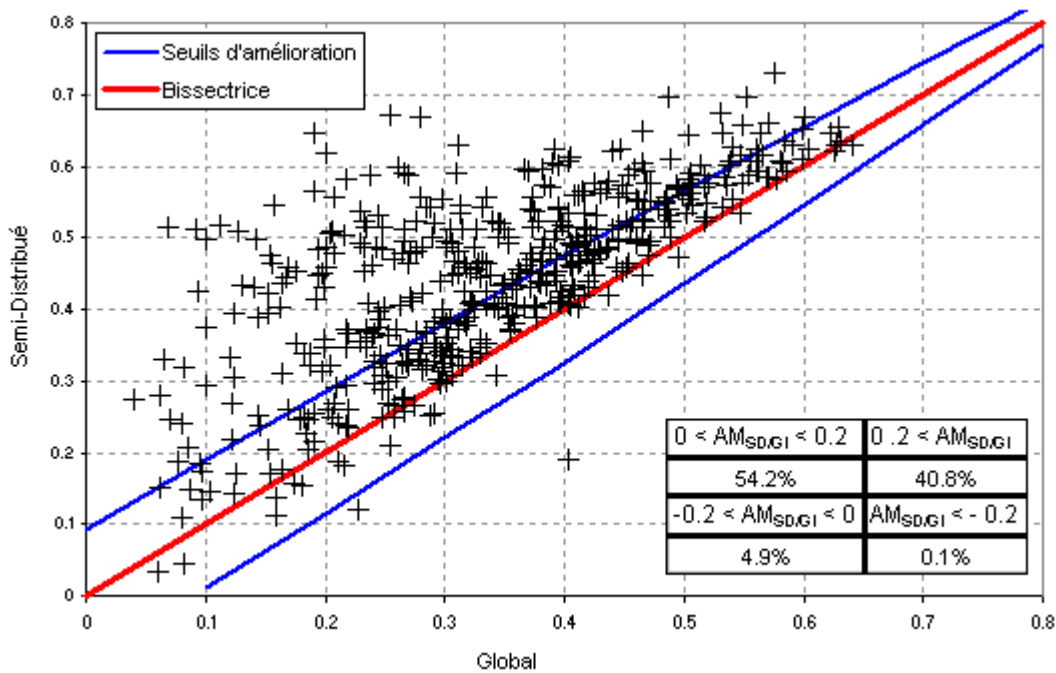


Figure 20 : Performances (en critère de persistance) du modèle GR3P en configuration semi-distribuée par rapport au même modèle en configuration globale.

Conclusion : dans le cas de bassins chimères (confluence sans bassin intermédiaire composée de deux sous-bassins déconnectés géographiquement), la qualité des prévisions est meilleure avec la version semi-distribuée du modèle GR (pluies et jeux de paramètres différents pour les deux sous-bassins).

5.2. Résultats sur les confluences idéalisées

En vue d'une application concrète de ces résultats, il convient évidemment d'évaluer si les mêmes tendances sont observées pour des confluences idéalisées (sans bassin intermédiaire) dont les deux sous-bassins composent une confluence existante. Pour ce sous-échantillon de confluences avec bassin intermédiaire nul, une tendance très marquée se dégage : les performances dans les trois configurations du modèle GR3P sont quasi-semblables (Fig. 21 et 22).

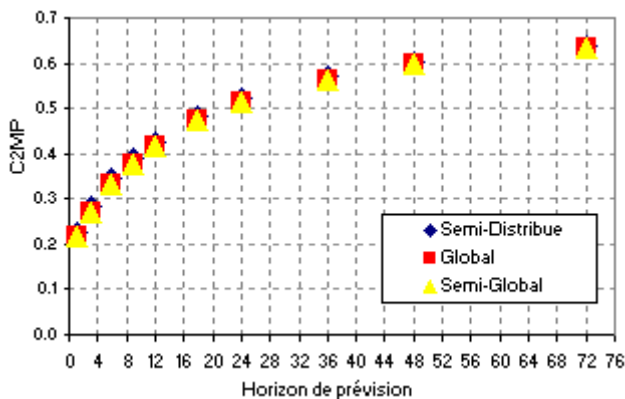


Figure 21 : Comparaison du critère de performance entre les différentes configurations en fonction de l'horizon de prévision

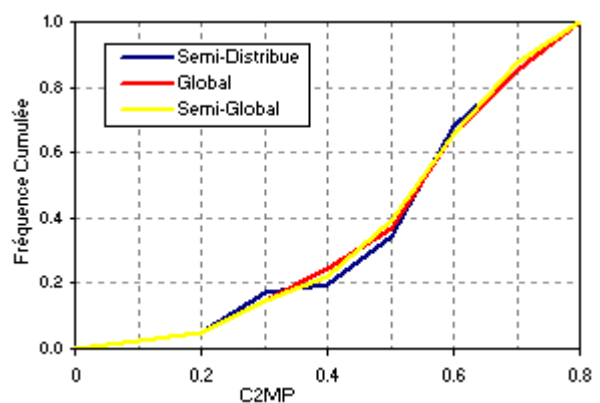


Figure 22 : Distribution cumulée des résultats entre les différentes configurations (prévision à 12h)

Sans pour autant que les performances soient dégradées avec la configuration semi-distribuée (22% des confluences idéalisées), les confluences dont les performances augmentent "significativement" avec la configuration semi-distribuée sont peu nombreuses (Fig. 23).

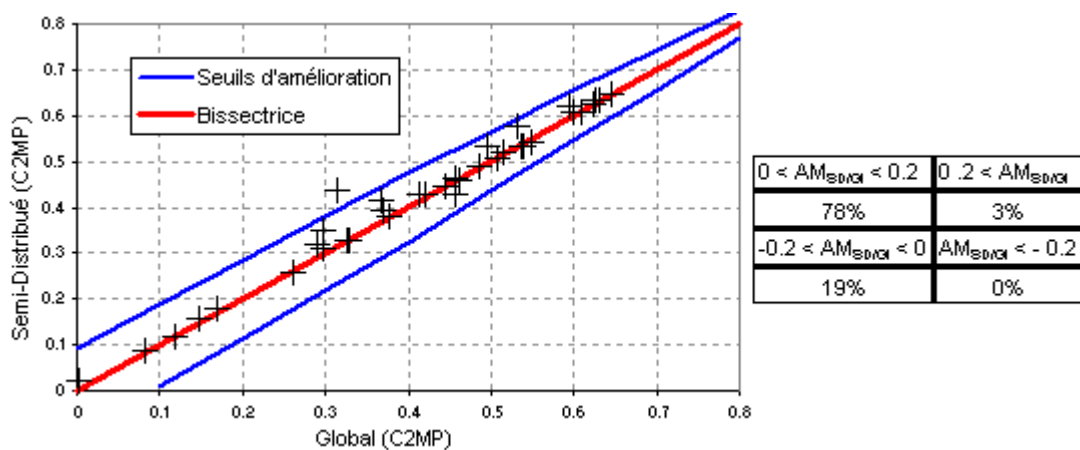


Figure 23 : Performances (en critère de persistance) du modèle GR3P en configuration semi-distribuée par rapport au même modèle en configuration globale – Délai de prévision : 12 heures.

Conclusion : L'amélioration de la qualité des prévisions avec le GR3P en version semi-distribuée est quasi-nulle dans le cas des confluences idéalisées. Le fait marquant est la différence de résultats observée entre les deux échantillons, bassins chimères et confluences réelles idéalisées. Il convient donc de comprendre et de tenter d'expliquer cet écart dans les résultats entre les deux sous-échantillons auxquels la même méthodologie a été appliquée.

5.3. Analyse et exploitation

Le seul paramètre qui change entre les deux sous-échantillons testés est l'origine des deux sous-bassins constitutifs de la confluence : sélectionnés de manière aléatoire dans une base de sous-bassins pour les chimères, et correspondant à des cas de confluences existantes pour le second sous-échantillon. L'objectif est donc d'analyser ce qui diffère fondamentalement entre ces bassins pour tenter d'associer une probabilité d'amélioration avec le modèle semi-distribué à une typologie de bassin. Il faut donc pour cela dégager des caractéristiques du bassin confluence et y confronter l'amélioration relative obtenue avec la configuration semi-distribuée du modèle.

5.3.1. Influence du rapport de taille des sous-bassins

Deux types de caractéristiques du bassin existent : des indices s'intéressant aux tailles (superficie, débit moyen etc.) des sous-bassins constitutifs, et des indices s'intéressant aux différences de comportement (régimes de pluie, temps de réaction, caractéristiques des sols, etc.). On s'intéresse ici dans un premier temps à la première catégorie citée. En classant (classes d'effectifs homogènes) les bassins en fonction du rapport de débit moyen des deux sous-bassins, ce rapport valant 1 pour des bassins de même taille, 0 pour des bassins de taille très différente, on observe que l'amélioration à attendre de l'utilisation de la version semi-distribuée augmente lorsque la similarité de taille des deux sous-bassins augmente (Fig. 24).

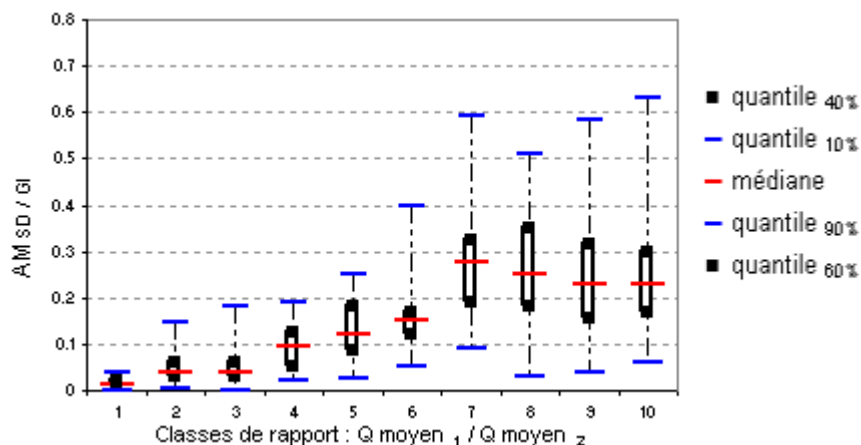


Figure 24 : Amélioration induite par l'utilisation du modèle semi-distribué en fonction des classes de rapport des débits moyens des deux sous-bassins constitutifs de la chimère (de 0 [bassins très différents] à 10 [bassins de même taille])

Ce résultat est interprétable. Comme évoqué dans la partie consacrée à la sélection des confluences, lorsqu'un des deux sous-bassins est très important par rapport à l'autre, le comportement de la confluence s'apparente beaucoup plus à celui de ce sous-bassin, rendant donc moins pertinente l'utilisation d'informations spatialisées, la modélisation du comportement du second sous-bassin n'ayant que peu d'influence dans le débit final à la confluence. Toutefois, la répartition de l'amélioration pour les chimères en fonction de ce rapport de taille montre que le fait d'avoir une confluence réunissant deux bassins très différents en terme de contribution relative en débit moyen n'est pas rédhibitoire pour obtenir un gain de performance avec le modèle semi-distribué (Fig. 25).

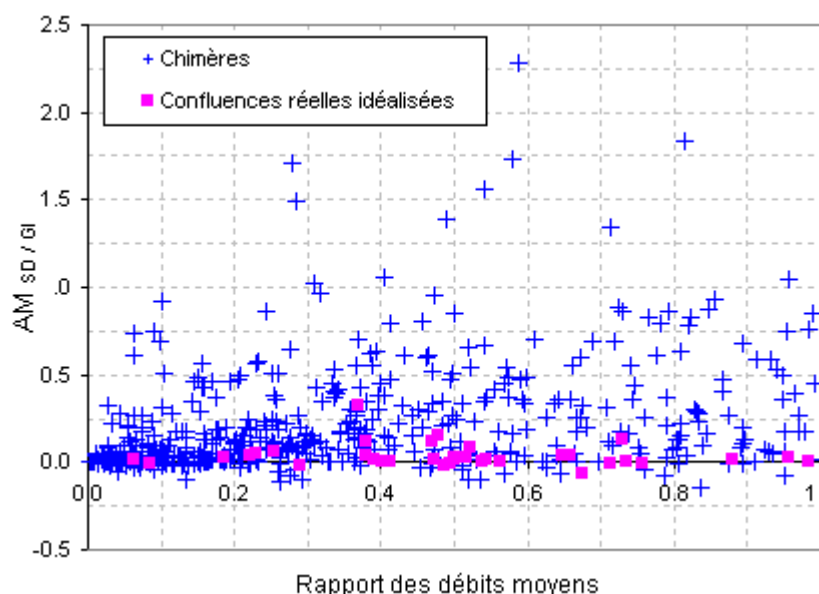


Figure 25 : Amélioration induite par l'utilisation du modèle semi-distribué en fonction du rapport des débits moyens des deux sous-bassins constitutifs du bassin confluence (rapport compris entre 0 [bassins très différents] à 1 [bassins de même taille])

Conclusion : Plus les deux sous-bassins constitutifs de la confluence sont de taille similaire (toutes choses égales par ailleurs), plus l'amélioration à attendre de l'utilisation de la version semi-distribuée est grande. Le fait d'associer deux bassins différents en terme de taille n'est cependant pas le seul facteur contrôlant les variations de performance et donc l'intérêt d'utiliser la configuration semi-distribuée.

5.3.2. Rôle de l'hétérogénéité des sous-bassins

De manière intuitive, l'idée première qui vient en tête est que l'intérêt de mettre en place un modèle semi-distribué augmente lorsque les deux sous-bassins constitutifs de la confluence ont des comportements différents. C'est donc ce que l'on tente de quantifier ici. L'objectif est de savoir si l'hétérogénéité entre les deux sous-bassins, censée être plus forte dans le cas des chimères que pour les confluences réelles, peut être un facteur explicatif du fait que le gain de performance obtenu est faible dans le cas des confluences idéalisées.

On veut donc déterminer si l'amélioration hypothétique apportée par les différentes configurations testées est une fonction de l'hétérogénéité du bassin considéré, si tant est qu'il soit possible de définir l'hétérogénéité d'un bassin versant de manière fiable. La détermination de l'hétérogénéité d'un bassin est en effet délicate, et demeure relativement subjective. Différents critères d'analyse des résultats ont été utilisés. Dans le cadre de cette étude, on ne dispose pour caractériser un bassin utilisé que :

- des coordonnées géographiques et de l'altitude de l'exutoire du bassin ;
- de sa superficie ;
- de la chronique des débits sur au moins 10 ans ;
- de la chronique des pluies sur au moins 10 ans.

Evidemment, il serait également envisageable d'accéder à d'autres données concernant les bassins à partir de Modèles Numériques de Terrains et de cartes d'occupations des sols. Toutefois, le travail consistant à quantifier l'hétérogénéité d'un bassin constitue un projet à part entière, qui a d'ailleurs plusieurs fois été mené par le passé (Plantier, 2003). On se contentera donc de celles évoquées ci-dessus.

A partir de ces données, une méthode envisageable pour quantifier l'hétérogénéité d'un bassin, issu de deux sous-bassins clairement identifiables, consiste à mettre en place une distance mathématique quantifiant l'hétérogénéité entre les deux bassins amont. Cette distance peut se calculer sur la base du rapport des pluies moyennes, du rapport des rendements des deux sous-bassins, etc. Le problème posé par cette méthode est que d'une part, il est difficile d'évaluer une pondération entre chacun des critères utilisés, d'autre part, l'information individuelle contenue dans ces différents rapports est "noyée" dans l'indice d'hétérogénéité mis en place.

En conséquence, une autre méthode a été envisagée : chacun des rapports évoqués ci-dessus (et explicités par la suite) sera étudié comme un facteur explicatif des résultats. L'amélioration éventuelle apportée par une configuration semi-distribuée du modèle sera confrontée à chaque indice d'hétérogénéité mis en place (Tableau 4).

Au préalable, il est à noter que pour chacun de ces critères (qui sont des rapports Caractéristique_{Bassins 1} / Caractéristique_{Bassin 2} noté ici C_1/C_2), le rapport a été transformé par une fonction mathématique, de manière à ce qu'il soit identique dans le cas du rapport C_1/C_2 et de celui C_2/C_1 . Par souci de simplicité, le choix s'est porté sur la fonction $f(C_1, C_2) = \min(C_1, C_2) / \max(C_1, C_2)$, utilisée par la suite.

Indice d'hétérogénéité	Rapport des rendements	Rapport des autocorrelations des débits à 6 heures	Rapport des altitudes
Code	R _{Rend}	R _{AcorQ}	R _{Alt}
Sens donné	Représentatif de l'hétérogénéité de comportement des deux sous-bassins	Caractéristique de l'hétérogénéité des dynamiques des deux sous-bassins	Peut caractériser l'hétérogénéité topographique du bassin
Caractéristique utilisé	Rend = Q _{moyen} / P _{moyen}	R ² corrélation [Q(t) ; Q(t+6h)]	Altitude exutoire
Formule de calcul	Rend _{ss-bass 1} / Rend _{ss-bass 2}	AcorQ _{ss-bass 1} / AcorQ _{ss-bass 2}	Alt _{ss-bass 1} / Alt _{ss-bass 2}
Indice d'hétérogénéité	Rapport des pluies moyennes	Rapport des indices de variabilité des pluies	Corrélation entre les séries de pluies moyennes mensuelles
Code	R _{Pmoy}	R _{IVP}	R ² Pmoy mensuelles
Sens donné	Caractéristique de la différence des régimes pluvieux entre les deux sous-bassins	Caractéristique de la différence des régimes pluvieux entre les deux sous-bassins	Caractéristique de la différence des régimes pluvieux entre les deux sous-bassins
Indice utilisé	Pluie moyenne (Pmoy)	(P _{mens max} - P _{mens min}) / P _{moyen}	Série des 12 Pmoy mensuel
Formule de calcul	Pmoy _{ss-bass 1} / Pmoy _{ss-bass 2}	IVP _{ss-bass 1} / IVP _{ss-bass 2}	R ² corrélation(Série _{ss-bas1} ; Série _{ss-bas2})
Indice d'hétérogénéité	Rapport des débits mensuels maximaux	Rapport des indices de variabilité du débit	Corrélation entre les séries de débits moyens mensuels
Code	R _{Qmax}	R _{IVQ}	R ² Qmoy mensuels
Sens donné	Caractéristique de la différence de régime hydrologique des deux sous-bassin	Caractéristique de la différence de régime hydrologique des deux sous-bassin	Caractéristique du déphasage entre régimes des deux sous-bassins
Indice utilisé	Max _i : 1.>12 [Qmoy (mois i)]	(Q _{mens max} - Q _{mens min}) / Q _{moyen}	Série des 12 Qmoy mensuel
Formule de calcul	Qmax _{ss-bass 1} / Qmax _{ss-bass 2}	IVQ _{ss-bass 1} / IVQ _{ss-bass 2}	R ² corrélation(Série _{ss-bas1} ; Série _{ss-bas2})

Tableau 4 : Description des indicateurs d'hétérogénéité du bassin confluence

Dans un premier temps, il est évidemment très intéressant de visualiser (Fig. 26) les différences qui peuvent exister entre l'échantillon de bassins chimères et l'échantillon de confluences idéalisées. En effet, cela permet d'observer les indices potentiellement explicatifs des différences de résultats largement évoquées auparavant. Le résultat le plus marquant est que les bassins correspondant à des confluences réelles idéalisées sont globalement plus homogènes que les chimères virtuelles. En revanche, le rapport des rendements et des indices de variabilité des débits ont des distributions de critères davantage similaires.

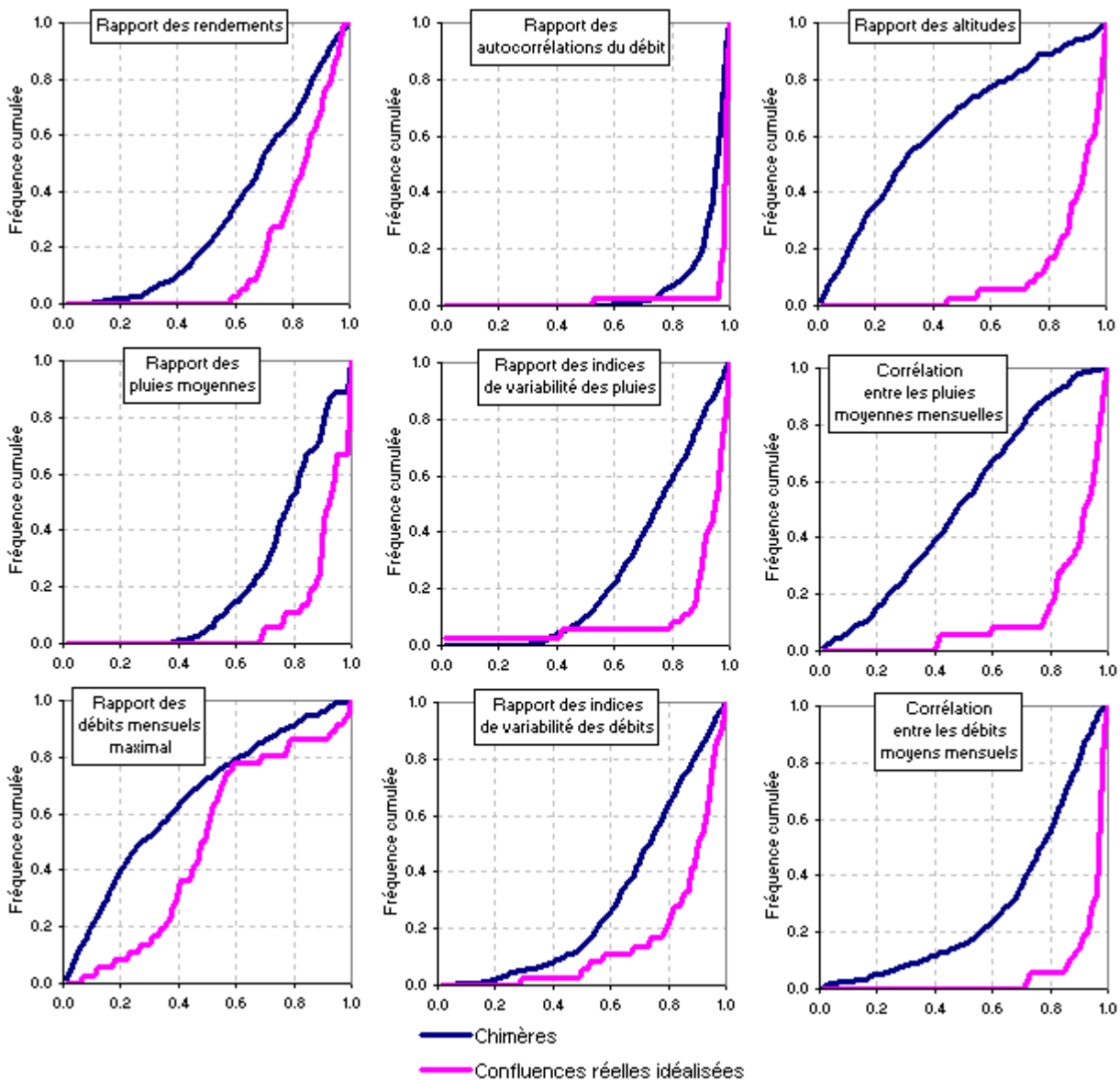


Figure 26 : Graphes comparés de distribution des différents indicateurs d'hétérogénéité entre l'échantillon chimère et l'échantillon de confluences réelles

Dans un second temps, il convient de s'attarder sur la répartition des améliorations induites par l'utilisation d'un modèle semi-distribué en fonction des différents indicateurs d'hétérogénéité (Fig. 27 et 28) :

- la tendance évoquée auparavant avec les graphes de distribution des indicateurs se confirme visuellement pour nombre d'entre eux : les bassins confluences réelles idéalisées sont plus homogènes que les bassins chimères (indice de variabilité des pluies, corrélations entre les séries de débits moyens mensuels et de pluies moyennes mensuelles) ;
- il semble difficile de dire que l'amélioration liée à l'utilisation du GR3P en configuration semi-distribuée par rapport à la configuration globale est une fonction directe d'un des indices. Le rapport des rendements, celui des débits moyens mensuels, le rapport des indices de variabilité des débits, celui d'autocorrélation des débits et la corrélation entre les séries de débits moyens mensuels semblent toutefois indiquer une évolution des performances en fonction de l'hétérogénéité des deux sous-bassins constitutifs.

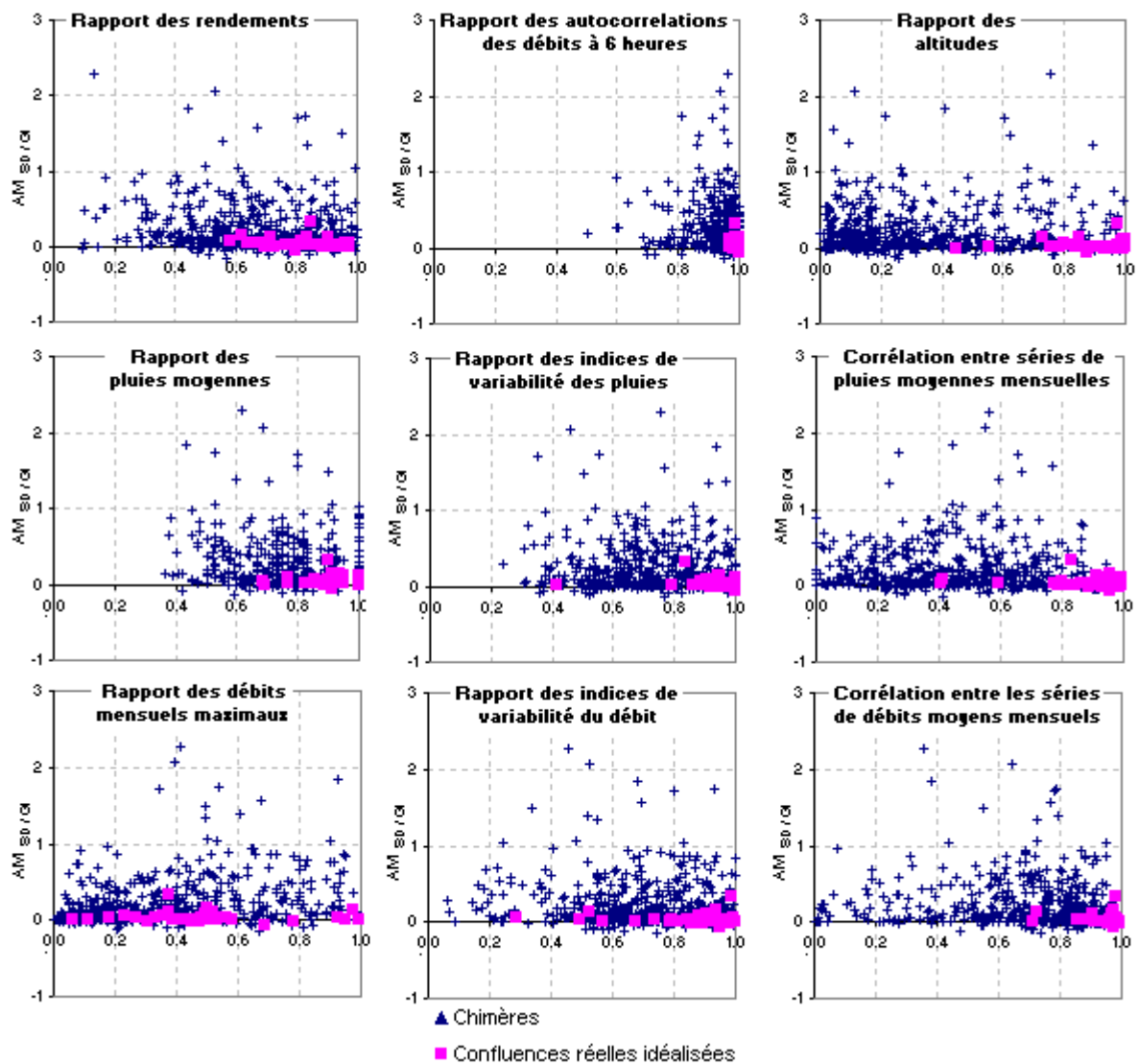


Figure 27 : Graphes représentant l'amélioration induite par l'utilisation d'un modèle semi-distribué en fonction des différents indicateurs d'hétérogénéité pour les deux échantillons de confluences (chimères et réels)

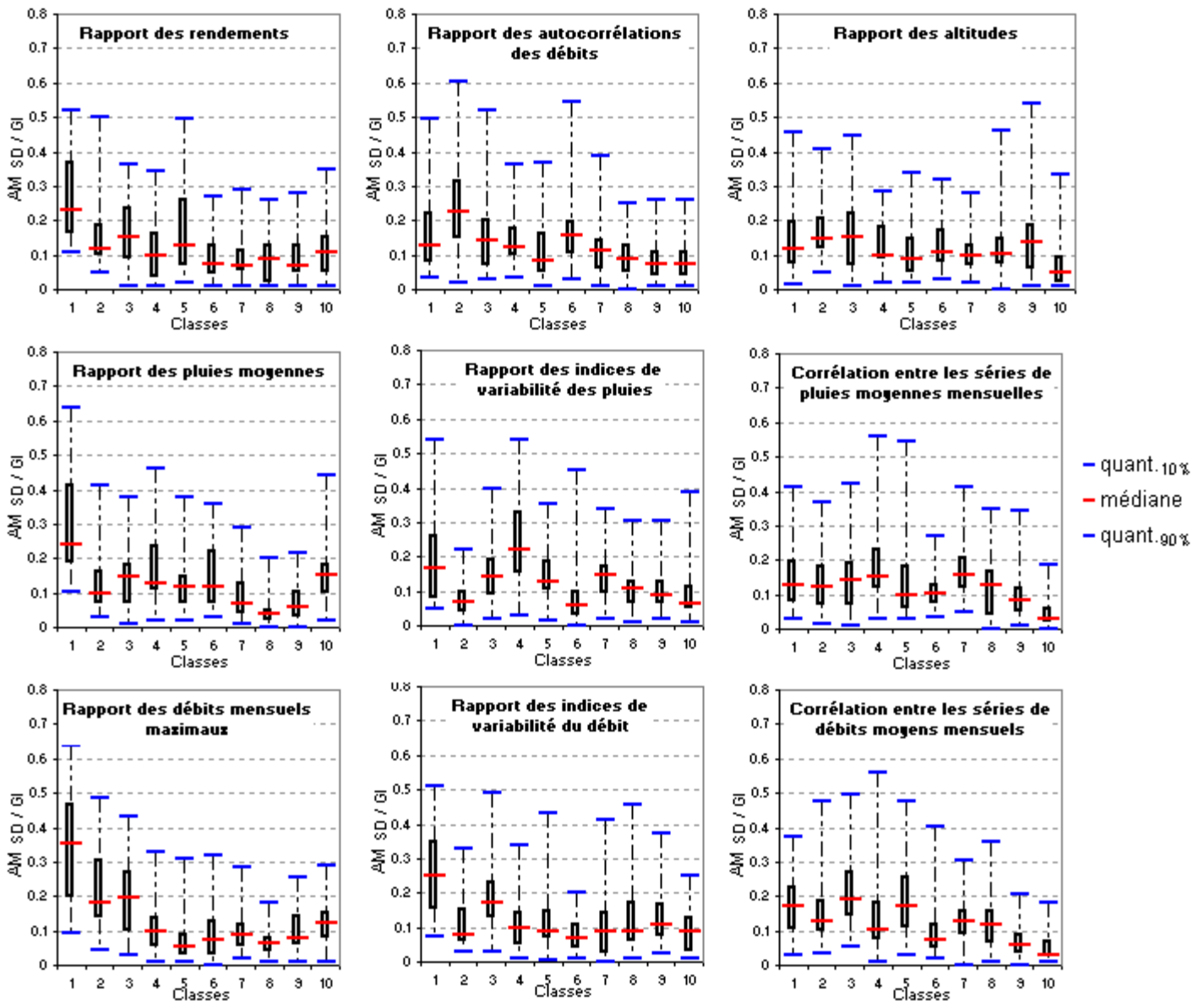


Figure 28 : Graphes représentant l'amélioration induite par l'utilisation d'un modèle semi-distribué en fonction des différents indicateurs d'hétérogénéité – bassins chimères réunis en classes d'effectifs homogènes

Conclusion : L'objectif était ici de tenter de définir des typologies de confluences (en fonction des caractéristiques des sous-bassins constitutifs) pour lesquelles la probabilité d'obtenir une amélioration significative de la qualité des prévisions est plus importante. Si les résultats montrent que cette probabilité augmente lorsque les deux sous-bassins sont de taille (en débit moyen) similaire, cela ne permet pas pour autant d'éliminer les confluences pour lesquelles ce n'est pas le cas.

En outre, ce gain de performance serait intuitivement supposé augmenter lorsque l'hétérogénéité des deux sous-bassins croît. Il n'a pas été possible de confirmer de manière nette cette intuition à l'aide des indices d'hétérogénéité utilisés, certains présentant une tendance dans ce sens. En revanche, deux arguments viennent donner du poids à cette hypothèse :

- les chimères sont en moyenne plus hétérogènes que les confluences idéalisées, et le gain de performance induit par l'utilisation de la version semi-distribuée du modèle est très nettement supérieur dans le cas du sous-échantillon de confluences chimériques ;
- l'autre argument venant étayer cette piste est le fait que, si l'homogénéité du bassin est poussée à l'extrême, aucune évolution de performances ne devrait être observée. Pour confirmer cela, un échantillon de confluences constituées de deux sous-bassins identiques a été créé. Pour ces confluences Amont₁ - Amont₁, les résultats sont éloquentes : les performances pour cet échantillon de bassin sont identiques quelle que soit la configuration utilisée, semi-distribuée ou globale.

6. Résultats sur les confluences réelles

6.1. Résultats

Les performances de la version semi-distribuée du modèle GR3P avec le module de routage sont comparées à la configuration de modélisation globale, utilisée en opérationnel dans certains SPC (Fig. 29). L'objectif est d'évaluer si cette version du modèle peut offrir des perspectives intéressantes pour la prévision des crues.

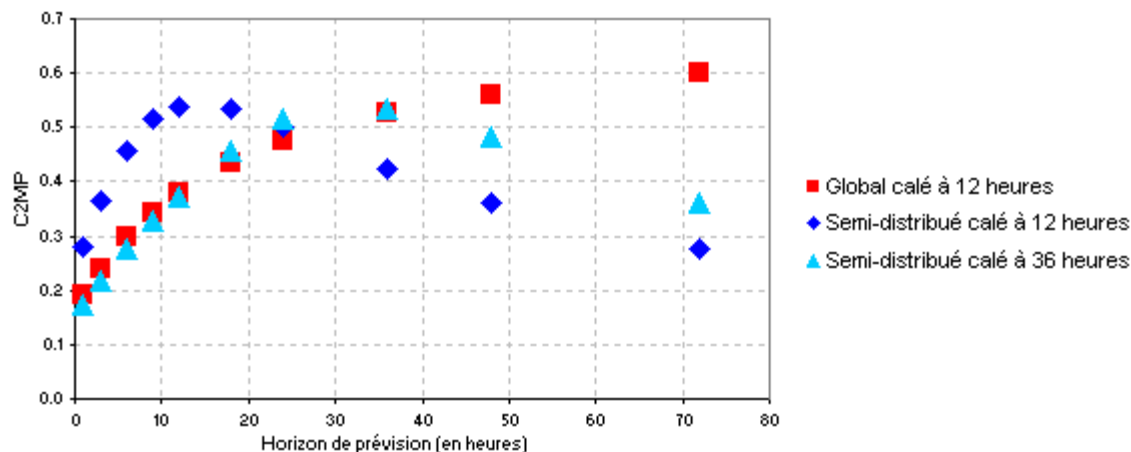


Figure 29 : Performance des différentes configurations de modélisation en fonction de l'horizon de prévision

Le modèle dans sa version semi-distribuée calé à 12 heures surpasse le GR3P global, pour des horizons de prévision jusqu'à 24 heures. Toutefois, il est également possible d'observer une dégradation des performances dès lors que l'on s'éloigne de l'horizon employé au calage, ce qui n'était pas observé pour les confluences chimiques. Les performances du modèle semi-distribué calé à 36 heures confirme que le pic de performance est très ciblé au niveau de l'horizon de calage ainsi que la dégradation des performances pour des horizons différents. Enfin, le troisième point marquant réside dans le fait que la configuration globale du modèle surpasse nettement la configuration semi-distribuée pour les longues échéances de prévision.

6.2. Analyse de sensibilité

Le fait que la version semi-distribuée surpasse la configuration de modélisation globale pour les faibles et moyens horizons de prévisions est conforme à ce que l'on pouvait intuitivement supposer. Pour de courtes échéances, le module hydraulique de propagation de l'eau de l'amont vers l'aval est censé prendre le pas en terme de performances sur un modèle pluie-débit. Pour des horizons de prévision en deçà du temps de propagation obtenu, le modèle fonctionne en effet uniquement à partir de données observées aux stations amont (on se sert en effet de toute l'information disponible au moment de la prévision).

En revanche, la dégradation des performances est le témoin du fait que ce modèle est très sensible à son horizon de calage (objectif d'application). Ces résultats laissent donc penser que le module de routage utilisé doit être utilisé avec précaution et que des solutions alternatives pourraient être testées. Cela laisse entrevoir des perspectives intéressantes d'amélioration : en effet, le module hydraulique mis en place était volontairement simple et facile à mettre en œuvre. Les résultats présentés ici laissent penser que le gain de performance observé aux courts et moyens horizons pourrait se généraliser en utilisant un module hydraulique plus stable rendant mieux compte de la dynamique de comportement du bassin versant intermédiaire.

Différents tests ont été menés pour tenter de mieux cibler les points faibles de la configuration de modélisation semi-distribuée avec module hydraulique, ou du moins de critiquer sa cohérence. De manière à critiquer le modèle, l'expérience des prévisionnistes du SPC SMYL a été mise à profit. Leur connaissance du bassin a permis de dégager des estimations des temps de propagation entre les stations amont et la station de confluence (pour le sous-échantillon de bassins se trouvant sur le territoire d'action du SPC SMYL), permettant ainsi d'avoir un avis critique sur les sorties du modèle de propagation.

Dans un premier temps, on peut s'interroger sur la validité du choix de procéder à l'approximation de n'utiliser qu'une seule célérité propre à la confluence (et pas une par station amont). La figure 30 montre que cette approximation ne semble pas induire de fortes erreurs. En effet, en n'utilisant une valeur moyenne pour les deux stations amont, l'estimation des temps de propagation semble très peu dégradée. Evidemment, il faut garder en tête que les résultats sont comparés à une référence, issue de l'expertise des prévisionnistes, et qui ne saurait être une vérité absolue.

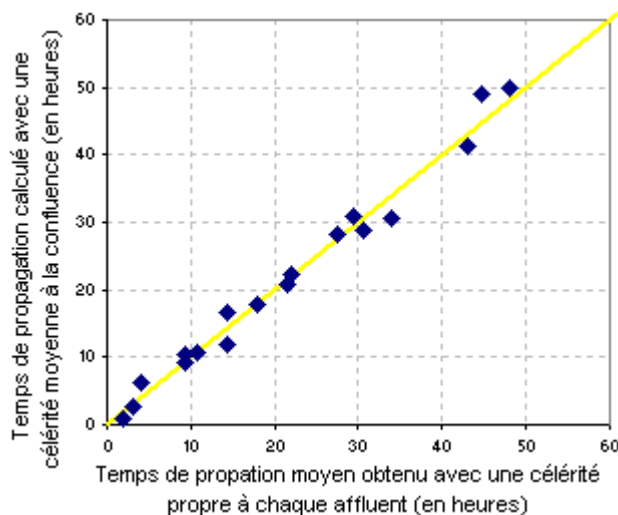


Figure 30 : Comparaison des temps de propagation obtenus en utilisant une célérité commune aux stations amont dans le module hydraulique en fonction de l'utilisation d'une célérité propre à chaque station amont

Ensuite, les délais de propagation obtenus avec le module de routage sont confrontés avec ces mêmes temps estimés par les prévisionnistes (Fig. 31). On peut observer que, pour un calage à 12 heures, les temps de propagation estimés via le modèle ont l'air de correspondre assez bien aux temps de propagation estimés a priori.

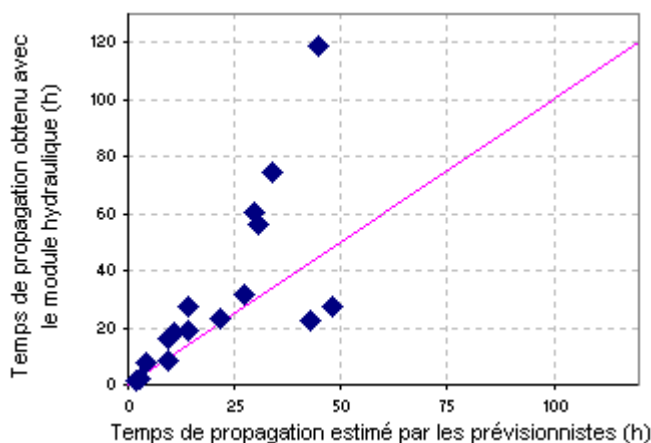


Figure 31 : Comparaison entre les temps de propagation obtenus avec le module hydraulique et ceux estimés par les prévisionnistes du SPC SMYL

Conclusion : Les résultats présentés ici semblent indiquer qu'à horizons lointains, la complexification du schéma de modélisation n'apporte pas de gain de performance, voire est contreproductive. Ce résultat semble assez logique au regard des résultats d'études antérieures menées en simulation, montrant la difficulté d'améliorer les résultats de modèles globaux par la distribution spatiale (cf. projet DMIP par exemple). En effet, à horizons lointains, l'effet de la mise à jour tend à diminuer et le comportement du modèle se rapproche d'un fonctionnement en simulation. L'amélioration constatée pour des horizons courts et moyens suggère quant à elle que mettre en place une approche de semi-distribution doit améliorer la fiabilité des prévisions à courte échéance.

7. Conclusion et perspectives

La question de l'intérêt de la distribution spatiale compte parmi les débats majeurs de la modélisation hydrologique. Dans ce contexte, l'objectif de cette étude était d'évaluer les différences de performances, dans un contexte de prévision, entre une modélisation hydrologique globale et une modélisation semi-distribuée pour des bassins constitués de deux sous-bassins clairement définis. L'application opérationnelle de cette configuration de bassin est la prévision des crues aux confluences. Les confluences sont en effet des cas d'études qui se prêtent bien au passage d'une modélisation globale à une modélisation semi-distribuée. Nous avons utilisé ici un modèle hydrologique global de prévision, le modèle GR3P, qui a été appliqué à chaque bassin ou sous-bassin.

Le problème a été traité en deux temps, l'intérêt de la semi-distribution étant tout d'abord évalué sur des confluences virtuelles sans bassin versant intermédiaire. Ce cadre théorique a notamment été étudié via des chimères, bassins virtuels dont on force le caractère hétérogène en le constituant de deux sous-bassins réels mais déconnectés géographiquement. Dans cette configuration idéalisée, la modélisation semi-distribuée surpasse nettement la modélisation globale. En revanche, pour des confluences réelles sans bassin intermédiaire, l'amélioration observée reste faible. L'analyse faite de ces résultats tend à montrer que la différence entre les deux sous-bassins constitutifs de la confluence joue un rôle important dans l'amélioration à attendre de la modélisation semi-distribuée. Ces résultats restent relativement cohérents avec ceux obtenus en simulation par [Andréassian et al, \(2004\)](#), la pluie ne semblant en revanche plus être aussi déterminante dans l'amélioration induite par l'utilisation d'une modélisation semi-distribuée. Ceci peut s'expliquer par le fait que l'assimilation de données donne plus d'importance au rôle intégrateur du bassin, et donc à ses caractéristiques, en pouvant minimiser l'erreur via la procédure d'assimilation des données de débit.

Un échantillon de confluences réelles a ensuite été utilisé comme cadre appliqué de l'étude, un module de propagation ayant été ajouté pour modéliser l'effet du bassin versant intermédiaire existant entre stations amont et aval. Les résultats ont montré que la version semi-distribuée du modèle surpasse nettement la version globale aux courts et moyens horizons de prévision, laissant entrevoir un potentiel certain. En revanche, les performances de la configuration semi-distribuée se dégradent en s'éloignant de l'horizon de prévision utilisé lors du calage, et pour les horizons de prévision lointains. Cela laisse penser que le module de routage simple mis en place doit être utilisé avec précaution et que des solutions alternatives pourraient être testées.

Une modélisation plus riche du bassin versant intermédiaire, incluant une modélisation des apports diffus ou ponctuels et leur introduction dans le module de propagation, pourrait être source d'amélioration des résultats de l'approche semi-distribuée en prévision, comme le suggèrent les travaux de [Lerat \(2008\)](#) justement consacrés à la prise en compte du bassin intermédiaire.

Les résultats présentés ici ouvrent donc d'intéressantes pistes de recherche pour l'amélioration de l'efficacité des modèles de prévision des crues. Des confrontations à d'autres types de modélisations (distribuées par exemple) permettraient d'avoir un regard complémentaire sur le niveau de discrétisation spatiale nécessaire pour exploiter au mieux les variabilités présentes sur les bassins.

8. Bibliographie

- Ajami, N.K., Gupta, H., Wagener, T. and Sorooshian, S., 2004. Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system. *Journal of Hydrology*, 298(1-4): 112-135.
- Andréassian, V., Oddos, A., Michel, C., Anctil, F., Perrin, C., Loumagne, C., 2004. Impact of spatial aggregation of inputs and parameters on the efficiency of rainfall-runoff models: A theoretical study using chimera watersheds. *Water Resources Research*, 40(5): W05209, doi:10.1029/2003WR002854.
- Baudez, J.C., Loumagne, C., Michel, C., Palagos, B., Gomendy, V., Bartoli, F., 1999. Modélisation hydrologique et hétérogénéité spatiale des bassins. Vers une comparaison de l'approche globale et de l'approche distribuée. *Etude et Gestion des Sols*, 6(4): 165-184.
- Berthet, L., Andréassian, V., Perrin, C. and Javelle, P., 2009. How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. *Hydrology and Earth System Sciences*, 13(6): 819-831.
- Boyle, D.P., Gupta, H.V., Sorooshian, S., Koren, V., Zhang, Z., Smith, M., 2001. Toward improved streamflow forecasts: value of semidistributed modeling. *Water Resources Research*, 37(11): 2749-2759.
- Breuer L., Huisman, J.A., Willems, P., Bormann, H., Bronstert, A., Croke, B.F.W., Frede, H.G., Gräff, T., Hubrechts, L., Jakeman, A.J., Kite, G., Lanini, J., Leavesley, G., Lettenmaier, D.P., Lindström, G., Seibert, J., Sivapalan, M., Viney, 2009. Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: Model intercomparison with current land use. *Advances in Water Resources*, 32(2): 129-146.
- Carpenter, T.M. and Georgakakos, K.P., 2006. Intercomparison of lumped versus distributed hydrologic model ensemble simulations on operational forecast scales. *Journal of Hydrology*, 329(1-2): 174-185.
- Chaubey, I., Haan, C.T., Grunwald, S. and Salisbury, J.M., 1999. Uncertainty in the model parameters due to spatial variability of rainfall. *Journal of Hydrology*, 220(1-2): 48-61.
- Das, T., Bardossy, A., Zehe, E. and He, Y., 2008. Comparison of conceptual model performance using different representations of spatial variability. *Journal of Hydrology*, 356(1-2): 106-118.
- El-Nasr, A.A., Arnold, J.G., Feyen, J. and Berlamont, J., 2005. Modelling the hydrology of a catchment using a distributed and a semidistributed model. *Hydrological Processes*, 19(3): 573-87.
- Khakbaz, B., Imam, B., Hsu, K. and Sorooshian, S., From lumped to distributed via semi-distributed: Calibration strategies for semi-distributed hydrologic models. *Journal of Hydrology*, In Press, Corrected Proof.
- Kitanidis, P.K. and Bras, R.L., 1980. Real-time forecasting with a conceptual hydrologic model. 2. Application and results. *Water Resources Research*, 16(6): 1034-1044.
- Klemeš, V., 1986. Operational testing of hydrological simulation models. *Hydrological Sciences Journal*, 31(1): 13-24.
- Kling, H. and Nachtnebel, H.P., 2009. A spatio-temporal comparison of water balance modelling in an Alpine catchment. *Hydrological Processes*, 23(7): 997-1009.
- Koren, V.I., Finnerty, B.D., Schaake, J.C., Smith, M.B., Seo, D.-J., Duan, Q.-Y., 1999. Scale dependencies of hydrologic models to spatial variability of precipitation. *Journal of Hydrology*, 217(3-4): 285-302.
- Le Moine, N., 2008. Le bassin versant de surface vu par le souterrain : une voie d'amélioration des performances et du réalisme des modèles pluie-débit ? Thèse de Doctorat Thesis, Université Pierre et Marie Curie, Paris, 324 pp.
- Lerat, J., 2009. Quels apports hydrologiques pour les modèles hydrauliques ? Vers un modèle intégré de simulation des crues. Thèse de Doctorat Thesis, Université Pierre et Marie Curie, Paris, 300 pp.
- Lopes, V.L., 1996. On the effect of uncertainty in spatial distribution of rainfall on catchment modelling. *Catena*, 28(1-2):107-119.
- Mathevet, T., 2005. Quels modèles pluie-débit globaux pour le pas de temps horaire ? Développement empirique et comparaison de modèles sur un large échantillon de bassins versants. Thèse de Doctorat Thesis, ENGREF (Paris), Cemagref (Antony), France, 463 pp.
- McIntyre, N. and Al-Qurashi, A., 2009. Performance of ten rainfall-runoff models applied to an arid catchment in Oman. *Environmental Modelling & Software*, 24(6): 726-738.
- Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. Part I - A discussion of principles. *Journal of Hydrology*, 27(3): 282-290.
- Oudin, L., 2004. Recherche d'un modèle d'évapotranspiration potentielle pertinent comme entrée d'un modèle pluie-débit global. Thèse de Doctorat Thesis, ENGREF (Paris) / Cemagref (Antony), 495 pp.
- Plantier, M., 2003. Prise en compte de caractéristiques physiques du bassin versant pour la comparaison des approches globale et semi-distribuée en modélisation pluie-débit, Mémoire du Diplôme d'Ingénieur de l'ENGEES et du DEA Mécanique et ingénierie option Sciences de l'eau, Université Louis Pasteur Strasbourg, 100 pp.
- Reed, S.M., Koren, V., Zhang, Z., Moredaa, F., Seo, D.J., DMIP Participants, 2004. Overall distributed model intercomparison project results. *Journal of Hydrology*, 298(1-4): 27-60.
- Refsgaard, J.C., 1997. Validation and intercomparison of different updating procedures for real-time forecasting. *Nordic Hydrology*, 28: 65-84.
- Singh, V.P. and Frevert, D.K. (Editors), 2002. *Mathematical Models of Small Watershed Hydrology and Applications*. Water Resources Publications, Highlands Ranch, Colorado, 972 pp.
- Smith, M.B., Koren, V., Zhang, Z., Reed, S.M., Panb, J.J., Moredaa, F., 2004. Runoff response to spatial variability in precipitation: an analysis of observed data. *Journal of Hydrology*, 298(1-4): 267-286.
- Tangara, M., 2005. Nouvelle méthode de prévision de crue utilisant un modèle pluie-débit global. Thèse de Doctorat Thesis, EPHE, Paris, 374 pp.
- Yang, X. and Michel, C., 2000. Flood forecasting with a watershed model: a new method of parameter updating. *Hydrological Sciences Journal*, 45(4): 537-546.