



Annexe : Calage automatique des modèles hydrologiques dans le cadre du projet MOSARH21

Evolutions des débits futurs sur le bassin du Rhin en contexte de changement climatique.

Une évaluation actualisée avec les simulations climatiques du 5^e rapport du GIEC

Guillaume Thirel, Kai Gerlinger, avec la participation de Charles Perrin, Gilles Drogue et Jean-Pierre Wagner

16/11/2017

Résumé

Cette annexe a pour objectif de présenter spécifiquement les résultats des calages automatiques réalisés dans le cadre du projet MOSARH21. Le calage automatique est une procédure réalisée de manière routinière avec les modèles GR. En revanche, il s'agit de la première application au modèle LARSIM, qui est normalement calé de manière manuelle suivant un protocole prédéfini. Nous soulignons ici le fait que les deux modèles sont appliqués ici au pas de temps journalier, en cohérence avec les données utilisées dans le cadre du projet. Ils diffèrent des versions horaires de ces modèles utilisées pour d'autres applications (par ex. la prévision en temps réel). L'application d'un calage automatique a pour but de permettre la réalisation des nombreux calages prévus dans le protocole de test des modèles décidé dans le projet, en particulier le calage sur des périodes différentes afin d'estimer l'incertitude qui en résulte sur des projections hydrologiques. Ces tests n'auraient pas été réalisables avec des calages manuels. Les résultats indiquent que les calages réalisés étaient plutôt robustes, notamment ceux de GRSD ainsi que ceux de LARSIM sur les stations qui ont été calées. La dégradation observée sur les différents critères (KGE, bilan, quantiles) reste limitée en contrôle. Les calages sur des périodes différentes présentent cependant des jeux de paramètres et des performances légèrement différents. Ceux-ci pourront donc aider à mesurer comment l'incertitude liée à cette étape de modélisation impacte les projections hydrologiques.

Contenu

Contenu	3
1 Introduction.....	4
2 Méthodologie	5
2.1. Les stations de calage.....	5
2.2. Le calage automatique des modèles	6
2.3. Les procédures de calage-contrôle	13
3 Analyse des résultats.....	15
3.1. Comparaison de la méthode de calage automatique appliquée à LARSIM avec la méthode manuelle.....	15
3.2. Réduction des jeux de données en entrée du modèle.....	17
3.3. Vérification du bon fonctionnement du calage automatique de LARSIM	18
3.4. Cartes des paramètres des modèles LARSIM et GRSD.....	26
3.5. Capacité des modèles à reproduire les hydrogrammes.....	28
3.6. Capacité des modèles à reproduire les courbes de régime	29
3.7. Cartographie des métriques d'erreur.....	30
3.8. Conclusions sur les calages des deux modèles.....	38
4 Calages complémentaires de GRSD : calages effectués sur des périodes glissantes.....	38
5 Conclusions.....	40
6 Références.....	41

1 Introduction

Le calage des modèles hydrologiques, c'est-à-dire la détermination d'un certain nombre de paramètres libres afin de représenter de la manière la plus précise possible des chroniques de débits, est un processus complexe. En effet, il existe rarement des observations permettant de déterminer de manière empirique chacun des paramètres libres (on pense par exemple à des cartes d'occupation du sol, de profondeur du sol, de conductivité hydraulique, etc.). Ainsi, en hydrologie, on se contente bien souvent de tester des valeurs pour ces paramètres qui, *in fine*, permettront d'obtenir les meilleurs débits, sans s'attarder sur d'autres variables du modèle. Les calages peuvent être réalisés de manière manuelle (le modélisateur décide lui-même des ajustements qu'il veut faire pour optimiser les paramètres) ou bien automatique (un algorithme est défini afin d'explorer au mieux l'espace des possibles pour les valeurs des paramètres). De plus, l'évaluation de la qualité des paramètres que l'on teste peut se faire de manière visuelle (en observant graphiquement la ressemblance entre les hydrogrammes simulés et ceux observés) ou bien de manière numérique (en utilisant une fonction-objectif, c'est-à-dire un critère de performance, dont on essaie d'optimiser la valeur). Le recours à un algorithme automatique requiert l'utilisation d'un critère numérique d'évaluation, appelé fonction objectif, même si cela n'empêche pas une analyse visuelle des hydrogrammes lorsque l'algorithme a terminé son exécution.

Depuis les années 1990, un algorithme de calage automatique est utilisé à Irstea pour le calage automatique des modèles GR (Michel, 1991). En revanche, à HYDRON, le calage du modèle LARSIM est réalisé de manière manuelle et évalué visuellement. En effet, différentes valeurs des paramètres sont testées selon l'expérience du modélisateur et en essayant d'améliorer successivement différents aspects de l'hydrogramme. Dans les deux cas, on s'intéresse à une période passée donnée, sur laquelle le modélisateur essaie d'obtenir la simulation de débits la plus proche des observations de débit disponibles. Si possible, on effectue une évaluation supplémentaire sur une période indépendante de celle utilisée en calage afin de tester la transposabilité temporelle des jeux de paramètres calés.

Pour les besoins du projet MOSARH21, il a été nécessaire de mettre en place des calages sur différentes périodes. En effet, il est désormais reconnu en hydrologie que le choix de la période de calage influence la qualité des débits simulés, et ce d'autant plus que l'on utilise le modèle calé sur une période climatiquement et hydrologiquement différente de la période de calage (Coron, 2013). En règle générale, plus les caractéristiques de la période d'étude sont éloignées de celles de la période de calage, moins les performances du modèle sont bonnes, ce qui signifie que les erreurs du modèle ont tendance à s'amplifier avec les contrastes climatiques. Dans MOSARH21, nous sommes dans ce cas : les modèles sont calés sur une période passée, et on réalise des projections de débits dans le futur. Les conditions futures étant très éloignées des conditions passées (la température de l'air augmente, les précipitations évoluent de manière incertaine), il est logique de s'attendre à ce que les erreurs que l'on fera sur l'estimation des débits futurs (si on avait les observations pour les quantifier) soient plus importantes que celles faites en conditions présentes. Donc, pour essayer de répondre à la question du niveau de confiance que l'on peut mettre dans la modélisation des projections hydrologiques, nous avons mis en place un protocole de calage des modèles sur différentes périodes. Ces différents calages seront évalués. De plus, en réalisant des projections utilisant chacun des calages différents, on a ainsi une estimation de l'incertitude résultant de cette

étape de modélisation. Par ailleurs, le fait d'utiliser deux modèles hydrologiques différents permet d'avoir une estimation de l'incertitude liée au choix du modèle.

Dans la suite de ce rapport, nous reviendrons brièvement sur les stations hydrométriques utilisées, le protocole mis en place, et les méthodes de calage. Ensuite, nous nous consacrerons à l'analyse et à la discussion des résultats.

2 Méthodologie

2.1. Les stations de calage

Les stations utilisées pour ces tests sont présentées dans le rapport final de MOSARH21 (voir section 3 et Annexe A du rapport final). Le modèle GRSD d'Irstea a été calé de manière automatique sur l'ensemble de ces stations. En raison de temps de calcul prohibitifs, seule une sélection de ces stations a été calée de manière automatique pour le modèle LARSIM d'HYDRON : il s'agit des stations indiquées comme appartenant à la catégorie 1 sur ce rapport final. On a représenté l'ensemble de ces deux jeux de stations de manière synthétique sur la Figure 1 du présent rapport (les stations calées dans LARSIM sont surlignées en vert). Il est donc important de noter, lorsque l'on effectue une comparaison entre les deux modèles ou lorsque l'on analyse les résultats, qu'une distorsion existe entre les deux modèles : il est en effet difficile d'attendre de la part de LARSIM le même niveau de performance sur l'ensemble du bassin versant notamment sur les zones où le calage aura moins d'influence.

HYDRON disposait aussi de paramètres calés manuellement pour les stations de catégorie 1 des bassins versants Moselle-Sarre. Par ailleurs, dans le cadre de MOSARH21, six stations supplémentaires ont été calées manuellement avec LARSIM sur le territoire de l'Alsace. Cependant, en raison d'une erreur, la station de Didenheim a été calée manuellement à la place de la station d'Holtzheim. A la différence des calages automatiques réalisés sur des périodes différentes, ces calages manuels n'ont été réalisés qu'une seule fois. Les calages manuels ont été réalisés sur la période 1970-2010 pour l'Alsace, avec l'accent mis sur les années récentes, et sur 1998-2003 pour le reste des stations.

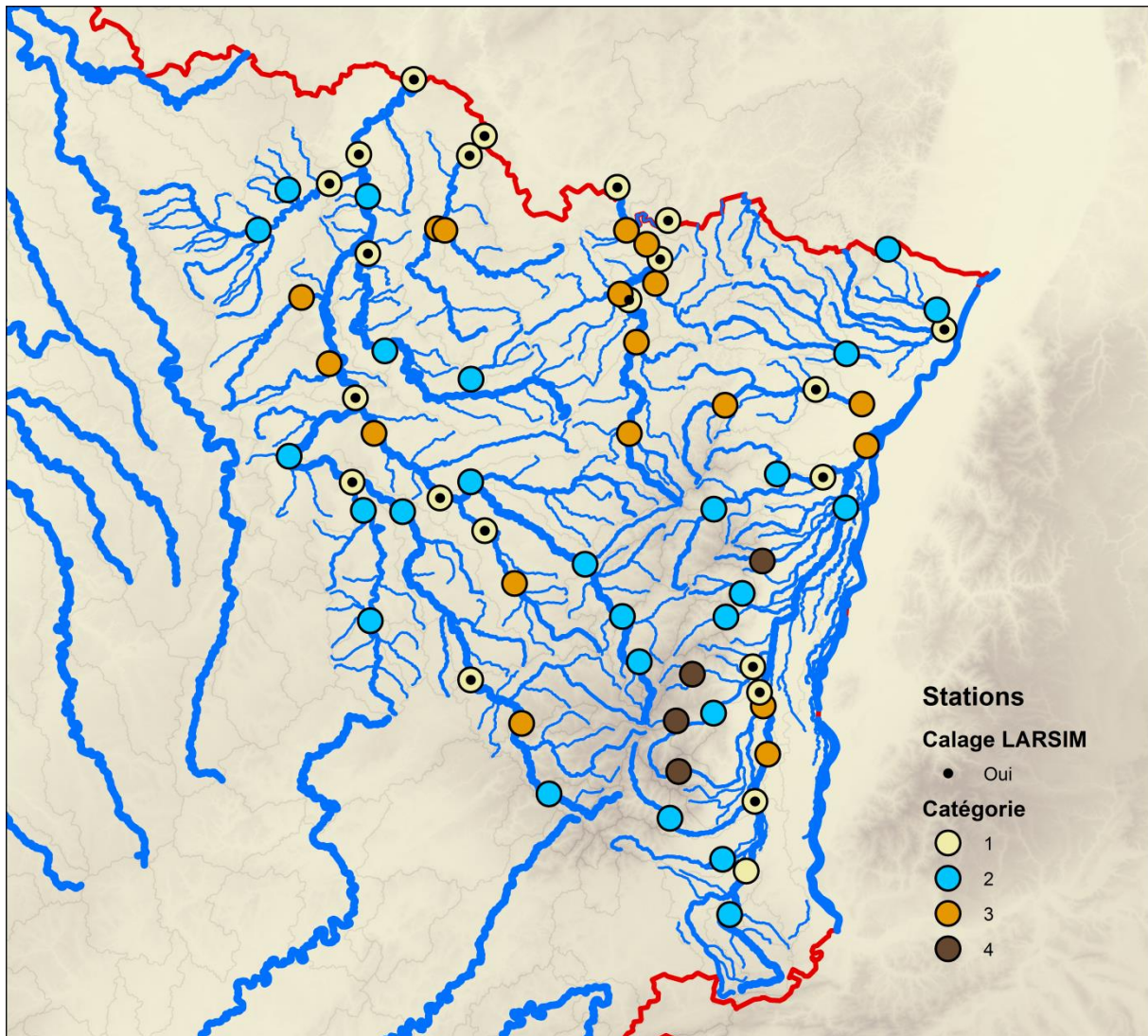


Figure 1 : Liste des stations. Toutes les stations sont calées avec GRSD dans le projet. Les stations calées de manière automatique pour LARSIM apparaissent avec un point noir.

2.2. Le calage automatique des modèles

Principe général du calage automatique

Le calage d'un modèle a pour objectif de rechercher dans un espace à p dimensions, p étant le nombre de paramètres libres du modèle, les valeurs optimales des paramètres conduisant à la meilleure qualité de simulation du modèle. L'espace à p dimensions pouvant être complexe à explorer, cette recherche requiert une stratégie prédéfinie, qui peut être automatisée sous la forme d'un algorithme numérique. Cet algorithme va, généralement de manière itérative, permettre une exploration progressive de l'espace à p dimensions en évaluant la qualité des jeux de paramètres correspondants, jusqu'à identifier ce qui peut être considéré comme un optimum au regard des critères de proximité entre les débits observés et simulés, qui auront été définis préalablement (choix d'une ou plusieurs fonctions objectif). On peut distinguer des procédures de calage locales, qui partent d'un point de l'espace des paramètres et explorent les environs, et les approches globales,

qui partent d'une population de jeux de paramètres et font évoluer cette population vers un jeu optimal. Les algorithmes de calage sont soumis, comme les procédures manuelles, au risque de ne pas identifier le vrai optimum de l'espace des paramètres, et de tomber sur un optimum dit secondaire, présentant une moindre performance du modèle. Les approches locales sont souvent davantage sujettes à ce genre de risque, du fait d'une possible dépendance au point de départ. Elles présentent cependant l'avantage d'être beaucoup moins gourmandes en temps de calcul.

Algorithme d'optimisation utilisé

Ici, l'approche de calage automatique est basée sur l'algorithme développé et utilisé à Irstea Antony depuis une vingtaine d'années. Cet algorithme a été très largement testé sur une large gamme de modèles présentant des niveaux de complexité différents (Nascimento, 1995 ; Perrin, 2000 ; Mathevet, 2005 ; Coron, 2013), et a montré sa capacité à produire des jeux de paramètres optimaux.

Il s'agit d'un algorithme en deux étapes, combinant les forces des approches globale et locale :

1. Un préfiltrage, s'inspirant d'une approche globale, qui permet de parcourir grossièrement l'espace des paramètres pour identifier la zone probable de localisation de l'optimum. Cette zone sert de point de départ à la deuxième étape,
2. Une recherche locale (de type gradient) dite « pas à pas » qui permet d'explorer de manière plus raffinée les environs de ce point de départ pour identifier l'optimum.

Cette procédure en deux temps minimise le risque de tomber sur un optimum secondaire grâce au préfiltrage, tout en maintenant des temps de calcul raisonnables grâce à l'efficacité de la recherche locale. Les risques de ne pas identifier le vrai optimum sont ainsi très limités. Le fonctionnement de l'algorithme est décrit plus en détail dans les paragraphes suivants.

Le préfiltrage des paramètres est le test de combinaisons de valeurs prédéfinies pour chacun des paramètres du modèle. Concrètement, on définit deux ou trois valeurs découpant de manière équilibrée la distribution a priori des paramètres pour chacun d'entre eux, puis on teste l'ensemble des combinaisons possibles de ces valeurs de paramètres pour garder celle qui permet d'obtenir les meilleures performances du modèle. Par exemple, si l'on a un modèle à quatre paramètres et que l'on définit trois valeurs possibles pour chacun d'entre eux, on testera $3^4 = 81$ jeux de paramètres correspondant à un échantillonnage large de l'espace des paramètres. Cette étape permet d'obtenir un point de départ de bonne qualité pour l'étape d'exploration locale, et limite les risques d'obtention d'un optimum secondaire.

La méthode « pas-à-pas » est une méthode locale développée à Irstea (Michel, 1989 ; Nascimento, 1995). Il a été montré que son rapport « qualité-prix » (avec la qualité étant la performance des modèles en validation, et le prix le coût en temps de calcul) est nettement supérieur à celui de méthodes globales qui explorent plus largement l'espace des paramètres. Une description plus détaillée de cette méthode, issue de la thèse de Perrin (2000), est donnée ci-dessous.

Le processus d'optimisation est itératif (Figure 2). La méthode adopte une stratégie de déplacement, la plupart du temps, le long des axes de l'espace des paramètres, avec un pas de recherche pouvant varier d'une itération à l'autre. L'amplitude du pas de recherche étant ici la même pour tous les paramètres, des transformations mathématiques préalables (transformations logarithmiques ou puissance par exemple) peuvent être appliquées pour garantir des sensibilités grossièrement

équivalentes à ce pas de recherche pour tous les paramètres. Ces transformations sur les paramètres sont choisies en tenant compte de la façon dont les paramètres interviennent dans le modèle, et sont propres à celui-ci.

La recherche démarre à partir d'un vecteur initial des paramètres x^0 , de composantes $(x_1^0, x_2^0, \dots, x_n^0)$. On calcule alors la valeur correspondante de la fonction-objectif. On fait ensuite varier successivement chacune des valeurs des paramètres d'une déviation initiale $\pm \Delta x_{ini}$ (ici on adopte $\Delta x_{ini} = 0,64$, ce qui revient à multiplier ou diviser par deux la valeur du paramètre lorsque la transformation est logarithmique). Lorsque la valeur du i^e paramètre est modifiée, les deux vecteurs de paramètres testés sont donc $(x_1^0, x_2^0, \dots, x_i^0 + \Delta x_{ini}, \dots, x_n^0)$ et $(x_1^0, x_2^0, \dots, x_i^0 - \Delta x_{ini}, \dots, x_n^0)$. A chaque fois la valeur de la fonction-objectif correspondante est calculée.

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

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

2. il n'y a aucune amélioration de la valeur de la fonction-objectif pour aucune des modifications des paramètres. On affine alors la recherche en divisant le pas de recherche par deux, et on réitère les modifications des paramètres sur le même vecteur initial x^0 avec ce nouveau pas de recherche.

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

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

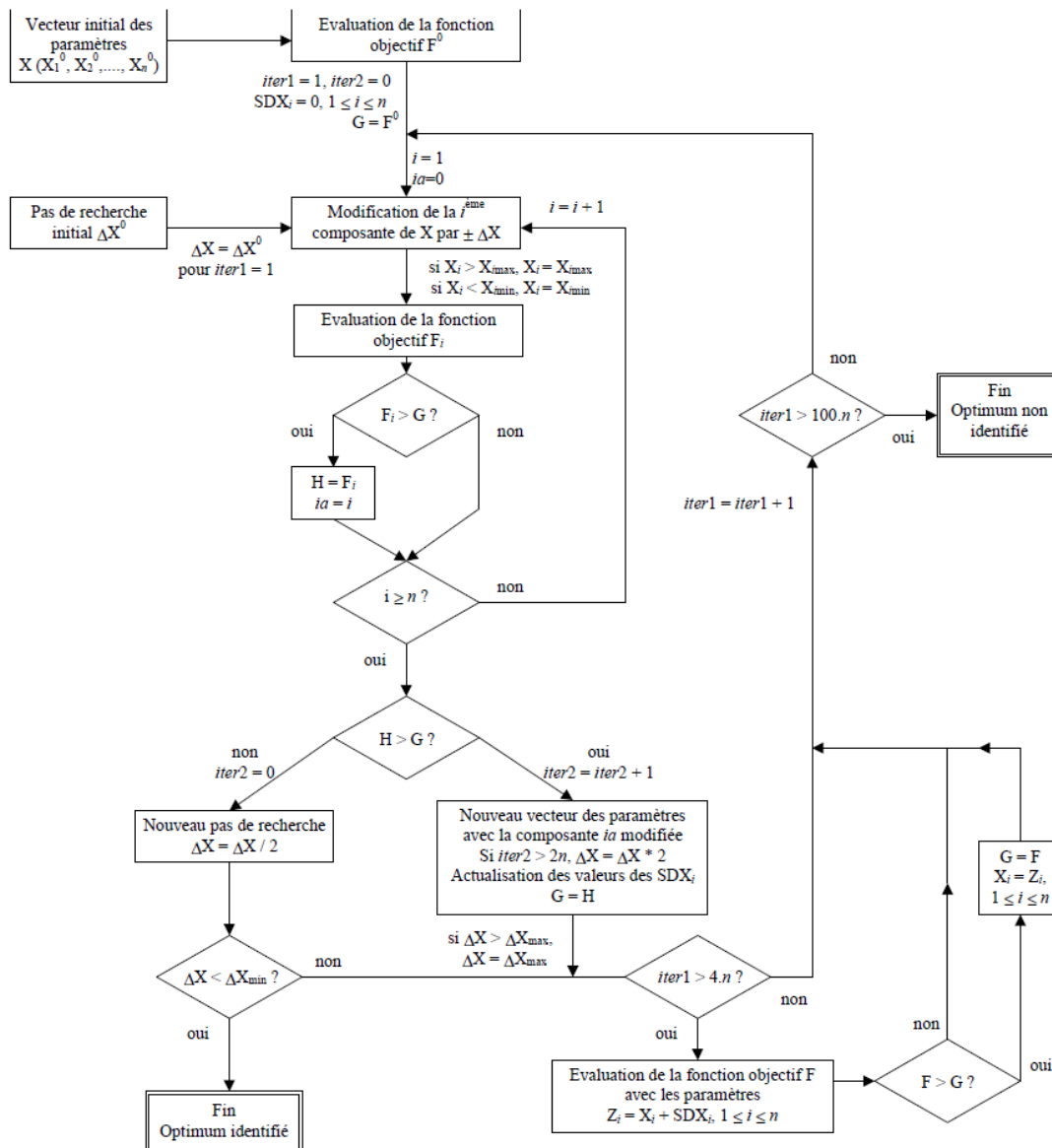


Figure 2 : Diagramme schématique de la méthode pas-à-pas (tiré de Perrin, 2000)

Application de l'algorithme au modèle GRSD

Pour GRSD, cet algorithme est mis en place pour optimiser les débits simulés en calant l'ensemble des paramètres de manière simultanée. Afin de ne pas donner un poids trop important aux débits les plus élevés, nous avons utilisé une fonction objectif réalisant la moyenne de deux critères : le critère KGE' classique, qui donne plus de poids aux forts débits, et un KGE' appliqué à une transformation de type inverse des débits, qui donne plus de poids aux faibles débits (voir section 3.6.1 du rapport final pour des informations complémentaires sur ces critères).

Application de l'algorithme au modèle LARSIM

L'algorithme a été adapté au modèle LARSIM en essayant de reproduire la manière dont le modèle est calé en mode manuel par HYDRON. En effet, ce calage manuel est habituellement réalisé en quatre étapes visant à reproduire successivement :

1. le bilan,
2. le débit de base,
3. le débit intermédiaire,
4. les pics de crue.

Sept des paramètres de LARSIM, jugés par HYDRON comme étant les plus sensibles pour reproduire les débits sur la zone d'étude, ont ainsi été calés automatiquement :

- KG, facteur de correction des précipitations
- beta, indice de drainage de sol profond
- EQB, constante de rétention du réservoir souterrain
- Dmax, indice de drainage du réservoir de sol supérieur
- EQI, constante de rétention du réservoir d'écoulement de subsurface
- EQD, constante de rétention du réservoir d'écoulement direct
- EQD2, 2^e constante de rétention du réservoir d'écoulement direct.

Les autres paramètres de LARSIM, qui sont moins sensibles, ont été fixés à des valeurs fournies par HYDRON. Les sept paramètres ont été calés en quatre étapes successives, par groupes, en fonction de leur rôle dans la détermination des quatre composantes de débit mentionnées plus haut :

1. Calage de KG (notion de bilan) : Le coefficient KG, qui sert à corriger les précipitations, et tout d'abord calé pour équilibrer le bilan, c'est-à-dire pour simuler des débits proches en moyenne des débits observés (Figure 3). Pour ce calage, nous avons utilisé un critère de bilan, c'est-à-dire le rapport entre la somme des débits simulés et la somme des débits observés, rapport que l'on cherche à obtenir le plus proche de un possible.

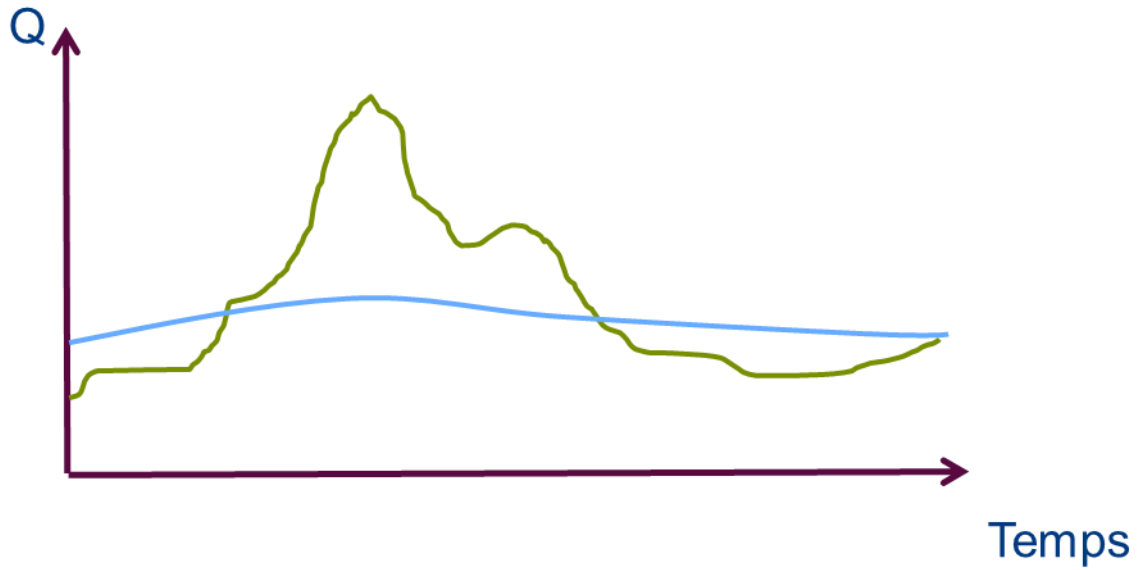


Figure 3 : Exemple d'un débit issu d'un calage complet (en vert) et d'un débit de base (en bleu) dont seul le coefficient KG a été calé

2. Calage de beta et EQB (débit de base) : Dans un deuxième temps, les paramètres beta et EQB, régissant le drainage du sol profond et le souterrain, sont calés. Afin de reproduire l'impact sur le débit de base qu'ont ces paramètres (Figure 4), ils ont été optimisés grâce à l'utilisation d'un KGE' appliqué sur le log des débits. Cette transformation donne un poids plus important aux erreurs commises sur les gammes de débits les plus faibles.

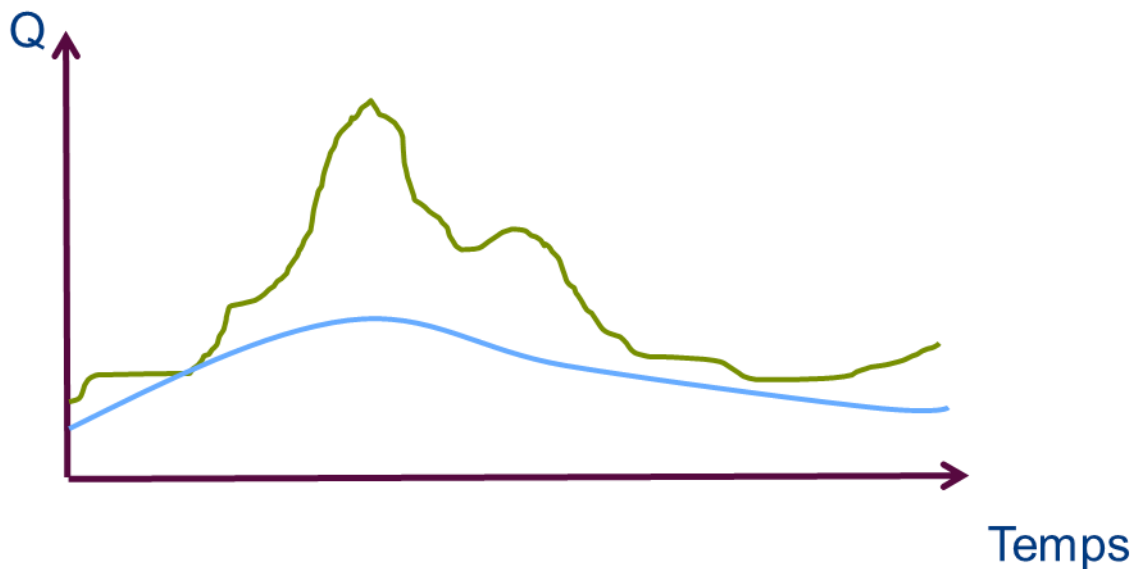


Figure 4 : Résultat du calage des coefficients beta et EQB qui impactent la composante long terme (débit de base) des débits. En vert le débit après calage complet, en bleu le débit de base après calage des deux coefficients

3. Calage de Dmax et EQI (débit intermédiaire) : Ensuite, les paramètres Dmax et EQI sont calés (Figure 5). Ces coefficients impactent le débit intermédiaire de LARSIM (visible sur la courbe rose de la Figure 5), puisqu'ils régissent le drainage dans le sol et le réservoir de subsurface. Pour refléter cet impact, une fonction-objectif portant sur le KGE appliqué sur la racine carrée des débits est utilisée. Cette transformation donne un poids relativement équivalent aux débits les plus forts et aux débits les plus faibles.

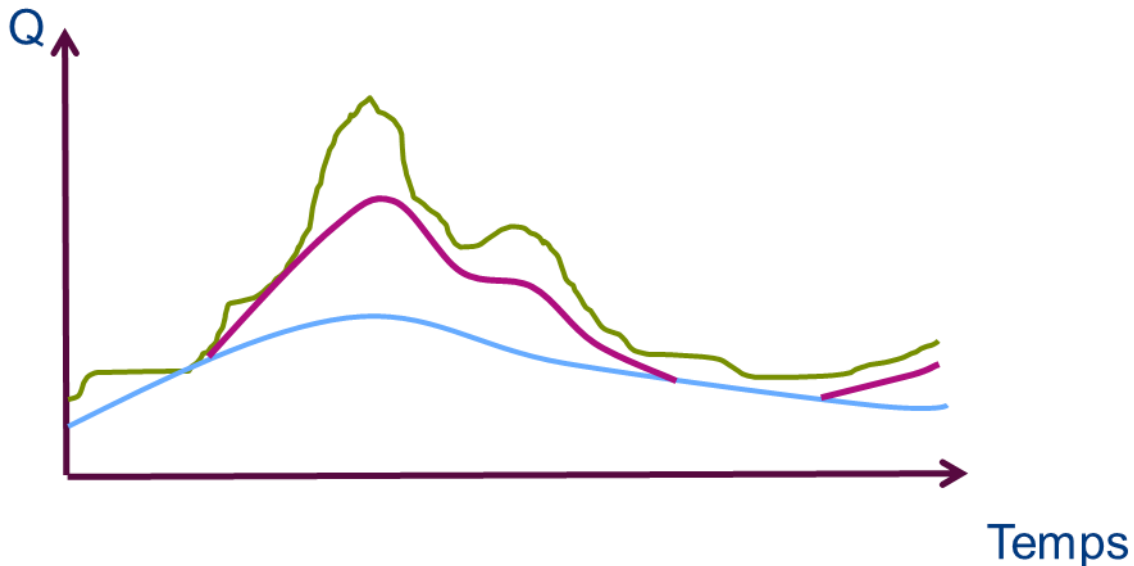


Figure 5 : Résultat du calage des coefficients Dmax et EQI qui impactent la composante intermédiaire des débits. En rose, le débit intermédiaire de LARSIM après calage des deux coefficients

4. Calage de EQD et EQD2 (pics de crues) : Les deux derniers paramètres calés, EQD et EQD2, régissent l'écoulement direct, c'est-à-dire la composante rapide des débits, qui intervient pour les débits les plus forts. Par conséquent, nous avons utilisés un KGE' normal (i.e. appliqué sur les débits non transformés) pour optimiser ces deux derniers paramètre. Sur la Figure 5, cela représente le passage de la courbe rose à la courbe verte finale.

Les quatre étapes du calage comprennent chacune un préfiltrage suivi d'un calage pas-à-pas des paramètres correspondant à l'étape considérée. Les différentes étapes pouvant avoir un impact sur les composantes des étapes précédentes, nous avons mis en place, après ce premier passage, une deuxième application de la méthode pas-à-pas aux quatre étapes sans en changer l'ordre. Il est donc clair que cette procédure de calage représente un compromis : on ne disposera pas pour chaque composante de l'hydrogramme du meilleur jeu de paramètres possible, mais on aura plutôt la combinaison de paramètres qui permettra d'obtenir des résultats satisfaisants pour l'ensemble des composantes du débit : il s'agit donc d'un **compromis**.

Nous avons donc calé de manière automatique sept des paramètres de LARSIM (les paramètres les plus sensibles), un pour le bilan et deux pour chacune des trois étapes suivantes, en utilisant la même procédure que pour GRSD. Cependant, il faut noter que le calage manuel de LARSIM comporte aussi une vérification des importances relatives des composantes du débit, ce qui n'a pas pu être mis en place ici. Cela signifie que certaines combinaisons de paramètres obtenues avec le calage

automatique peuvent paraître « irréalistes » et n’auraient jamais été identifiées par un modélisateur humain car elles se situeraient hors des combinaisons normales. Comme l’objectif de ce calage automatique est de systématiser des calages sur des périodes différentes pour prendre en compte cette incertitude sur les projections hydrologiques, nous considérons cependant que cette limitation n’est pas bloquante. Par ailleurs, seuls sept des paramètres de LARSIM sont calés : les autres sont fixés à des valeurs expertisées le plus souvent, ce qui peut revenir à considérer le calage automatique comme un perfectionnement du calage manuel dans une certaine mesure.

Notons que la mise en place de cette procédure de calage spécifique de LARSIM a conduit au développement d’un code spécifique à Irstea, construit avec les conseils d’HYDRON sur les démarches de calage manuel utilisées.

2.3. Les procédures de calage-contrôle

Si l’utilisation de calages sur différentes périodes pour réaliser des projections permet d’avoir une estimation de l’incertitude liée à cette étape, il est important au préalable d’évaluer ces calages sur des périodes passées indépendantes : il s’agit du calage-contrôle. Cette étape permet d’estimer la robustesse des calages, c’est-à-dire leur capacité à être utilisés sur des périodes aux caractéristiques climatiques et hydrologiques différentes.

La période de référence de nos expériences (1971-2000) sera ainsi divisée en deux parties égales : 1971-1985 (P1) et 1986-2000 (P2), que nous utiliserons tour-à-tour pour le calage automatique et pour le contrôle (Figure 6). De plus, des calages sur la période complète (1971-2000) ont été réalisés.

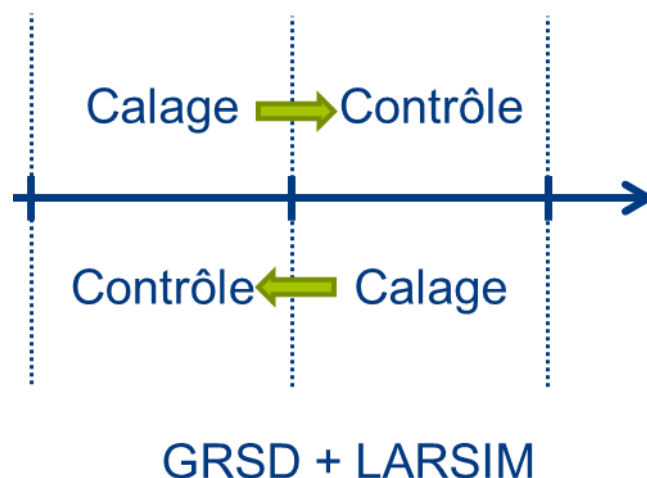


Figure 6 : Procédure de calage-contrôle utilisée pour GRSD et LARSIM

Les périodes P1 et P2 présentent des conditions climatiques qui diffèrent l’une de l’autre. En effet, les précipitations SAFRAN annuelles moyennes sur la totalité de la zone d’étude sont de 900 mm environ sur P1 et de 960 mm sur P2. La température de l’air augmente quant à elle de 9.1°C à 9.7°C. L’évapotranspiration potentielle augmente de 630 à 655 mm. Si ces variations peuvent être considérées comme relativement limitées à première vue, elles représentent déjà une évolution significative sur la température en particulier. Il est intéressant de mesurer la capacité des modèles à s’y adapter. Nous pouvons observer sur la Figure 7 les évolutions mensuelles. Il apparaît que la période P2 est bien la plus chaude, quel que soit le mois considéré. Par ailleurs, la période P1 se caractérise par des hivers et des automnes en moyenne plus secs que P2.

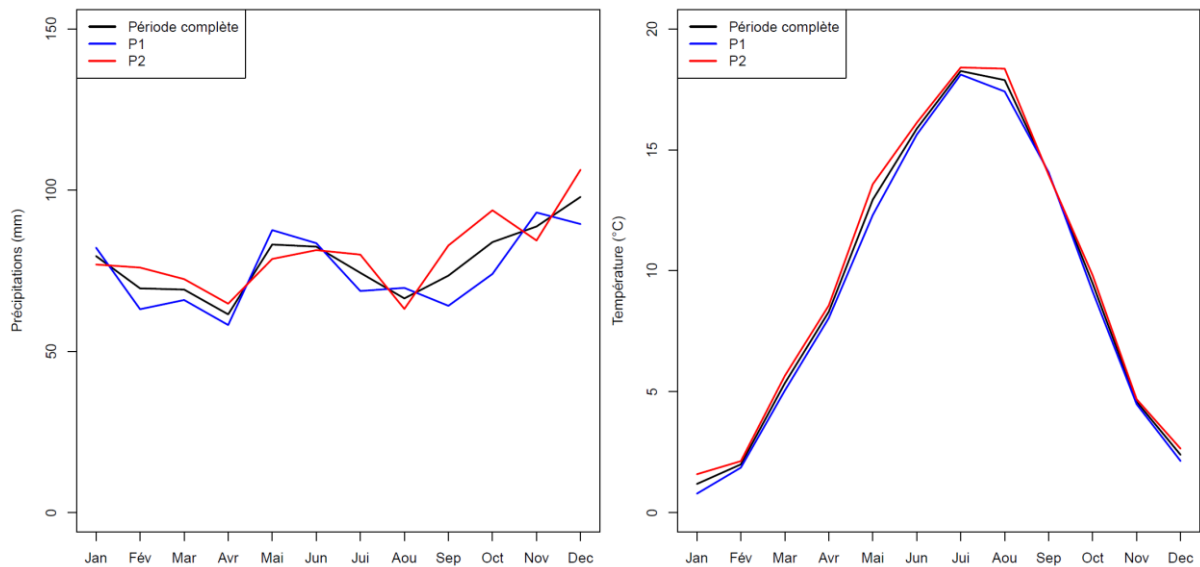


Figure 7 : Evolution mensuelle des précipitations et de la température sur les périodes d'étude (P1 : 1971-1985 ; P2 : 1986-2000) sur l'ensemble de la zone du projet MOSARH21. L'évapotranspiration potentielle suivant les mêmes évolutions que la température, elle ne figure pas sur ces graphes.

En plus de cela, des calages supplémentaires ont été réalisés : il s'agit de calages sur des périodes glissantes de 10 ans (Figure 8). Ce type de calage permet d'observer d'éventuelles évolutions graduelles de performances. Ceux-ci ont été réalisés avec GRSD, seul modèle permettant un nombre important de calages pour des raisons de temps de calcul (en effet, on réalise alors 21 calages au maximum au lieu de deux).



Figure 8 : Procédure de calage sur des périodes glissantes de 10 ans pour GRSD (tiré de de Lavenne et al., 2016)

3 Analyse des résultats

3.1. Comparaison de la méthode de calage automatique appliquée à LARSIM avec la méthode manuelle

Nous avons cherché à éprouver l'efficacité du calage automatique ainsi mis en place pour les sept paramètres de LARSIM, pour évaluer s'il est opportun de l'utiliser dans le but de réaliser les calages sur des périodes contrastées. En effet, cette méthode de calage est nouvellement appliquée à LARSIM. Nous l'avons testée sur le bassin versant de la Moselle à Custines (Figure 9). Cette station se situe juste après la confluence entre la Moselle et la Meurthe et correspond à un bassin versant d'environ 6800 km². Trois stations incluses dans ce bassin ont été prises en compte afin d'effectuer un calage de manière similaire à ce qui est réalisé pour GRSD, c'est-à-dire un calage séquentiel allant des stations amont aux stations aval. Ces stations sont situées sur la Moselle à Epinal, la Mortagne à Gerbéviller et la Meurthe à Damelevières.

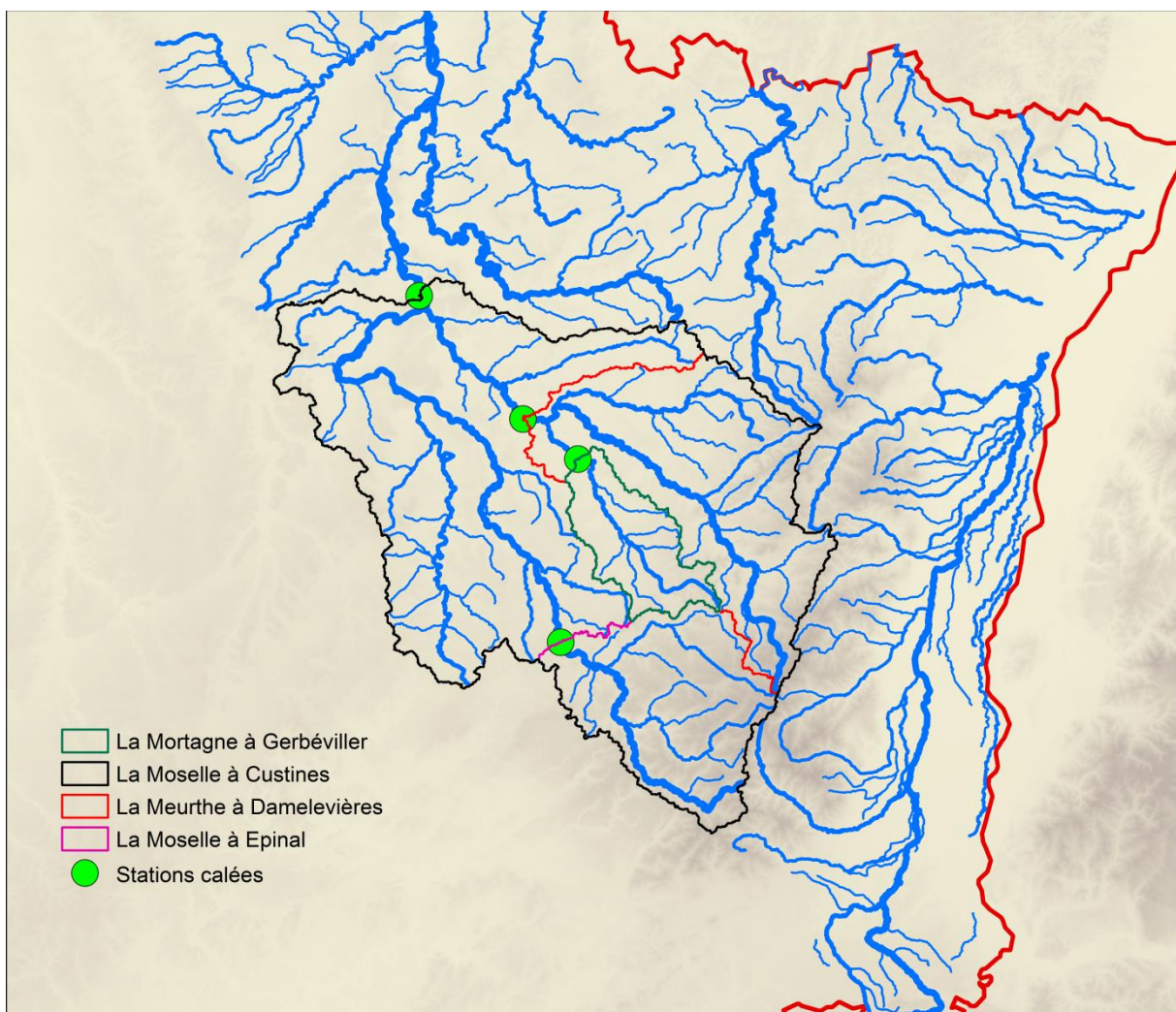


Figure 9 : Bassin versant de la Moselle à Custines. Le test de calage de LARSIM a été effectué sur ces quatre stations.

Afin de servir de point de référence permettant d'évaluer la pertinence de la procédure de calage automatique, les partenaires d'HYDRON ont fourni des jeux de paramètres calés par leurs soins ou

par un autre bureau d'étude sur ces quatre stations (voir Tableau 1 pour les principaux paramètres). Le jeu de paramètres fourni a été déterminé pendant un calage manuel avec d'autres données d'entrée et pour une période différente. Ces paramètres ont alors servi à la fois à effectuer une simulation de référence avec LARSIM, mais aussi à fournir des valeurs fixées pour les paramètres non calés par l'algorithme.

Tableau 1 : Valeurs des principaux paramètres de LARSIM tels qu'utilisés pour servir de référence (en gris, les paramètres qui seront calés automatiquement)

Station	EQB	EQI	EQD	EQD2	A2	EKM	EKL	EKR	BSF	beta	Dmin	Dmax	KG	TGr
Epinal	40000	12000	1200	300	1.5	0.8	0.8	0.8	0.3	0.012	0	2	1	0.5
Gerbeviller	30000	1600	200	30	1	0.5	0.5	0.5	0.3	0.011	0	0.5	1	1
Damelevières	70000	14000	800	100	2	0.7	0.7	0.7	0.2	0.012	0	1.5	1	1
Custines	50000	3500	1200	150	2.5	1.2	1.2	1.2	0.3	0.003	0	1.2	1	1

Les résultats de calages automatiques préliminaires sont présentés pour les quatre stations considérées en ce qui concerne les paramètres (Tableau 2) et des critères d'évaluation (Tableau 3). Nous pouvons observer que les paramètres optimaux sont assez proches des paramètres fournis par HYDRON en règle générale, et surtout qu'ils respectent des ordres de grandeurs habituellement utilisés. Par ailleurs, les performances en bilan, KGE' sur les forts débits et KGE' sur les faibles débits, des calages automatiques sont très proches voire meilleures que celles des calages manuels fournis par HYDRON.

Tableau 2 : Valeurs des sept paramètres à optimiser sur les quatre stations pour la référence, et pour des calages automatiques sur 1971-1975 et 1976-1980

Station	Période de calage	Paramétrisation	KG	beta	EQB	Dmax	EQI	EQD	EQD2
Epinal	-	Manuelle	1	0.012	40000	2	12000	1200	300
	1971-1975	Automatique	0.98	0.021	18697	1.9	13748	1645	674
	1976-1980		0.96	0.012	26825	3.3	12000	1200	541
Gerbeviller	-	Manuelle	1	0.011	30000	0.5	1600	200	30
	1971-1975	Automatique	0.9	0.011	17610	3.6	16770	1838	569
	1976-1980		0.9	0.011	22985	2.4	3715	670	74
Damelevières	-	Manuelle	1	0.012	70000	1.5	14000	800	100
	1971-1975	Automatique	0.96	0.0082	30112	1.5	14000	2823	41
	1976-1980		0.96	0.0039	17493	2.1	17403	1325	134
Custines	-	Manuelle	1	0.003	50000	1.2	3500	1200	150
	1971-1975	Automatique	0.96	0.0039	19985	0	3500	1987	150
	1976-1980		0.96	0.0039	87393	0.6	3500	1200	201

Tous ces éléments indiquent que le calage automatique de LARSIM fonctionne et pourra être utilisé pour l'exercice de calage sur des périodes différentes, tel qu'il est prévu dans le cadre de MOSARH (voir dans la suite de ce rapport).

Tableau 3 : Performances des différentes méthodes de calages sur les quatre stations pour la référence et pour des calages automatiques sur 1971-1975 et 1976-1980. Le bilan est optimal pour une valeur égale à 1 (au-dessus il y a surestimation, en-dessous sous-estimation), et le KGE' est lui aussi optimal lorsqu'il vaut 1, plus sa valeur est faible moins la performance est bonne.

Station	Paramétrisation	Période de calage	Période de validation	Bilan	KGE' Q	KGE' log Q
Epinal	Manuelle	-	1981-1990	1.077	0.892	0.936
			1991-2000	1.131	0.827	0.926
	Automatique	1971-1975	1981-1990	1.059	0.856	0.905
			1991-2000	1.116	0.870	0.893
		1976-1980	1981-1990	1.020	0.902	0.951
			1991-2000	1.071	0.913	0.943
Gerbeviller	Manuelle	-	1981-1990	1,471	0,029	0,759
			1991-2000	1,423	0,200	0,706
	Automatique	1971-1975	1981-1990	1,253	0,668	0,787
			1991-2000	1,224	0,691	0,727
		1976-1980	1981-1990	1,220	0,703	0,808
			1991-2000	1,188	0,726	0,769
Damelevières	Manuelle	-	1981-1990	1,213	0,700	0,838
			1991-2000	1,233	0,689	0,866
	Automatique	1971-1975	1981-1990	1,144	0,499	0,566
			1991-2000	1,191	0,530	0,633
		1976-1980	1981-1990	1,106	0,744	0,876
			1991-2000	1,131	0,791	0,902
Custines	Manuelle	-	1981-1990	1,063	0,908	0,941
			1991-2000	1,104	0,861	0,946
	Automatique	1971-1975	1981-1990	1,007	0,740	0,815
			1991-2000	1,059	0,808	0,876
		1976-1980	1981-1990	0,992	0,817	0,942
			1991-2000	1,038	0,886	0,937

3.2. Réduction des jeux de données en entrée du modèle

L'exécution de LARSIM étant particulièrement lente lorsque l'on travaille sur de grands bassins en raison de la lecture des données d'entrée, il a été décidé de tester une méthode de réduction du nombre de pixels sur les cartes de données. En effet, les données HYRAS de précipitations sont à la résolution de 1 km et si l'on fournit à LARSIM des données à une résolution plus faible, celui-ci les interpole, ce qui est plus rapide en termes de temps de calcul. Par conséquent, une méthode de sélection aléatoire de données d'entrées a été testée. Celle-ci consiste en la conservation de seulement 10% des données de précipitations, et 5% des autres données (Figure 10). Cette méthode repose sur l'hypothèse que l'hétérogénéité spatiale des précipitations, et a fortiori des autres données, n'est pas assez importante pour nécessiter forcément une résolution si élevée. On peut voir sur la Figure 10 que la quantité de données reste raisonnable et couvre correctement la zone d'étude.

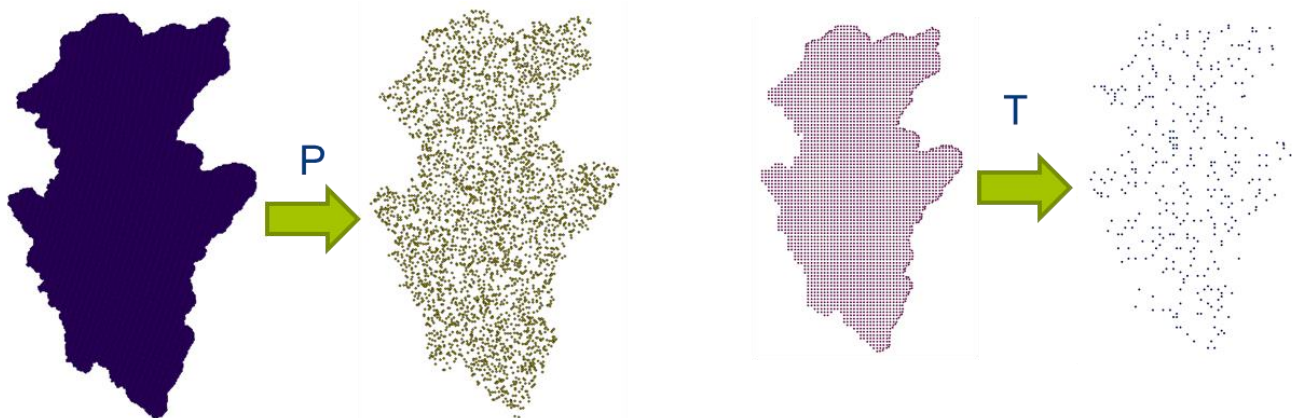


Figure 10 : Impact de la réduction des jeux de données sur les précipitations et les températures

Tableau 4 : Paramètres optimisés de manière automatique et performances sur la période 1981-1990 de LARSIM pour la Moselle à Epinal en utilisant l'ensemble des données ("Tout") ou le jeu de données réduit comme sur la Figure 10 ("Réduit")

Période de calage	Données	KG	beta	EQB	Dmax	EQI	EQD	EQD2	Bilan	KGE' Q	KGE' log Q
1971-1975	Tout	0.98	0.021	18697	1.9	13748	1645	674	1.06	0.86	0.91
1971-1975	Réduit	0.98	0.021	19985	1.9	13243	1645	674	1.05	0.85	0.90
1976-1980	Tout	0.96	0.012	26825	3.3	12000	1200	541	1.02	0.90	0.95
1976-1980	Réduit	0.97	0.012	26825	3	12000	1200	541	1.03	0.91	0.95

Le Tableau 4 montre les paramètres et performances obtenus lors de calages automatiques de LARSIM pour la station de la Moselle à Epinal lorsque l'on utilise l'ensemble des données (i.e. valeurs issues du Tableau 2 et du Tableau 3) et les jeux de données réduits. Ces résultats montrent deux choses : les paramètres obtenus avec le jeu réduit sont très proches des paramètres avec le jeu complet, et les scores d'évaluation sont aussi très proches. Cela montre qu'il est possible de réduire les jeux de données sans impacter les calages et les simulations. Par ailleurs, il est à noter que le temps de calcul nécessaire aux calages est ainsi réduit de 75%. Pour les calages effectivement réalisés et utilisés dans le cadre de MOSARH21, nous réduirons donc les jeux de données utilisés.

3.3. Vérification du bon fonctionnement du calage automatique de LARSIM

Comme nous l'avons détaillé précédemment, le calage automatique de LARSIM s'effectue en quatre étapes, en optimisant (préfiltrage + recherche locale) successivement :

- le paramètre lié au bilan,
- puis les deux liés au débit de base,

- puis les deux liés au débit intermédiaire,
- et enfin les deux liés aux pics de crue.

Chacune de ces étapes comporte une fonction-objectif qui lui est propre (voir sections précédentes). Ensuite, on répète une fois l'étape de recherche locale, afin de réajuster les composantes du débit qui auraient pu être trop détériorées par le calage de paramètres liés aux autres composantes. Par conséquent, nous allons vérifier que l'optimisation d'une composante du débit n'est pas trop néfaste pour les autres composantes. Cela sera effectué par le biais de l'analyse de l'évolution des scores et des paramètres lors du processus de recherche locale (donc préfiltrage exclu) uniquement.

La Figure 11 présente les valeurs obtenues pour différents critères lors des itérations successives de l'algorithme de calage. On appelle itération le test d'un nouveau jeu de paramètres, ce nouveau jeu de paramètre ne comprenant en général qu'un paramètre différent du meilleur jeu retenu à l'itération précédente. Certains jeux de paramètres ne donnent pas lieu à une amélioration du critère considéré et n'ont donc pas été gardés pour la suite des itérations, même s'ils sont représentés dans les graphes suivants.

Nous illustrons le calage avec la station de la Bruche à Holtzheim (676 km²) pour un calage et une évaluation sur la période 1971-2000. Les quatre critères présentés correspondent aux quatre composantes du débit utilisées pour le calage. On rappelle que le KGE' (version du critère KGE prenant le ratio des coefficients de variation plutôt que celui des écarts-types) est optimal pour une valeur égale à 1 et ne peut pas aller au-delà. Le bilan (balance) est lui aussi optimal pour une valeur égale à 1, mais une valeur supérieure à 1 indique une surestimation des débits, et une valeur inférieure à 1 indique une sous-estimation des débits.

Sur la Figure 11, nous pouvons observer une augmentation générale du KGE (*i.e.* une amélioration de la performance du modèle) pour les trois fonctions objectif, ainsi qu'un bilan qui se rapproche de 1. Ces résultats montrent un fonctionnement correspondant à ce qui est attendu du calage automatique. Notons toutefois que des valeurs de KGE' très légèrement supérieures à celles conservées au final ont pu être observées au cours du calage. Elles n'ont pas été conservées car elles correspondaient alors à des valeurs plus faibles pour l'un des autres critères. Ceci montre bien le compromis qui est recherché par le calage.

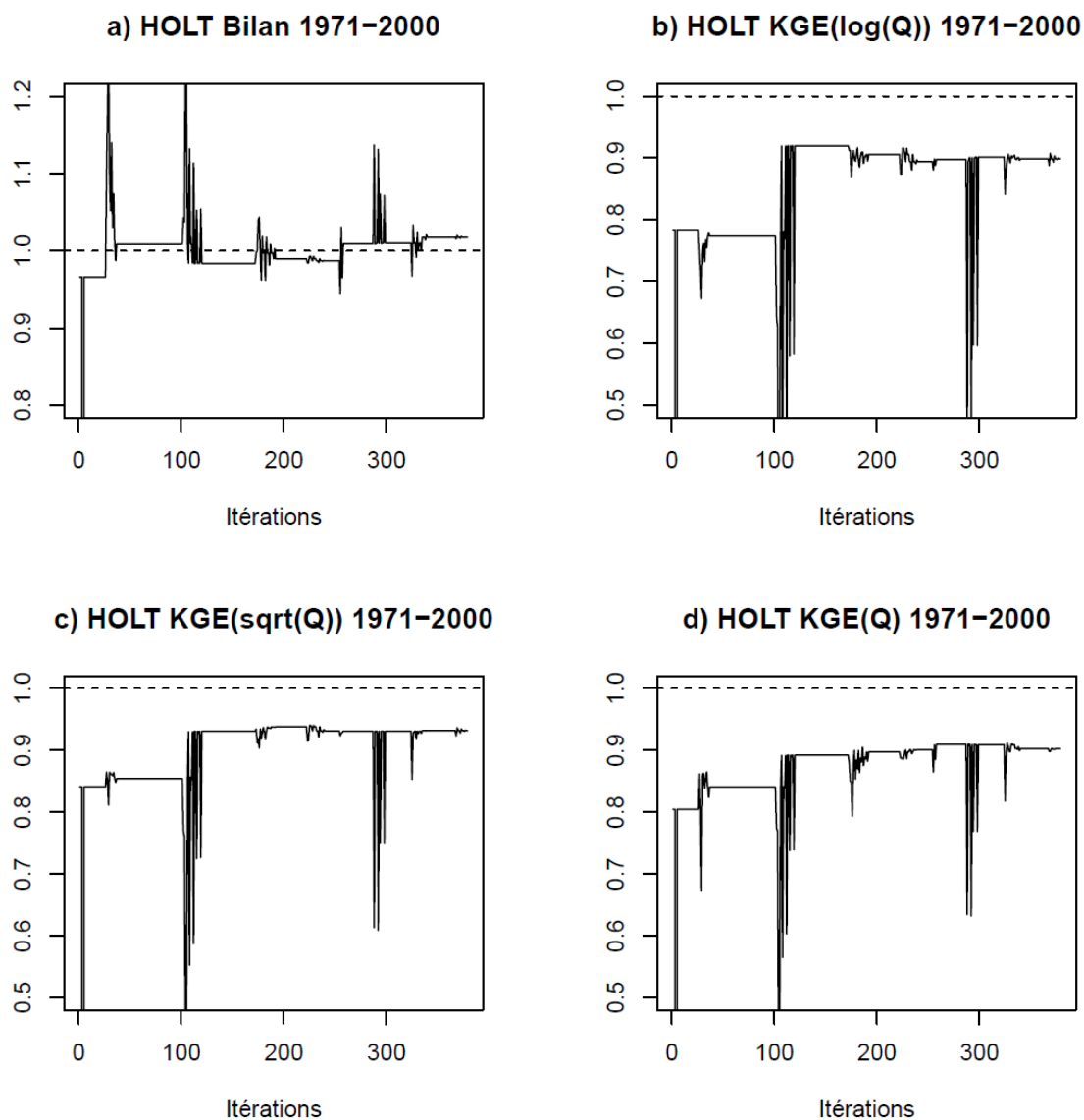


Figure 11 : Evolution des performances de LARSIM à chaque itération successive pour le bilan (a), pour le débit de base (b), pour le débit intermédiaire (c) et pour les hauts débits(d). Exemple de la station de la Bruche à Holtzheim pour un calage sur la période 1971-2000. Les valeurs des critères d'optimisation sont celles calculées sur la période de calage (1971-2000).

La Figure 12 présente des résultats similaires pour un calage et une évaluation effectués sur la période 1971-1985 pour la même station. Les conditions hydroclimatiques étant différentes, les valeurs des critères d'optimisation sont différentes, et évoluent de manière un peu différente.

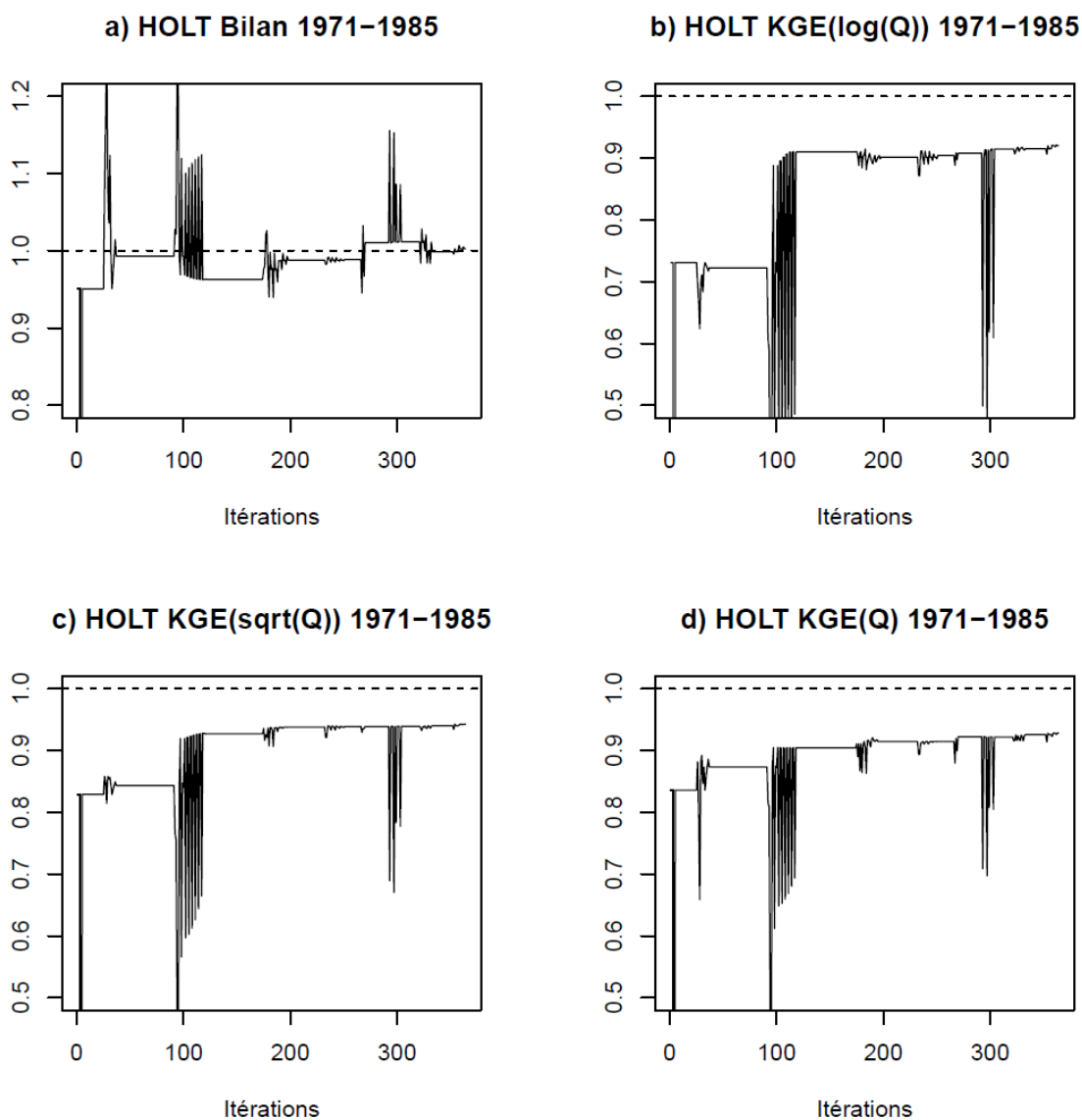


Figure 12 : Idem que la Figure 11, pour un calage et des critères d'optimisation calculés sur la période 1971-1985.

Enfin, la Figure 13 s'intéresse quant à elle à un calage sur 1986-2000. Dans ce cas, on voit que l'algorithme plafonne assez vite après quelques itérations à des valeurs de critère inférieures à celles obtenues sur les deux autres périodes, ce qui montre que le modèle est moins performant sur cette sous-période. On notera également le léger biais systématique sur le critère de bilan, le modèle surestimant systématiquement le débit, ce qui explique la moindre performance en termes de KGE' (Figure 12d). On peut expliquer ce problème de bilan par une surestimation des précipitations, une surestimation des débits, ou bien une gamme de variation du paramètre KG trop restreinte (on a en effet fixé la borne inférieure à 0.9).

Le comportement de l'algorithme de calage semble montrer l'intérêt d'utiliser une archive de données la plus longue possible. Cela montre aussi que les calages sur les deux sous-périodes fonctionnent de manière plutôt satisfaisante. Le comportement observé sur la période complète représente le comportement idéal de la méthode de calage (l'ensemble des critères s'améliorent

progressivement et sont proches de 1) et est observé sur de nombreuses stations. Le comportement montré sur les deux sous-périodes (l'une où le calage est très efficace et l'autre où il l'est moins) est observé sur différentes stations. Aucun problème majeur de calage n'a été observé cependant.

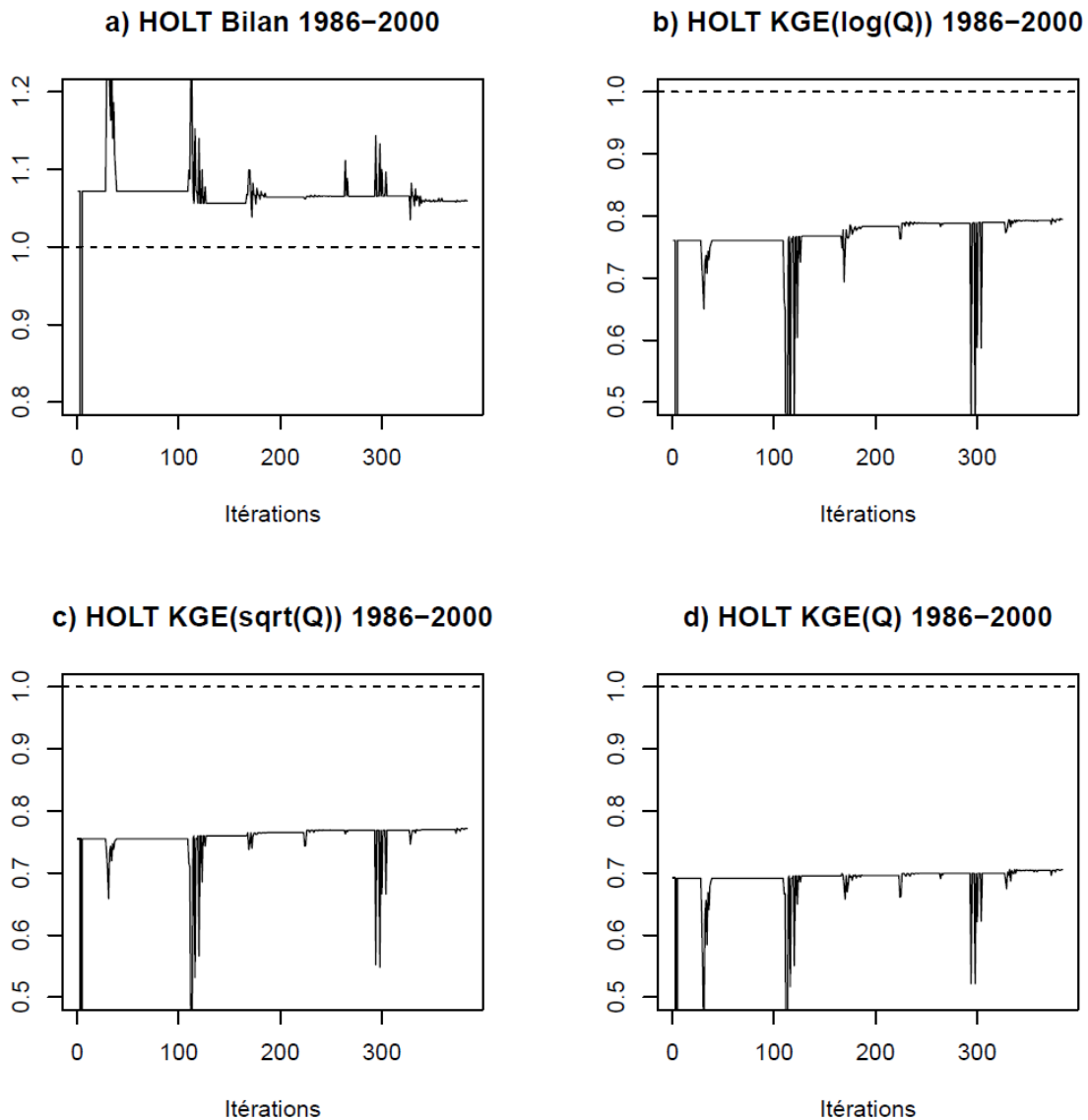


Figure 13 : Idem que la Figure 11, pour un calage et des critères d'optimisation sur la période 1986-2000.

Sur la Figure 14, nous montrons les résultats du calage sur les trois périodes pour la Moselle à Perl (11 556 km²). Il s'agit, dans notre projet, de l'exutoire de la Moselle. Ce bassin versant est modélisé par un modèle LARSIM qui utilise des débits (simulés) de plusieurs stations situées en amont, pour lesquelles les paramètres des bassins versants respectifs sont calés (voir Figure 1). Ainsi, nous pouvons vérifier le bon fonctionnement du calage sur un bassin plus complexe que la Bruche à Holtzheim, et dont les débits dépendent fortement de calages indépendants réalisés en amont. On observe une amélioration générale des critères au cours des itérations et de très bonnes valeurs de KGE', ce qui est très satisfaisant. On note également un léger biais systématique sur le critère de

bilan, le modèle surestimant systématiquement le débit (Figure 14d). Le calage évolue de manière très similaire sur les trois périodes de calage (voir Figure 15).

Nous ne montrerons pas dans ce rapport les graphiques pour l'ensemble des stations et des calages. Les résultats sont globalement bons, avec des critères KGE élevés et des bilans généralement proches de 1. Cependant, sur certaines stations, les performances sur un score donné (bilan par exemple) sont plus mauvaises. Ces moins bonnes performances (qui restent toutefois acceptables) ne peuvent pas être expliquées *a priori* par l'algorithme en lui-même, car les itérations successives donnent des améliorations. L'explication la plus plausible réside dans la non-représentation d'un processus hydrologique important dans le modèle, ou bien dans la présence d'un biais dans les données hydrologiques ou météorologiques.

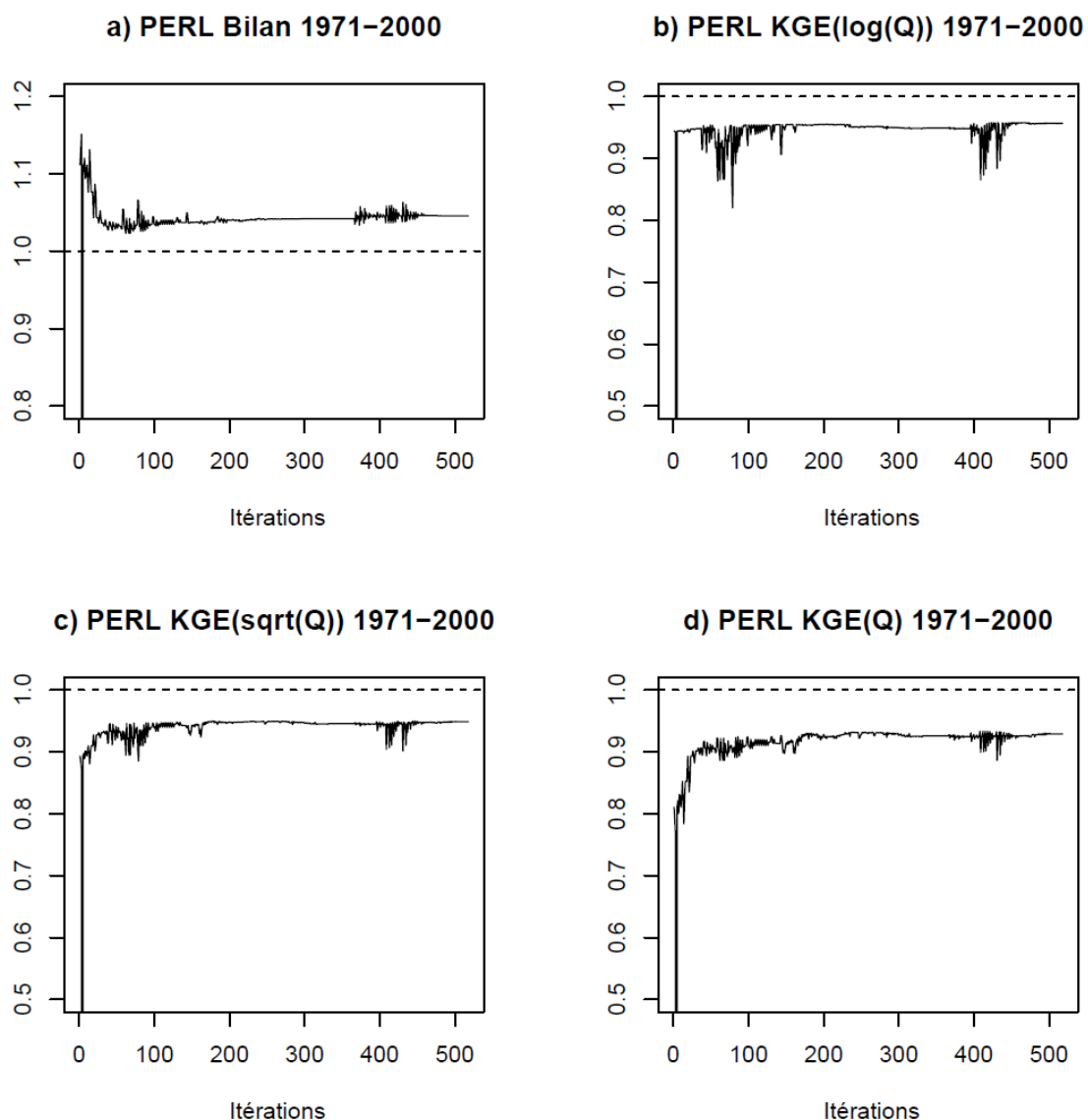


Figure 14 : Evolution des performances de LARSIM à chaque itération successive pour le bilan (a), pour le débit de base (b), pour les débits intermédiaires (c) et pour les hauts débits (d). Exemple de la station de la Moselle à Perl (DE) pour un calage sur la période 1971-2000. Les valeurs des critères d'optimisation sont celles calculées sur la période de calage (1971-2000).

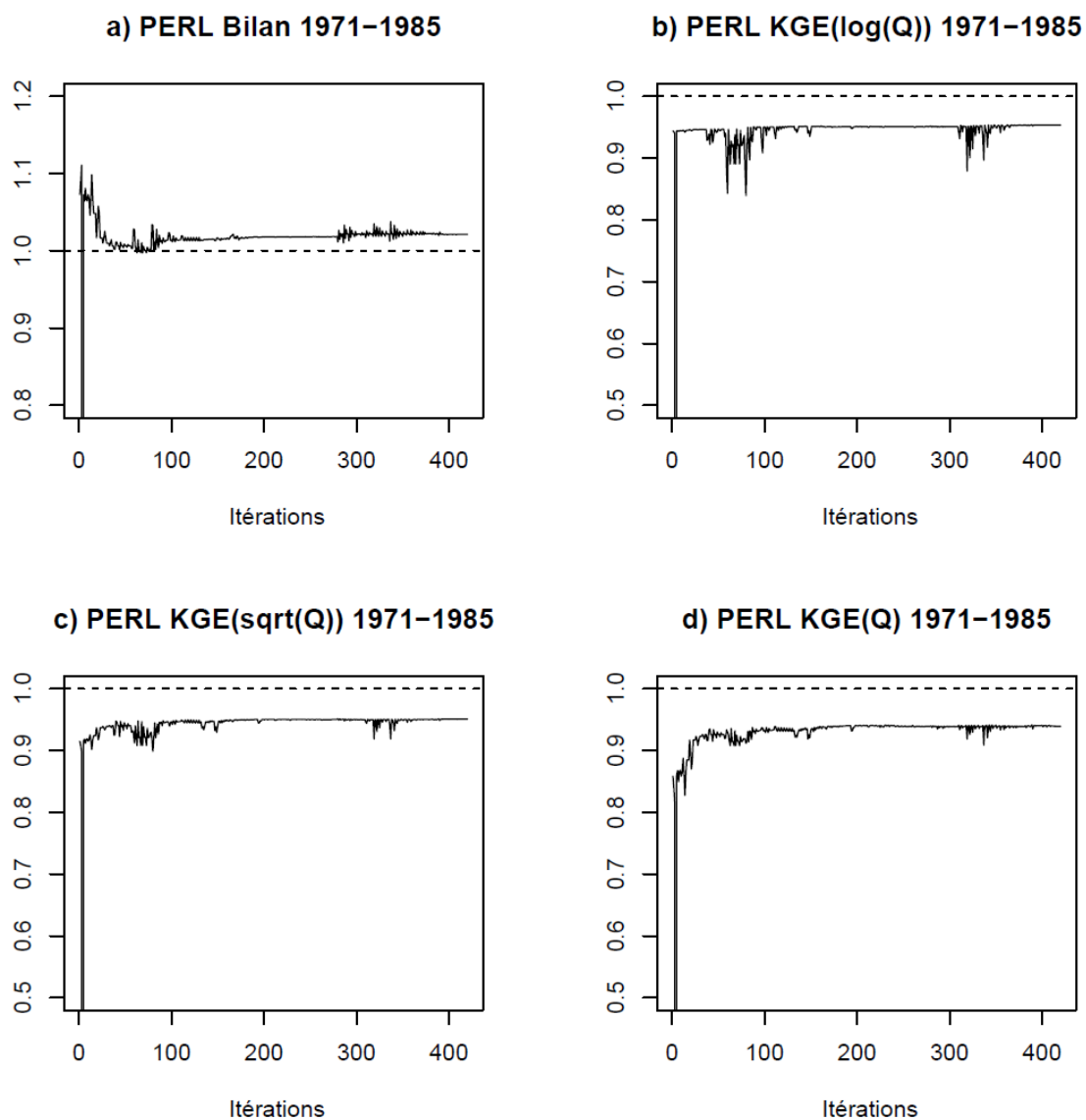


Figure 15 : Idem que la Figure 14, pour un calage et des critères d'optimisation sur la période 1971-1985.

Nous montrons l'un des graphes décrivant l'évolution des valeurs des sept paramètres de LARSIM calés dans le cadre de MOSARH21 sur la Figure 16 à titre d'illustration. Cependant, nous n'analyserons pas l'ensemble de ces graphiques ici. L'évolution des paramètres présente une certaine stabilité, ce qui est rassurant sur le fait que l'algorithme ne diverge pas vers des valeurs aberrantes.

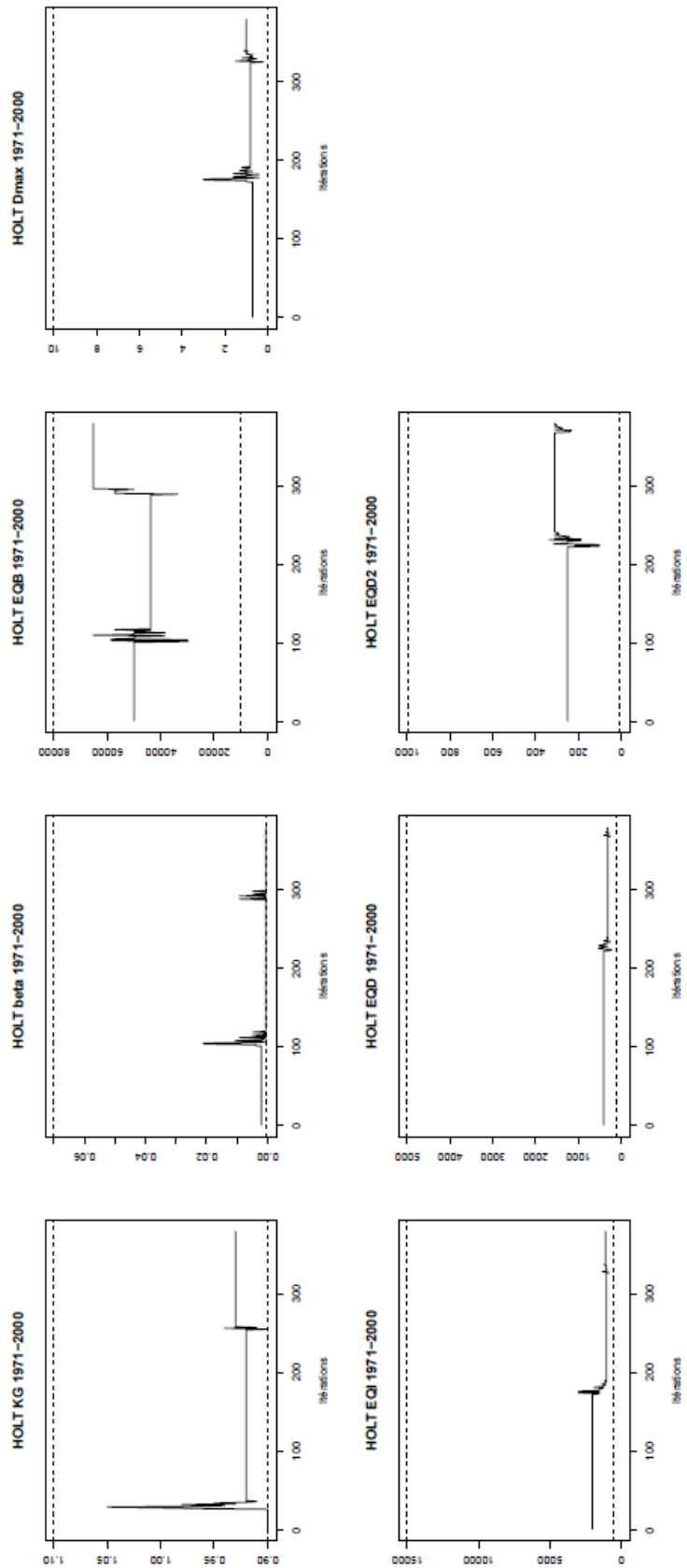


Figure 16 : Evolution des paramètres de LARSIM à chaque itération successive. Exemple de la station de la Bruche à Holtzheim pour un calage sur la période 1971-2000.

3.4. Cartes des paramètres des modèles LARSIM et GRSD

Des cartes de paramètres des deux modèles hydrologiques ont été produites.

Il n'est pas possible de présenter une analyse détaillée de l'ensemble des paramètres des deux modèles dans ce rapport. Des exemples de ces cartes sont présentés sur la Figure 17 pour des calages réalisés sur la période complète. Nous rappelons ici que, si GRSD est calé sur chacun des sous-bassins versants de manière séquentielle, LARSIM ne l'est pas : les paramètres de certains sous-bassins versants sont fixés, avec pour certaines stations un maillage différent de celui de GRSD ; le calage dispose donc d'un levier moins important pour ajuster les débits qui sont parfois influencés par des débits en amont. Les résultats qui suivent n'ont donc pas de but comparatif strict, ils doivent être vus comme une vérification supplémentaire du bon fonctionnement des calages des deux modèles ainsi que comme une tentative d'analyse de potentielles dépendances régionales. L'incertitude de modélisation étant prise en compte par l'utilisation de ces deux modèles hydrologiques dans MOSARH21, nous ne cherchons pas à identifier le meilleur modèle !

La Figure 17 permet de noter des différences entre les paramètres EQB (constante de rétention du réservoir souterrain) de LARSIM en calage manuel et ceux en calage automatique. Par ailleurs, on peut observer une structure du paramètre X1 (capacité du réservoir de production) de GRSD sur la nappe d'Alsace, celui-ci étant plus élevé ce qui reflète une plus grande capacité de stockage de l'eau dans le sol et une dynamique plus lissée de la relation pluie-débit. Pour les bassins apparaissant en gris, des paramètres EQB existent et ont été fournis par HYDRON. Ils ont été utilisés mais ne sont pas représentés ici car on ne veut montrer que les paramètres obtenus par le calage automatique. De plus, la discrétisation spatiale des paramètres de LARSIM peut différer du découpage des 70 bassins versants, mais on ne les montre pas ici car nous ne disposons pas des contours de tous ces bassins.

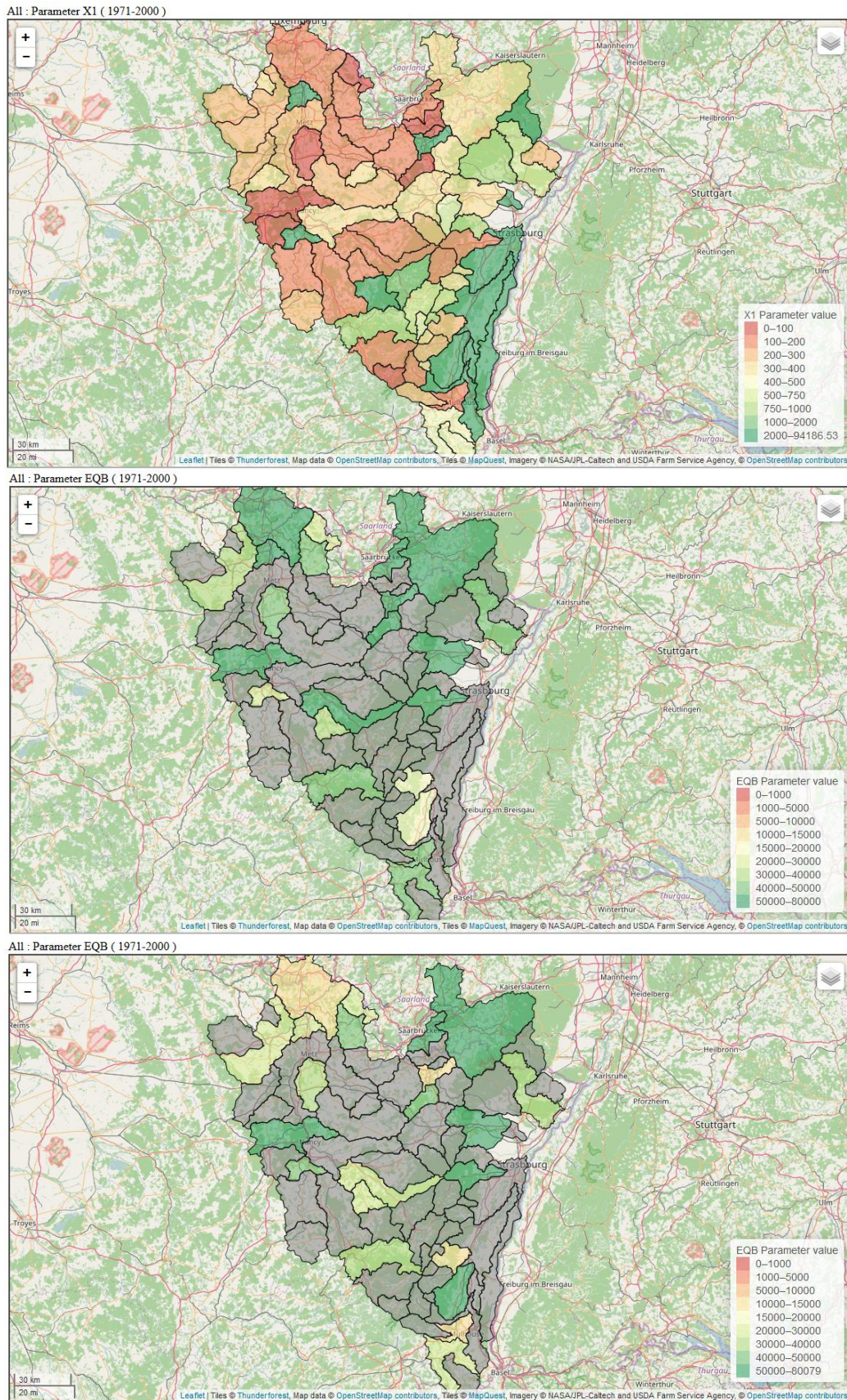


Figure 17 : Cartes des paramètres X1 (pour GRSD, en haut) et EQB (pour LARSIM calage manuel au milieu, LARSIM calage automatique en bas). Il faut noter que le maillage des paramètres de LARSIM ne correspond pas entièrement à celui de GRSD : certains sous-bassins autres que les 70 du projet comprennent des paramètres fixés ; cela laisse moins de flexibilité au calage automatique de LARSIM pour bien simuler les débits aux stations calées.

3.5.Capacité des modèles à reproduire les hydrogrammes

Nous n'avons pas effectué une analyse détaillée de chacun des hydrogrammes pour toutes les stations. Ils semblent généralement bons, les différences entre les modèles et avec les observations de débit restant raisonnables.

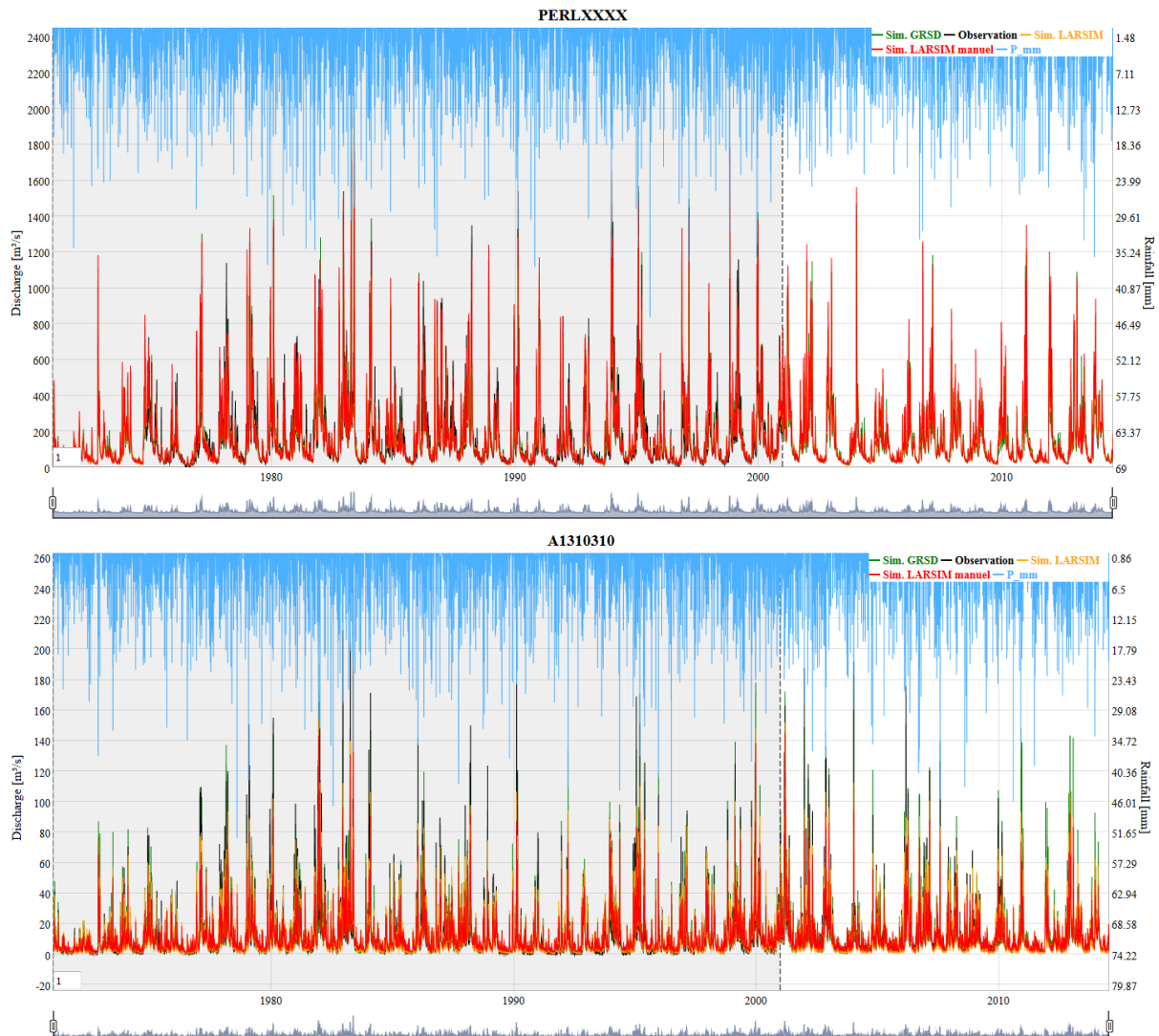


Figure 18 : Hydrogrammes de la Moselle à Perl (en haut) et de l'III à Ensisheim (en bas). En vert, les simulations de GRSD, en orange les simulations de LARSIM calé en automatique, en rouge les simulations de LARSIM calé en manuel, en noir les débits observés, en bleu les précipitations. Les deux modèles sont calés sur la période complète (1971-2000).

Nous pouvons voir sur la Figure 18 des exemples d'hydrogrammes pour la période 1971-2013. Sur cette figure, la période de calage (1971-2000) apparaît en grisé. Nous présentons ainsi sur la Figure 19 un zoom sur la sécheresse de 1976 pour la Moselle à Perl et sur des crues de l'hiver 2007-2008 pour l'III à Ensisheim (A1310310). Nous pouvons observer un bon comportement de GRSD et LARSIM lors de cette sécheresse exceptionnelle de 1976. En effet, même si toutes les variations ne sont pas reproduites par les modèles, les tendances sont bien représentées et l'intensité de l'étiage semble fidèlement modélisée. En ce qui concerne les crues de l'hiver 2007-2008, nous observons tout d'abord que, selon le pic de crue, le modèle le plus performant peut différer. Cependant, on peut noter une sous-estimation globale des pics de crue de la part du modèle LARSIM calé manuellement. En revanche, les deux modèles calés automatiquement semblent tour-à-tour surestimer ou sous-

estimer les différents pics de crue. Par ailleurs, le modèle LARSIM calé manuellement est le modèle qui semble représenter le mieux les périodes de décrue. Il est important de noter que ces conclusions sur des cas particuliers ne sont pas généralisables à l'ensemble des hydrogrammes, et que l'utilisation de métriques d'erreur est nécessaire pour cela.

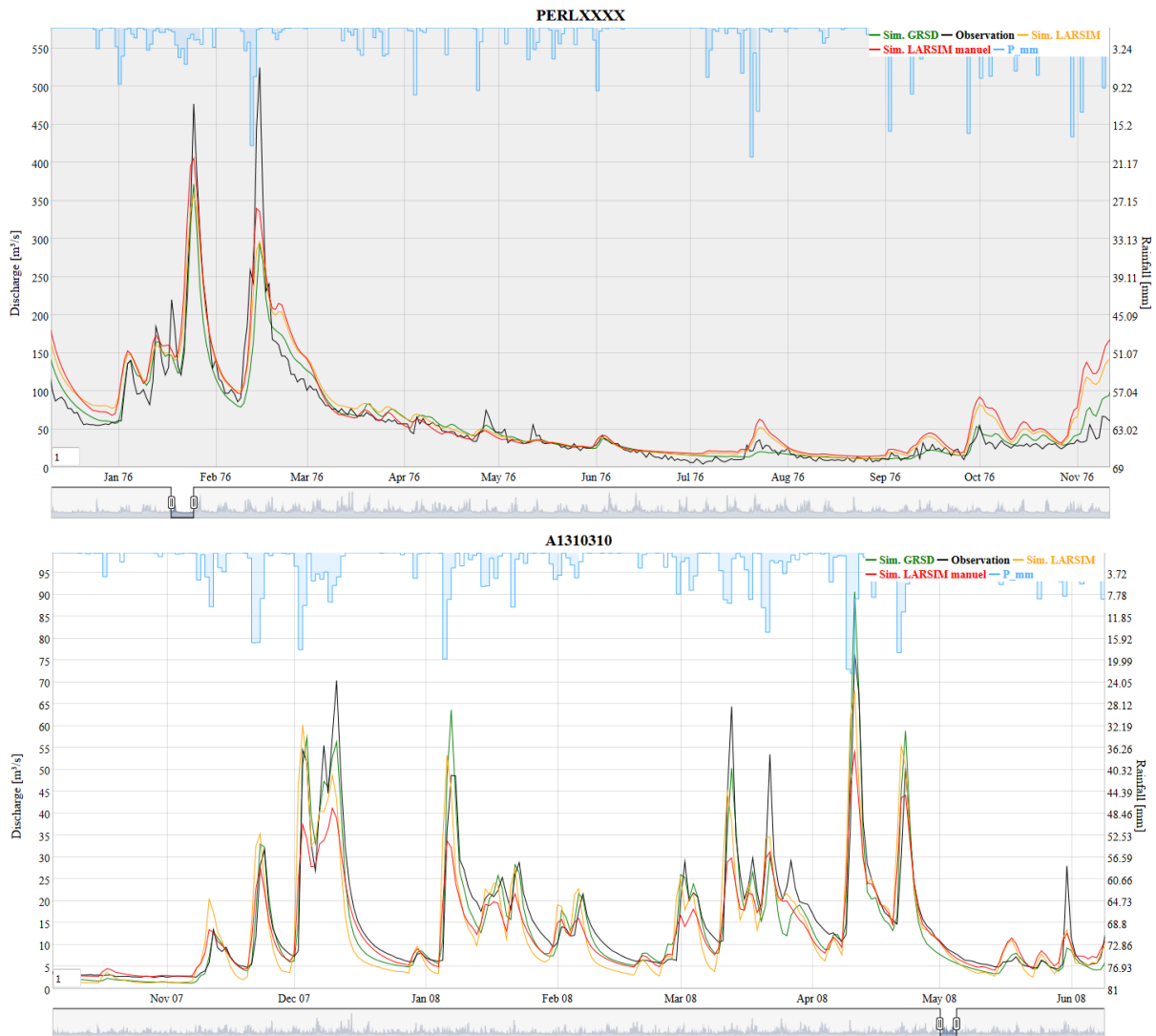


Figure 19 : Hydrogrammes de la Moselle à Perl (en haut) pendant la sécheresse de 1976 et de l'III à Ensisheim (en bas) pendant une période de crues de l'hiver 2007-2008. Ces courbes sont extraites de celles de la Figure 18.

3.6. Capacité des modèles à reproduire les courbes de régime

La Figure 20 montre que les régimes de la Moselle à Perl et de l'III à Ensisheim sont généralement bien reproduits, notamment en ce qui concerne la Moselle. En effet, excepté en novembre et décembre où le modèle LARSIM calé manuellement, et dans une moindre mesure le modèle LARSIM calé automatiquement divergent, tous les modèles sont très proches des observations. Les résultats sont plus disparates pour l'III à Ensisheim : la période d'étiage est assez mal reproduite par LARSIM et GRSD montre une légère surestimation. On peut suspecter pour cette période une influence ou erreur de mesure sur les débits observés, le paramètre d'ajustement KG de LARSIM étant optimisé à

une valeur de 0.9 (on retire 10% de la précipitation), sa valeur la plus basse permise par l'algorithme. Sur la période des hautes eaux, les modèles se comportent mieux.

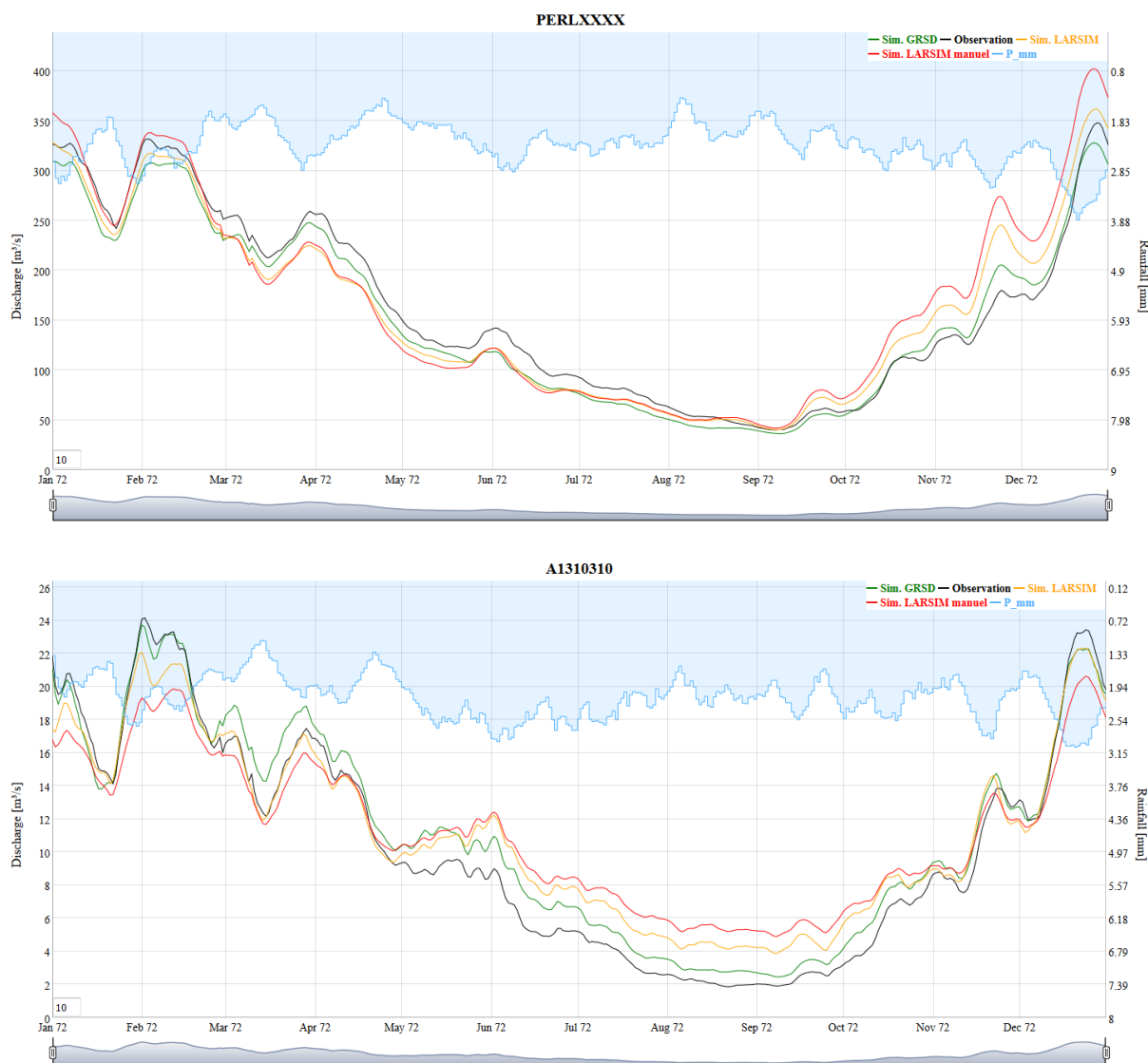


Figure 20 : Courbes de régimes journaliers interannuels pour la Moselle à Perl (en haut) et l'III à Ensisheim (en bas) sur la période 1971-2000.

3.7. Cartographie des métriques d'erreur

De la même manière que pour les paramètres, des cartes de performances ont été produites. Neuf critères différents sont disponibles (Tableau 5).

Tableau 5 : Liste des métriques utilisés.

Nom du critère	Acronyme	Composante du débit analysée
Nash-Sutcliffe	NSE	Forts débits
Coefficient de corrélation linéaire de Pearson	r	Dynamique
Efficacité volumétrique	VE	Temporalité
Kling-Gupta efficiency	KGE	Forts débits (corrélation, volume et écart-type)

KGE sur l'inverse des débits	KGE_INV	Faibles débits
Moyenne entre le KGE et le KGE sur l'inverse des débits	KGE_DOUBLE	Ensemble de la gamme des débits
Bilan	Bilan	Biais relatif en termes de débits moyens
Biais relatif en termes de quantile 5 %	Q05_ratio	Faibles débits
Biais relatif en termes de quantile 95 %	Q95_ratio	Forts débits

Les scores sont présentés pour cinq cas :

- GRSD en période de calage
- GRSD en période de contrôle
- LARSIM en période de calage
- LARSIM en période de contrôle
- LARSIM calé manuellement.

Cela représente donc trente cartes par période de calage, sauf pour le calage sur la période complète qui ne comprend pas de contrôle et donc ne comporte que 18 cartes.

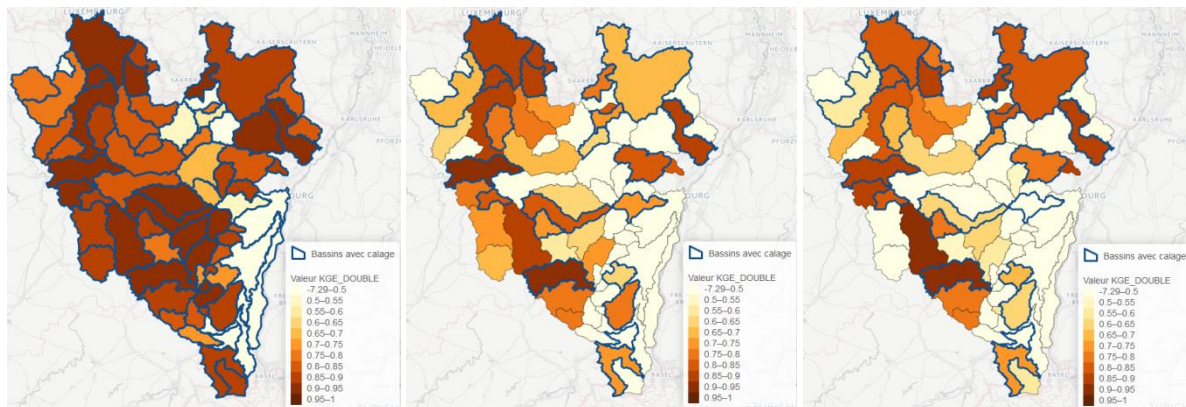


Figure 21 : Valeurs de KGE_DOUBLE (hauts et bas débits) calculées sur la période 1971-2000 pour les modèles calés sur la période 1971-2000 (à gauche GRSD, au centre LARSIM calage automatique, à droite LARSIM calage manuel). Un ton foncé indique les meilleures performances, un ton clair indique de mauvaises performances. Les bassins calés sont entourés de bleu.

La Figure 21 présente le KGE_DOUBLE pour les modèles calés sur la période complète. Nous remarquons sur cette figure la difficulté de l'ensemble des modèles à bien représenter les débits sur la plaine d'Alsace. Cette difficulté est en partie due à la présence d'influences anthropiques dans les débits observés utilisés à des fins d'évaluation. Il est important de noter que les paramètres de ces sous-bassins ne sont pas issus d'un calage, mais sont repris de sous-bassins situés en amont, la mauvaise qualité des débits observés ayant été repérée en amont des travaux de calage. Par ailleurs, nous observons la meilleure performance générale de GRSD par rapport à LARSIM calé automatiquement, qui lui-même donne de meilleures performances que LARSIM calé manuellement. Cette hiérarchie peut s'expliquer par le plus grand nombre de stations utilisées par GRSD pour le calage (voir Figure 17). Par ailleurs, la recherche de paramètres optimaux pour LARSIM de manière plus systématique (par le calage automatique) montre ici encore son intérêt.

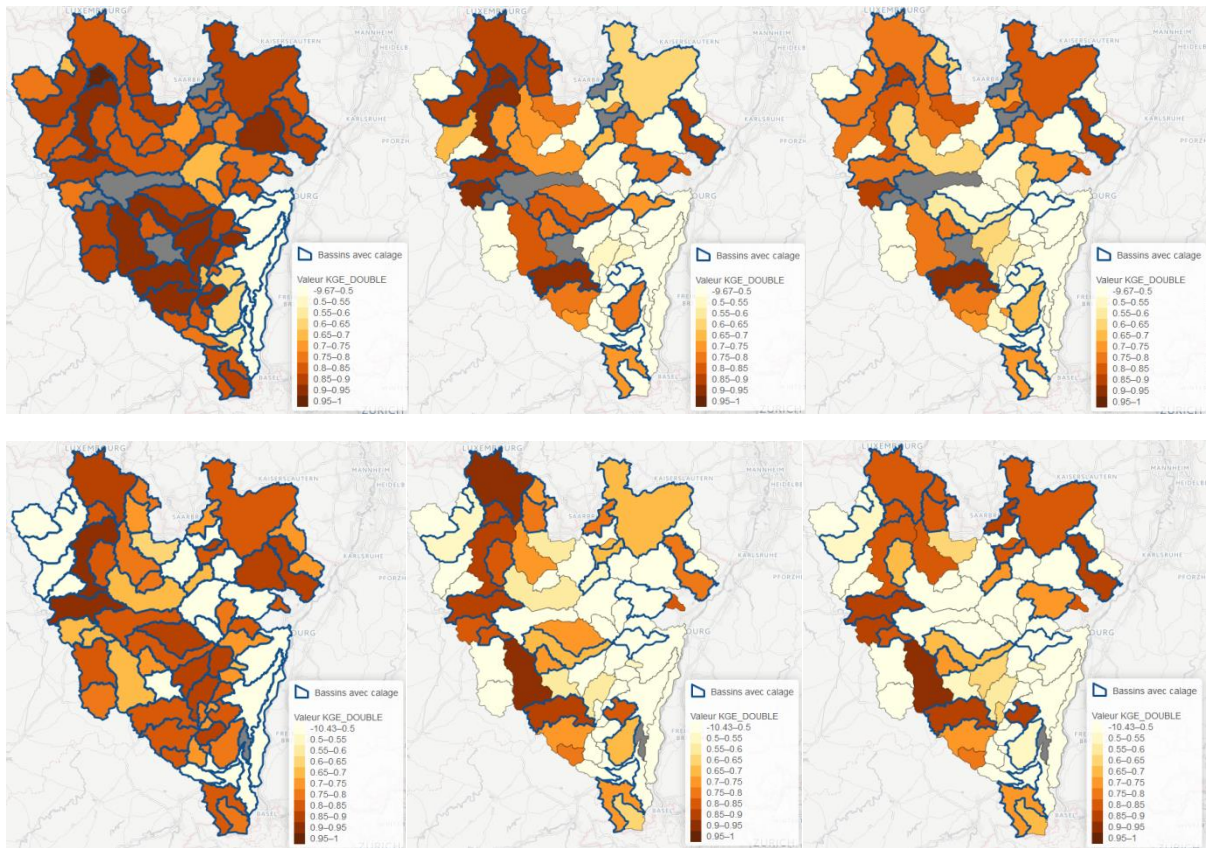


Figure 22 : Valeurs de KGE_DOUBLE calculées sur la période 1971-1985 (période de calage, en haut) et 1986-2000 (période de contrôle, en bas) pour les modèles calés sur la période 1971-1985 (à gauche GRSD, au centre LARSIM calage automatique, à droite LARSIM calage manuel). Un ton foncé indique les meilleures performances, un ton clair indique la pire performance. Les sous-bassins grisés correspondent à des sous-bassins sans débits observés disponibles pour la période d'intérêt. Pour le modèle LARSIM calé manuellement, il ne s'agit pas à proprement parler de calage et de contrôle.

La Figure 22 montre les résultats en calage et en contrôle des modèles (sauf pour LARSIM calé manuellement). On observe une dégradation logique des performances en contrôle par rapport à la période de calage, mais celle-ci reste raisonnable. On observe aussi que la dégradation du modèle LARSIM calé automatiquement semble plus faible que celle de GRSD. Cela peut être expliqué par une meilleure robustesse ou alors par le fait que les performances partent de plus bas.

De nombreux scores ont été produits lors de cette analyse des modèles. Nous nous contenterons de présenter en plus du KGE_DOUBLE ci-dessus, des cartes de bilan, de quantiles 95 % (hauts débits) et de quantiles 5 % (bas débits). Ces trois cartes seront affichées en pourcentages de biais, ce qui fait que par exemple une valeur de 9.5 signifie que le critère étudié est de 9.5% plus élevé lorsqu'il est simulé par le modèle que lorsqu'il est observé (en d'autres termes, il y a une surestimation de 9.5% par le modèle).

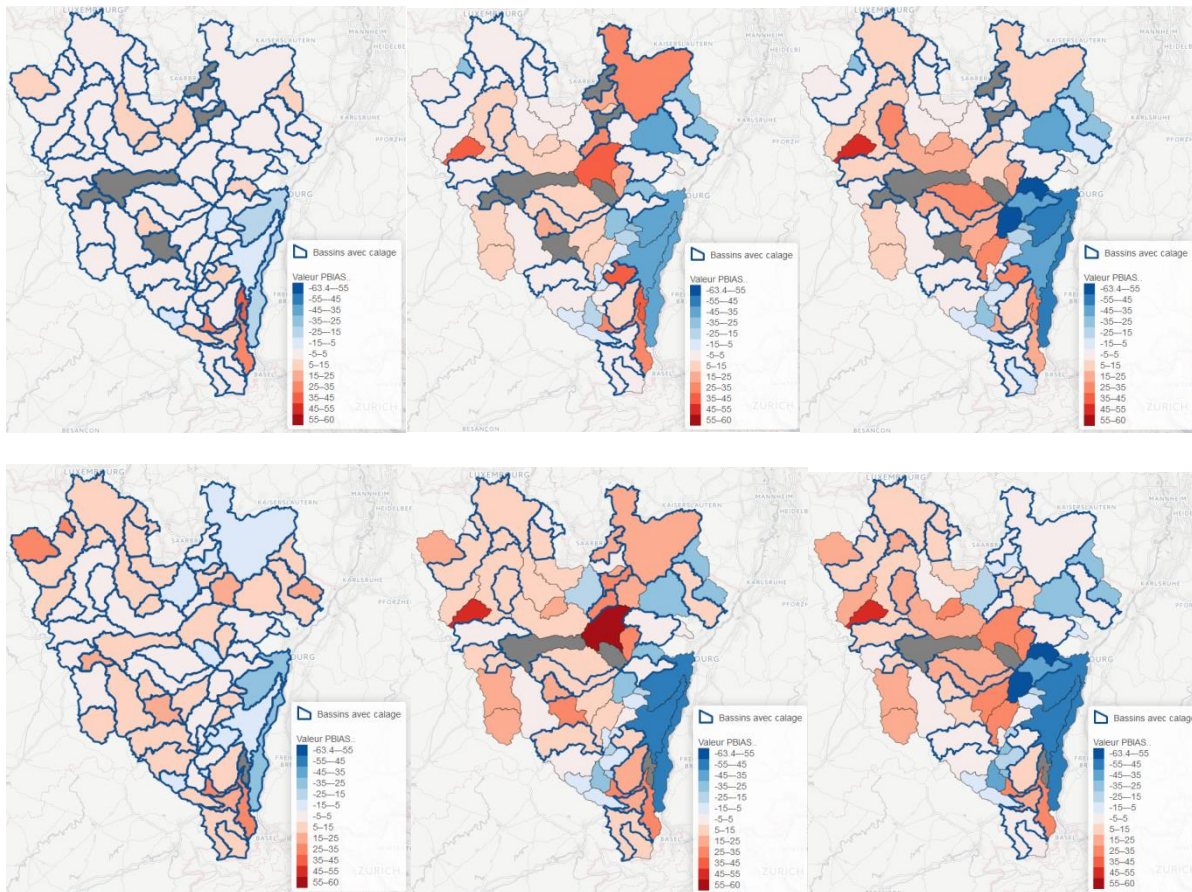


Figure 23 : Idem que Figure 22 mais pour le bilan (biais du débit moyen).

Les cartes de bilan (Figure 23) montrent de moindres biais pour GRSD que pour LARSIM en calage. On observe aussi que les bassins calés de LARSIM se comportent mieux en règle générale que les bassins non calés, ce qui est attendu. Une des raisons qui explique ceci réside dans le calage des modèles : LARSIM comprend un coefficient KG qui ajuste la quantité de pluie et qui a été limité entre 0.9 et 1.1 (c'est-à-dire un ajustement de plus ou moins 10%) afin de reproduire ce qui est habituellement fait par les modélisateurs d'HYDRON. En revanche, GRSD comprend un terme d'échange souterrain qui n'est pas limité ; cela laisse donc plus de marge de manœuvre au calage pour obtenir de meilleurs bilans, sachant que le bilan est l'une des composante du KGE' utilisé pour l'optimisation des paramètres. Par ailleurs, il faut noter que la comparaison des performances entre les deux modèles ne peut pas être réalisée de manière complètement équitable : certains paramètres de LARSIM, jugés peu sensibles, sont fixés lors du calage automatique, et les calages manuels ainsi que certains calages automatiques de LARSIM ont été réalisés sur des maillages différents de ceux de GRSD, ce qui peut induire moins de flexibilité pour déterminer l'optimum des paramètres. En résumé, si la performance pure de LARSIM dans ce projet est moins bonne que celle de GRSD, cela est impacté par la méthodologie mise en œuvre et n'a que peu d'influence sur la robustesse, et donc le bienfondé de l'utilisation de LARSIM pour effectuer des projections hydrologiques.

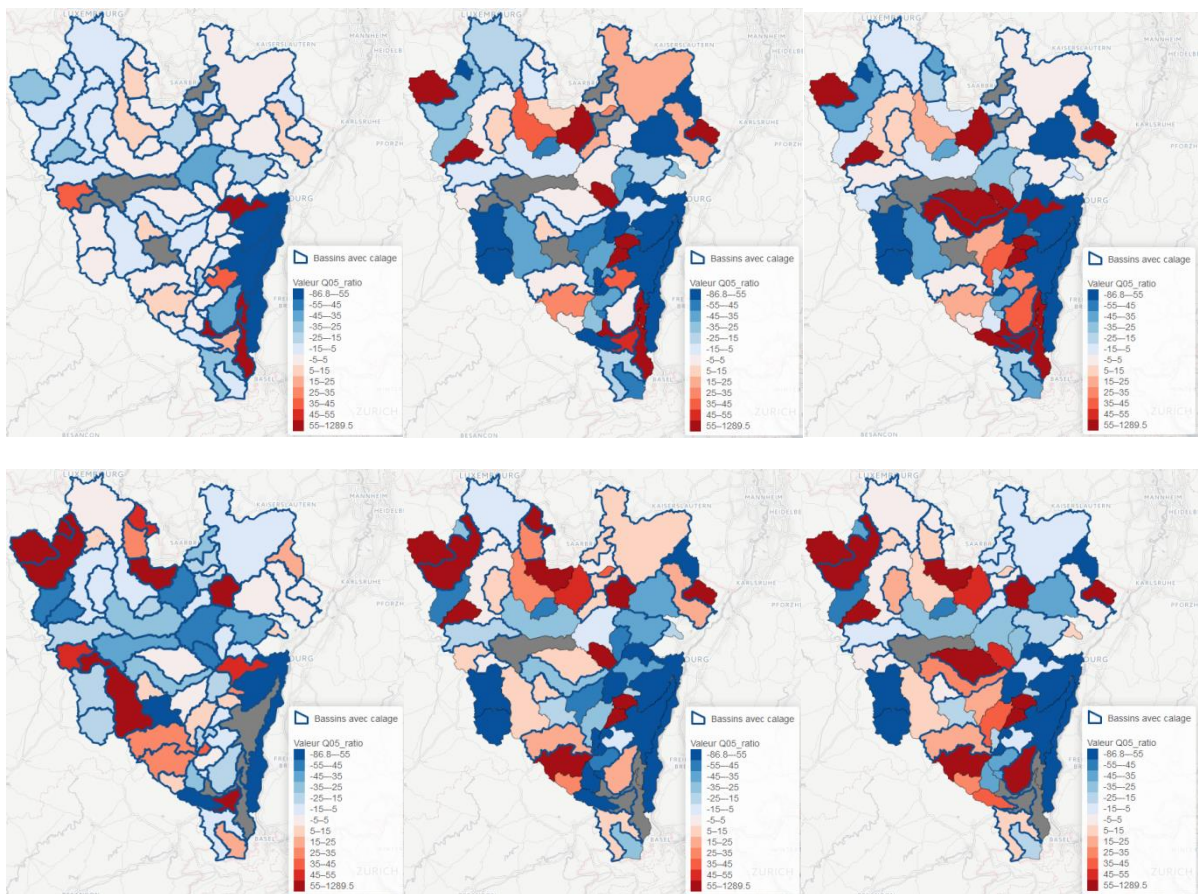
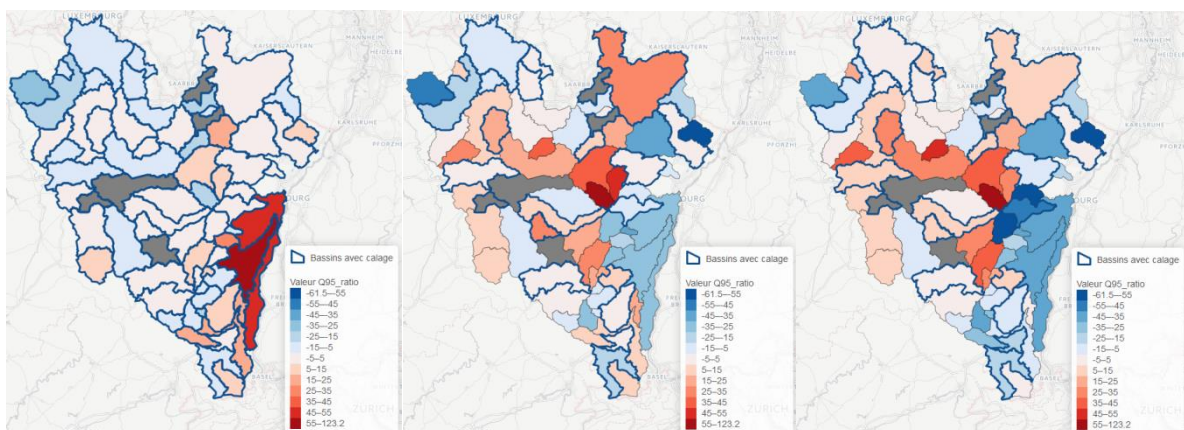


Figure 24 : Idem que Figure 22 mais pour le quantile 5 % (bas débits).

Les cartes de quantiles 5 % (Figure 24) montrent des biais plus importants. Ceci était attendu, car des erreurs faibles en termes de débits peuvent provoquer des erreurs relatives élevées. Nous voyons que sur la période de calage, les biais sont importants pour LARSIM sur les stations non calées, mais plus faibles ailleurs tout comme pour GRSD. En revanche, la dégradation de GRSD sur certaines stations en contrôle est plus élevée que celle de LARSIM.



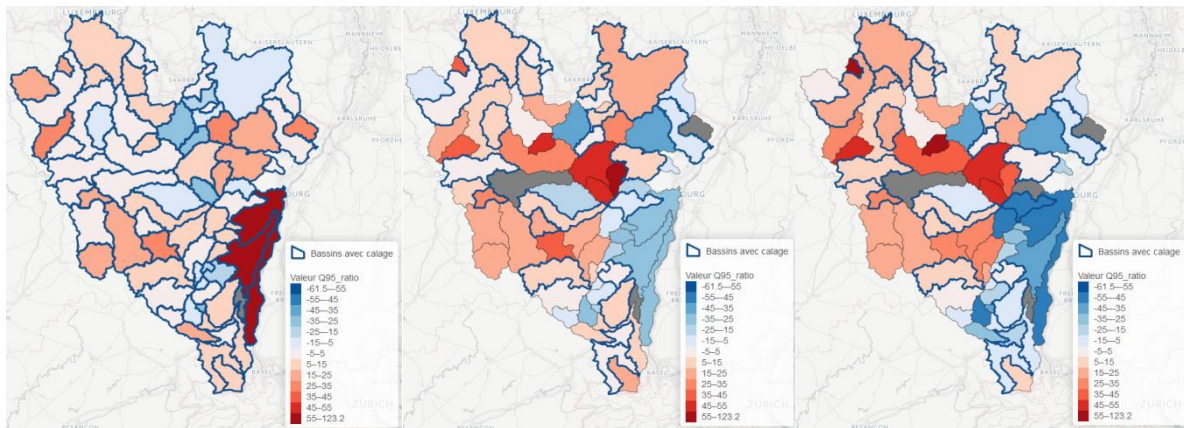


Figure 25 : Idem que Figure 22 mais pour le quantile 95 % (hauts débits).

Sur la Figure 25, les biais des quantiles 95 % sont plus faibles que pour les quantiles 5 %. Les deux modèles semblent présenter des biais faibles en calage et en contrôle sur les stations sur lesquelles ils sont calés, en revanche les biais sont plus importants sur les stations non calées de LARSIM.

Nous présentons quelques-uns des résultats de performances ci-dessous, pour la Moselle à Perl (Tableau 6) et pour la Zorn à Waltenheim (Tableau 7) pris comme exemples d'un grand bassin et d'un petit bassin.

Tableau 6 : Performances des modèles en calage et contrôle pour la Moselle à Perl. P1 : 1971-1985 ; P2 : 1986-2000.

Modèle	Période de calage	KGE hauts débits		KGE hauts et bas débits		KGE bas débits	
		Calage	Contrôle	Calage	Contrôle	Calage	Contrôle
GRSD	Complète	0.96	-	0.93	-	0.90	-
	P1	0.95	0.90	0.82	0.86	0.69	0.83
	P2	0.98	0.89	0.95	0.84	0.92	0.79
LARSIM	Complète	0.93	-	0.88	-	0.83	-
	P1	0.94	0.89	0.87	0.90	0.80	0.91
	P2	0.91	0.92	0.89	0.80	0.87	0.69
LARSIM manuel	Complète	0.83		0.81		0.78	
	P1	0.79		0.84		0.89	
	P2	0.87		0.78		0.68	

Le Tableau 6 montre plusieurs choses :

- Bonne performance générale des modèles, en calage et en contrôle.
- Dégradation raisonnable des performances en contrôle par rapport au calage.
- La dégradation de LARSIM en contrôle semble moindre.
- Meilleure performance générale du calage automatique de LARSIM par rapport au calage manuel.
- Résultats proches entre GRSD et LARSIM : GRSD semble un peu meilleur pour les forts débits et LARSIM semble un peu meilleur pour les faibles débits.

Tableau 7 : Performances des modèles en calage et en contrôle pour la Zorn à Waltenheim. P1 : 1971-1985 ; P2 : 1986-2000.

Modèle	Période de calage	KGE hauts débits		KGE hauts et bas débits		KGE bas débits	
		Calage	Contrôle	Calage	Contrôle	Calage	Contrôle
GRSD	Complète	0.90	-	0.82	-	0.73	-
	P1	0.90	0.79	0.80	0.39	0.69	-0.01
	P2	0.72	0.65	0.80	0.55	0.89	0.45
LARSIM	Complète	0.88	-	0.81	-	0.74	-
	P1	0.82	0.88	0.75	0.44	0.69	0.01
	P2	0.71	0.67	0.78	0.61	0.86	0.55
LARSIM manuel	Complète	0.83		0.79		0.75	
	P1	0.86		0.75		0.63	
	P2	0.81		0.73		0.64	

Le Tableau 7 présente des conclusions à peu près similaires, à ceci près que le calage manuel de LARSIM semble plus proche du calage automatique. On remarque cependant une dégradation de la performance en bas débits des modèles en contrôle, notamment pour GRSD. Un certain nombre de données manquantes pour les faibles gammes de débits pour cette station lors de la période de calage peut expliquer cette baisse de performance.

La Figure 26 montre les performances en termes de KGE agrégées sur toutes les stations (gauche) et sur les stations calées dans LARSIM (droite). Cette figure montre logiquement les meilleures performances de LARSIM sur les stations sur lesquelles il est calé. Cependant, et même si les performances de LARSIM sont moindres par rapport à GRSD lorsque l'on considère l'ensemble des stations, celles-ci restent acceptables et montrent une dégradation limitée lorsque l'on évalue le modèle sur une période indépendante (perf. en contrôle). Le choix de la période de calage a un impact négligeable sur la performance en calage ou en contrôle lorsque l'on considère la totalité des stations. En revanche, l'impact est plus grand si l'on ne considère plus que les stations calées dans LARSIM : le calage sur P2 est le plus performant pour GRSD alors que le calage sur P1 est le plus performant pour LARSIM.

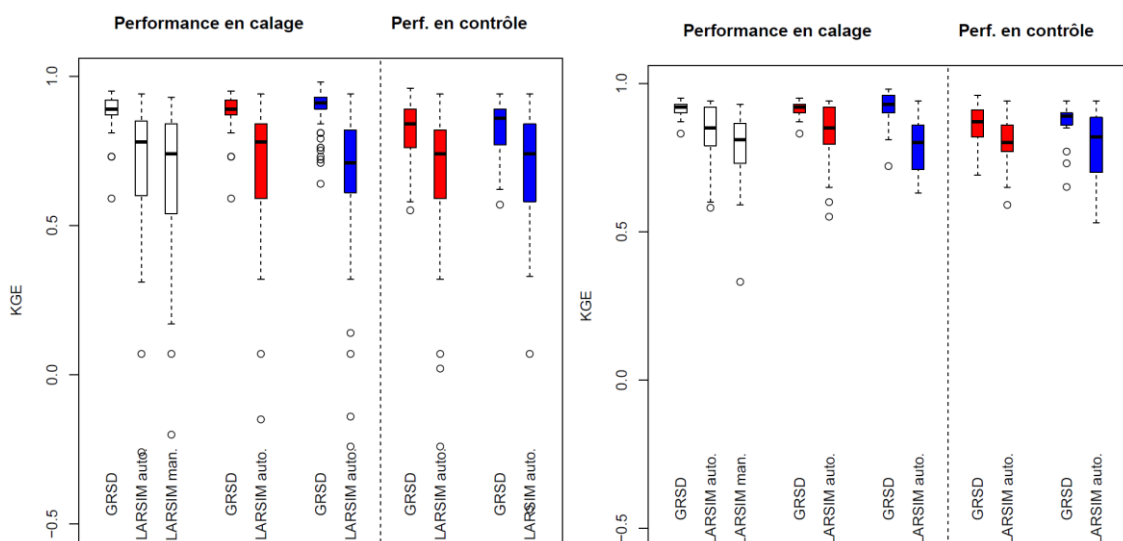


Figure 26 : Distributions des valeurs de KGE pour les différentes configurations des modèles en calage et en contrôle. En blanc les calages sur la période complète, en rouge les calages sur P1, en bleu les calages sur P2. Gauche : les 70 stations.

Droite : uniquement pour les stations calées dans LARSIM. La boîte des boxplots représente les valeurs entre 25 et 75%, le trait plein est la médiane, les extrêmes des moustaches sont les quantiles 5 et 95 % et les cercles sont les outliers.

Les performances des modèles en termes de KGE_DOUBLE, KGE_INVQ et biais (Figure 27, Figure 28 et Figure 29) montrent des évolutions similaires à la Figure 26. Cependant, la performance de LARSIM en bas débits, notamment en période de contrôle et pour les stations calées, est plus proche de celle de GRSD.

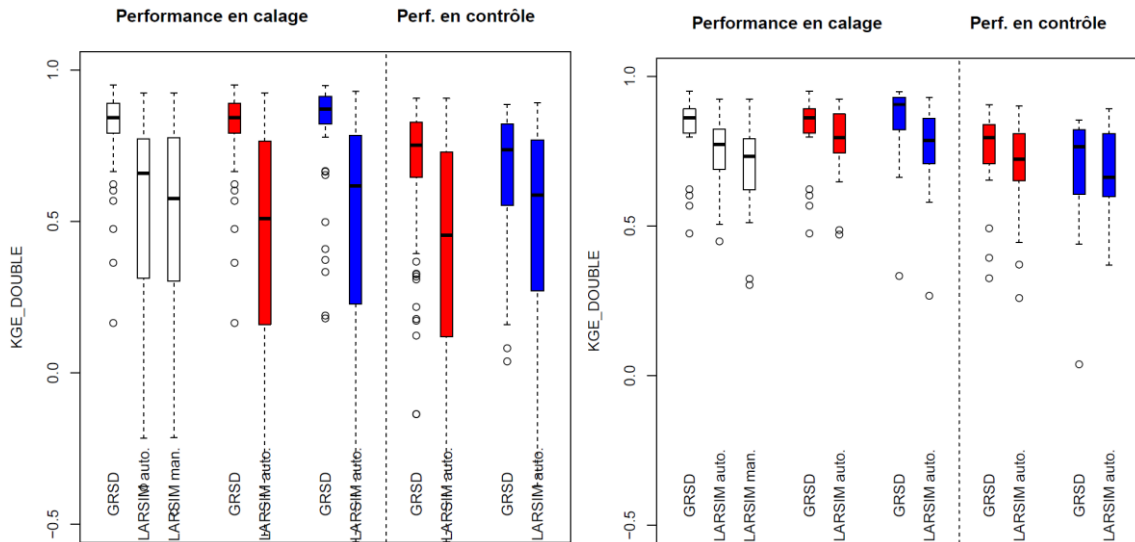


Figure 27 : Idem que Figure 26 pour le KGE_DOUBLE (hauts et bas débits). Gauche : les 70 stations. Droite : uniquement pour les stations calées dans LARSIM.

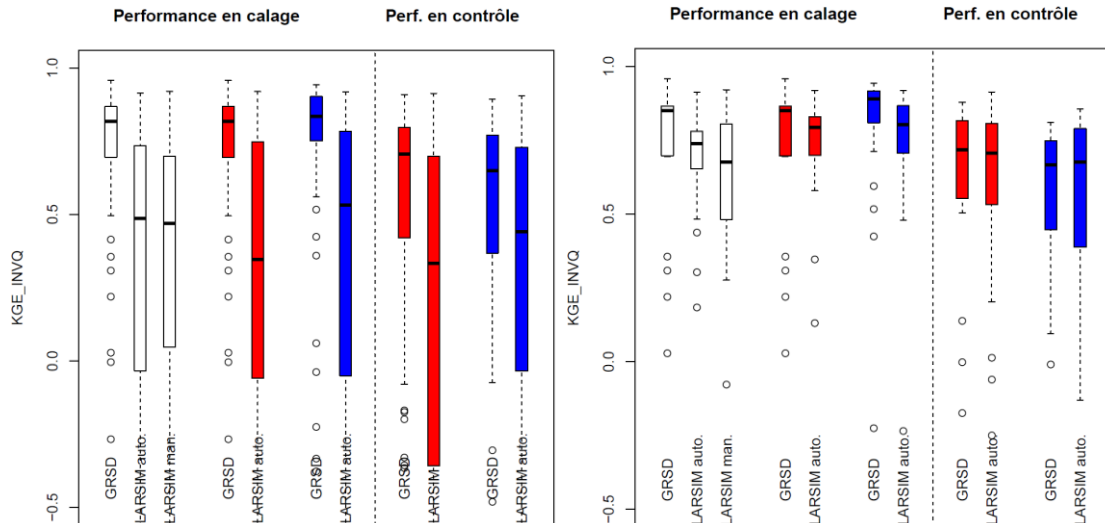


Figure 28 : Idem que Figure 26 pour le KGE_INVQ (bas débits). Gauche : les 70 stations. Droite : uniquement pour les stations calées dans LARSIM.

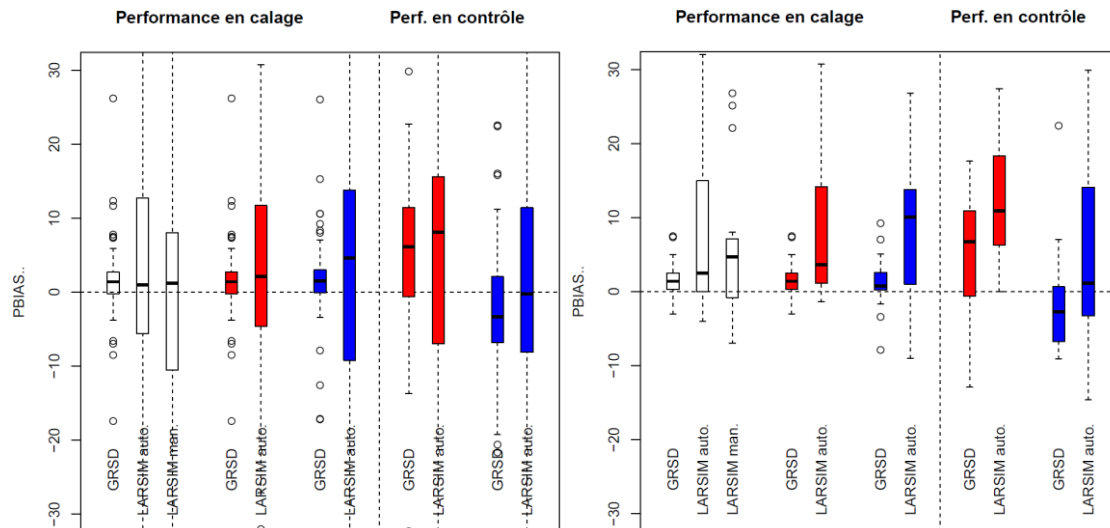


Figure 29 : Idem que Figure 26 pour le biais en pourcentage (des valeurs positives indiquent une surestimation des débits, alors que des valeurs négatives indiquent une sous-estimation des débits). Gauche : les 70 stations. Droite : uniquement pour les stations calées dans LARSIM.

3.8. Conclusions sur les calages des deux modèles

Les modèles GRSD et LARSIM ont été calés automatiquement sur trois périodes : 1971-2000, 1971-1985 et 1986-2000. Ces calages ont permis de valider la méthode adoptée pour le calage automatique : ces calages peuvent donc être utilisés pour réaliser les projections hydrologiques.

L'analyse spatiale de la performance des modèles a montré que les modèles donnaient des résultats éloignés des débits observés disponibles sur la plaine d'Alsace. Ceux-ci étant très influencés, il a été décidé d'utiliser pour ces bassins des paramètres issus de bassins non influencés situés en amont. Les modèles utilisés étant semi-distribués, on peut s'attendre à ce que les débits alors simulés représentent de manière suffisante l'hydrologie naturelle pour les bassins influencés, car une partie importante des débits amont est propagée pour obtenir les débits aval.

Pour mettre en œuvre une chaîne de modélisation de l'impact du changement climatique sur les débits, il est important de disposer de modèles robustes. Cette robustesse a été évaluée par le contrôle des calages P1 et P2 sur des périodes indépendantes. Nous avons montré que la dégradation des performances, qui est logique, reste faible, et qu'elle est légèrement plus importante pour GRSD que pour LARSIM.

4 Calages complémentaires de GRSD : calages effectués sur des périodes glissantes

Ces calages sur des périodes glissantes de 10 ans entre 1971 et 2000 (21 calages différents) ont pour but de pouvoir étudier de manière plus continue la robustesse des jeux de paramètres du modèle ainsi optimisés. En effet, de simples calages sur P1 et P2 ne permettent pas de mettre en évidence d'éventuelles tendances dans les valeurs optimales de paramètres que l'on pourrait lier par exemple à des tendances climatiques. Les calages sur des périodes glissantes ont été effectués pour GRSD uniquement, le temps de calcul de LARSIM ne permettant pas de mettre en place ces calages pour ce modèle.

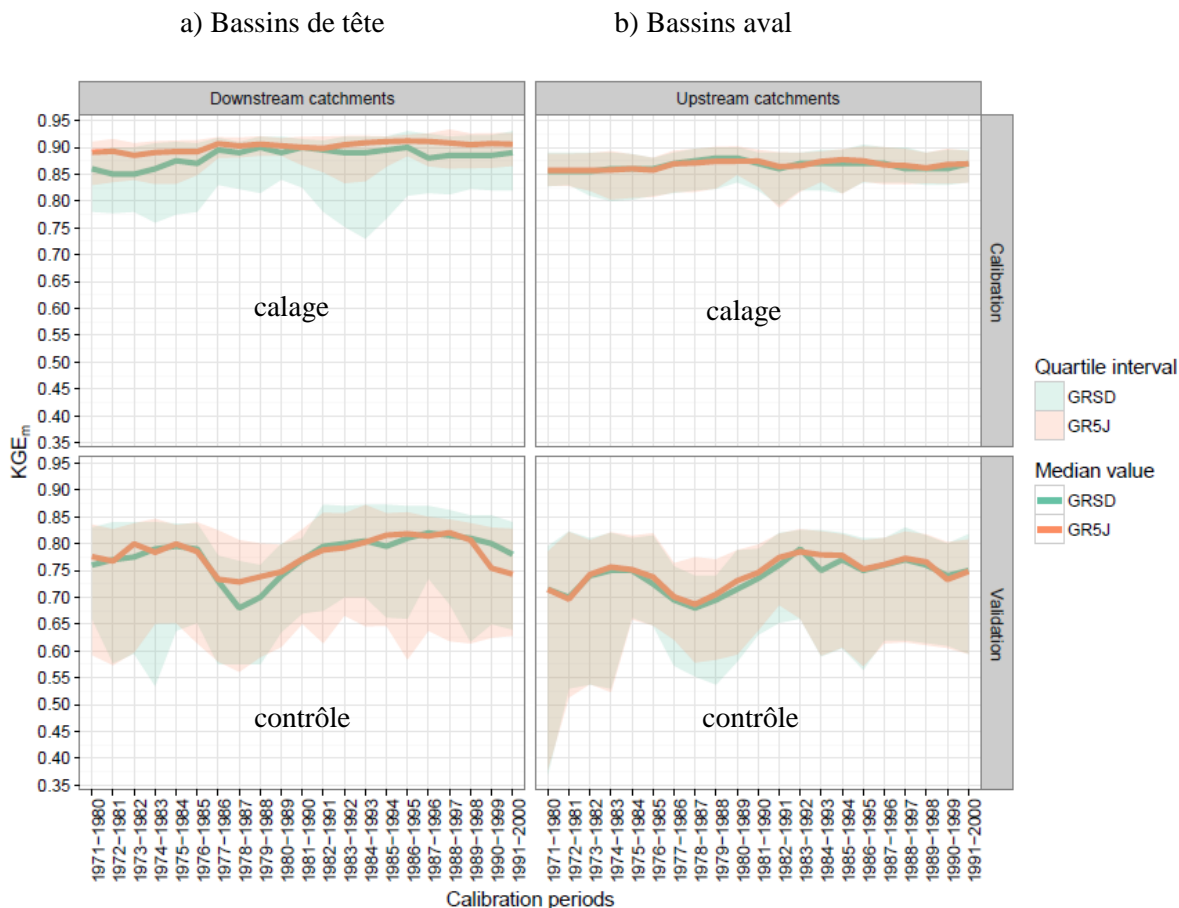


Figure 30 : Evolution de la performance (KGE) des calages de GRSD sur les périodes glissantes. Comparaison avec le modèle global GR5J sur lequel se base GRSD. a) les bassins aval ; b) les bassins de tête. Les bassins de la plaine d'Alsace ont été retirés de cette analyse. D'après de Lavenne et al. (2016).

Dans un premier temps, nous comparons les performances des différents calages de GRSD à ceux de GR5J (Figure 30). Logiquement, leurs performances sont équivalentes pour les bassins de tête. En revanche, les performances sur les bassins aval ne sont pas meilleures pour GRSD que pour GR5J, contrairement à ce à quoi on aurait pu s'attendre. Plusieurs explications sont possibles : les performances du modèle global sont déjà très bonnes ; l'hétérogénéité spatiale de la zone d'étude de MOSARH21 est faible ; des artefacts numériques d'optimisation peuvent pénaliser le modèle semi-distribué. En ce qui concerne l'évolution temporelle de la performance des modèles, elle est très stable en calage, mais l'est un peu moins en contrôle (baisse entre 1974 et 1989). Cependant, aucune tendance nette sur l'ensemble de la période n'est observable et il n'a pas été possible de lier cette légère baisse à une raison particulière. Il convient également de noter que l'intervalle interquartile 25-75 % augmente en contrôle pour les deux catégories de bassins ce qui traduit une dispersion plus forte des valeurs de KGE qu'en calage.

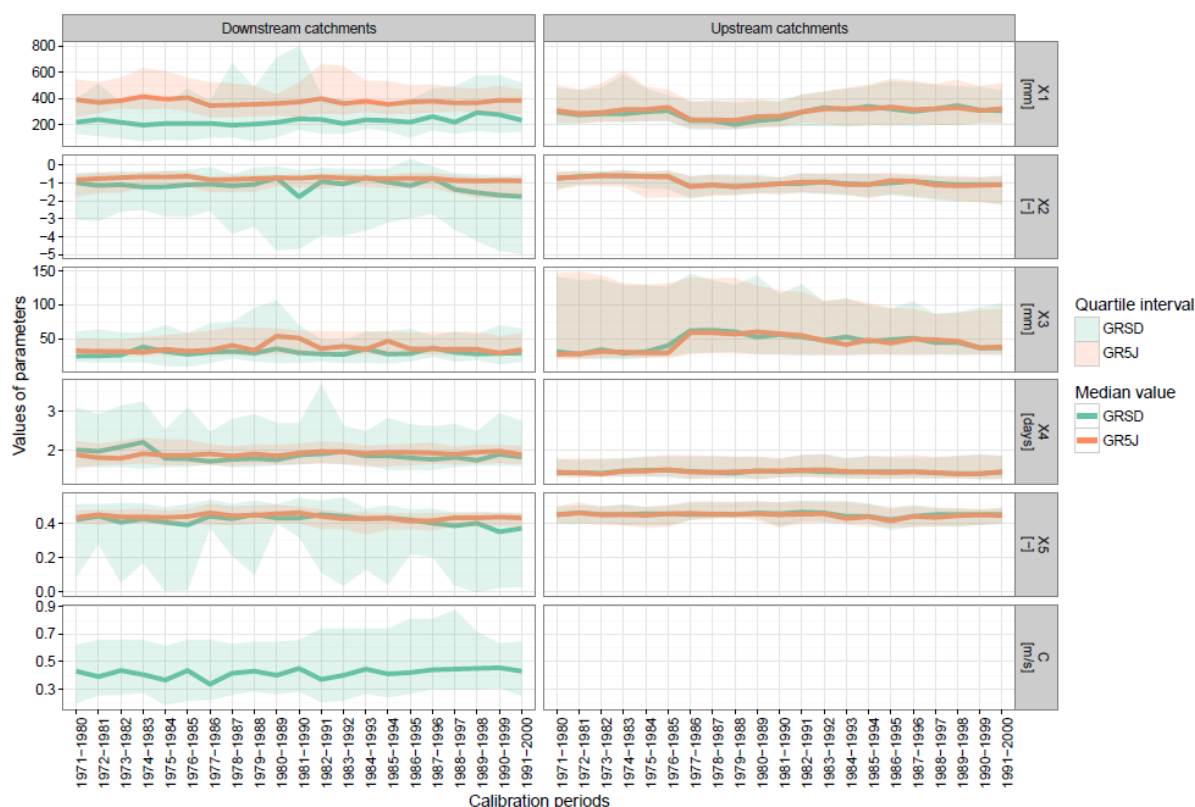


Figure 31 : Idem que Figure 30 (sans période de contrôle) pour les valeurs de paramètres de GRSD. D'après de Lavenne et al. (2016).

Une analyse comparable a été réalisée sur les paramètres de GRSD (Figure 31). Seul le paramètre X1 (capacité du réservoir de production du modèle, c'est-à-dire paramètre réglant le bilan d'eau et l'évaporation du modèle) présente une différence entre GRSD et GR5J. En ce qui concerne la tendance des paramètres, seule une légère diminution des paramètres X2 et X5 (paramètres contrôlant l'échange souterrain en eau) est notable. Cependant, cette baisse reste faible et probablement non significative. Une analyse plus complète de ces résultats est présentée par de Lavenne et al. (2016).

5 Conclusions

Les calages de référence de GRSD et LARSIM ont servi à mettre en place les projections. Nous avons pu déterminer que les calages réalisés étaient plutôt robustes, notamment ceux de GRSD ainsi que ceux de LARSIM sur les stations qui ont été calées. La dégradation observée sur les différents critères (KGE, bilan, quantiles) reste limitée en contrôle. Les calages sur des périodes différentes présentent cependant des jeux de paramètres et des performances légèrement différentes. Ceux-ci pourront donc aider à mesurer comment l'incertitude liée à cette étape de modélisation impacte les projections hydrologiques.

6 Références

Coron, L., 2013. Les modèles hydrologiques conceptuels sont-ils robustes face à un climat en évolution ? Thèse de doctorat, Irstea (Antony), AgroParisTech (Paris), 364 pp.

de Lavenne, A., Thirel, G., Andréassian, V., Perrin, C., and Ramos, M.-H.: Spatial variability of the parameters of a semi-distributed hydrological model, Proc. IAHS, 373, 87-94, doi:10.5194/piahs-373-87-2016, 2016.

Mathevet, T., 2005. Quels modèles pluie-débit globaux pour le pas de temps horaire ? Développement empirique et comparaison de modèles sur un large échantillon de bassins versants. Thèse de Doctorat, ENGREF (Paris), Cemagref (Antony), France, 463 pp.

Michel, C. : Hydrologie appliquée aux petits bassins versants, Edit. CEMAGREF, Antony, 413 p., 1991.

Nascimento, N.O., 1995. Appréciation à l'aide d'un modèle empirique des effets d'action anthropiques sur la relation pluie-débit à l'échelle du bassin versant. Thèse de Doctorat, CERGRENE/ENPC, Paris, 550 pp.

Perrin, C., 2000. Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative. Thèse de Doctorat, INPG, Grenoble, 530 pp.

FUTURE EVOLUTION OF RIVER DISCHARGE FOR THE FRENCH RHINE BASIN IN A CONTEXT OF CLIMATE CHANGE

AN UPDATED EVALUATION BASED ON THE AR5 IPCC CLIMATE SIMULATIONS

Evolution future des débits des cours d'eau dans le bassin du Rhin en contexte de changement climatique

Une évaluation actualisée avec les simulations climatiques issues du 5^{ème} rapport du GIEC

Guillaume Thirel¹, Charles Perrin¹, Gilles Drogue², Kai Gerlinger³, Julia Krumm³, Jean-Pierre Wagner⁴

¹Irstea, HBAN, 1 rue Pierre-Gilles de Gennes CS 10030, 92761 Antony CEDEX, France
guillaume.thirel@irstea.fr, charles.perrin@irstea.fr

²Laboratoire LOTERR, Université de Lorraine, Metz, France
gilles.drogue@univ-lorraine.fr

³HYDRON, Karlsruhe, Allemagne
Kai.Gerlinger@hydron-gmbh.de, Julia.Krumm@hydron-gmbh.de

⁴DREAL Lorraine, Metz, France
Jean-Pierre.WAGNER@developpement-durable.gouv.fr

ABSTRACT

This paper presents the MOSARH21 (Evolution of the discharge of the Rivers MOselle, Sarre and tributaries of the River RHine in France during the 21st century) project funded by the Rhine-Meuse Water Agency, which aims at making an evaluation of the future impacts of climate changes on river discharge for the French tributaries of the River Rhine. The MOSARH21 project uses the most recent climate projections produced during the last IPCC climate modeling exercise. The assessment of the impacts of these new climate projections on the Rhine basin hydrology will help verifying to which extent they corroborate previous projections and how they could contribute to stakeholders decisions, e.g. in terms of adaptation strategies. This will be achieved by comparing the MOSARH21 results with those obtained from impact studies previously made in the French part of the Rhine basin. To do so, a prospective study is performed following a top-down ensemblist multi-model approach (using two hydrological models and several parameter sets) in conjunction with an ensemble of downscaled climate projections. The hydrological impacts of future climate changes will be quantified for different time slices (with a particular focus on the mid-21st century) based on indices of regimes, floods and low flows. The quantification of the uncertainties associated with the modelling chain will be a major aspect of this work.

KEY WORDS

Multi-model approach, hydrological impact, CMIP5 climate modeling experiment, Rhine basin, France

RESUME

*Evolution future des débits des cours d'eau futurs sur dans le bassin du Rhin en contexte de changement climatique
Une évaluation actualisée s'appuyant sur les simulations climatiques issues du 5^{ème} rapport du GIEC
Cet article présente le projet MOSARH21 (Evolution des débits de la Moselle, de la Sarre et des affluents du Rhin en France au cours du 21^{ème} siècle), soutenu par l'Agence de l'Eau Rhin-Meuse, qui a pour ambition de réaliser une évaluation des impacts futurs des changements climatiques sur les débits de la partie française des affluents du Rhin, en utilisant les projections climatiques les plus récentes produites dans le cadre des travaux du GIEC. Il paraît en effet opportun d'évaluer les impacts de ces nouvelles projections climatiques sur l'hydrologie du bassin du Rhin, pour voir dans quelle mesure elles permettent d'affiner les précédentes projections et in fine d'éclairer la décision publique. Les résultats seront comparés à ceux obtenus dans le cadre d'études d'impact antérieures réalisées sur le bassin. Pour cela,*

¹ Corresponding author: Guillaume Thirel, guillaume.thirel@irstea.fr

une étude de prospective hydrologique sera effectuée dans une logique ensembliste et descendante grâce à la mise en œuvre d'une approche multi-modèles (exploitant deux modèles hydrologiques et un ensemble de paramètres) utilisée conjointement avec un ensemble de projections climatiques désagrégées à l'échelle d'intérêt. Les impacts hydrologiques du changement climatique futur seront quantifiés à différents horizons temporels (notamment celui de 2050) sur la base de divers indicateurs relatifs aux régimes, aux crues et aux étiages. Une attention particulière sera donnée à la quantification des incertitudes associées à la chaîne de modélisation.

MOTS-CLEFS

Approche multi-modèles, impact hydrologique, expérience numérique CMIP5, bassin du Rhin, France

1. INTRODUCTION

The fifth Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC) has recently been published [IPCC, 2014]. Summarizing the reports of three working groups (WGs), respectively dealing with physical science basis, impacts, adaptation and vulnerability and mitigation of climate change, the synthesis report provides a state-of-the-art view on scientific and socio-economic aspects of climate change. It qualifies the warming of oceans and atmosphere as unequivocal, and states that human influence is extremely likely to be the main cause of the observed warming. These conclusions strengthen the conclusions of the former report (AR4) published by the IPCC [2007].

Rising temperature together with modifications of precipitation regimes have an impact on water cycle at many scales. Surface hydrology may already be altered by these changes [Renard et al., 2006]. River discharge is one of the most important water sources for many water uses: hydro-electricity, agriculture, drinking water, power plants cooling, etc. It also represents a risk, during floods and droughts. Therefore assessing the long term (over the next decades) evolution of flows in a context of climate change is essential. Climate projections are usually designed through a forcing modelling chain: possible future emissions of greenhouse gases are established based on socio-economic and demographic scenarios, then these scenarios are used as inputs to General Circulation Models (GCMs). The outputs from GCMs are downscaled at finer spatial and temporal resolutions to be used for regional modelling purposes. To transform the climate projections into hydrological projections, a further step of hydrological modeling is implemented.

Such a top-down modeling chain has already been used for assessing the impact of climate change on hydrology for several large-scale areas in France [Chauveau et al., 2013], Germany [Görger et al., 2010], Europe [Forzieri et al., 2014] and Africa [Roudier et al., 2014]. These impact studies of climate change allow public water authorities, like e.g. the French « Directions Régionales de l'Environnement, de l'Aménagement et du Logement » (DREALs) or the German Länder, to get scientifically consistent estimations of relevant discharge indices that are useful to anticipate the management of water resources. However, each country making its own choices on climate scenarios, hydrological models and discharge indices (especially for low water), and data sharing being sometimes difficult between countries, the comparability of results and studies for transnational basins is usually limited. Yet the need for an international coordination within transboundary river basins is important for the identification of the likely impact of climate change on the occurrence of droughts or floods which shall be taken into account in the forthcoming review of the preliminary flood risk assessment for the purpose of the European Directive 2007/60/EC on the assessment and management of flood risks² and the review of the characteristics of the river basin district for the purpose of the European Directive establishing a framework for Community action in the field of water policy³. This requires a significant degree of coordination and harmonization of the existing regional tools and methodologies for producing hydrological scenarios [Bauwens et al., 2014]. Moreover, impact studies should be updated with the most recent generation of climate projections to better assess the adaptation measures.

This article presents the first outcomes of the MOSARH21 project (Evolution of the discharge of the Rivers MOselle, SARre and tributaries of the River RHine in France during the 21st century). The project is supported by the Rhine-Meuse Water Agency whose general objectives are to enhance water management in a broad sense in French parts of the Rhine and Meuse basins. By fostering the collaboration between hydrological modellers from France and Germany, and by setting up a common methodology to produce

² See article 14 of the Floods Directive / Reporting sheet for the implementation of the Floods Directive (2007/60/EC) Endorsed by Water Directors 30 November 2009 / Reporting sheets for the implementation of the Floods Directive (2007/60/EC) Endorsed by Water Directors: February 2011 /

³ See article 5 of the Water Framework Directive and Technical Report n° 2009 – 040 for the Common Implementation Strategy for the Water Framework Directive (2000/60/EC) "Guidance Document n°24 - River basin management in a changing climate"

hydrological projections with the most recent climate projections, this project intends to give a new insight on water resources vulnerability to climate change during the 21st century. A link with previous impact studies made either on the French part or on the German part of the Rhine basin will be established. To this aim, hydrological projections will be produced by using both AR4 and AR5 climate projections. The comparison between impact studies is based on a common set of metrics.

The paper is organized as follows: first, previous studies including the French Rhine tributaries are analyzed and compared, with a focus on the methodological aspects. Then, the two hydrological models used in the MOSARH21 project are presented. The study area and the data are described in the third section. The last section focuses on the methodological aspects of the project.

2. SETTING THE SCENE: AN OVERVIEW OF RECENT HYDROLOGICAL IMPACT STUDIES IN THE RIVER RHINE BASIN

In this section, several studies assessing the impact of climate change on river discharges in areas including the French Rhine tributaries are presented. The expression “French Rhine tributaries“ will be used to refer to the French part of the River Moselle, the French part of the River Sarre, and the French tributaries of the section of the Rhine making the border with Germany (Ill, Moder and Sauer).

Project	Rheinblick2050	VULNAR	Explore 2070	FLOW MS
Greenhouse gases emission scenario	AR4 A1B	AR4 A1B, A2 and B1	AR4 A1B	AR4 A1B
Number of GCMs	2	7	7	1
Downscaling method	Regional Climate Models	Statistical weather-type approach [Boé et al., 2007]	Statistical weather-type approach [Boé et al., 2007]	Regional Climate Model
Hydrological model(s)	HBV [Bergström, 1976]	MODCOU [Ledoux et al., 1989] et HPP-inv [Lecluse, 2014]	GR4J [Perrin et al., 2003] and Isba-Modcou [Habets et al., 2008]	LARSIM [Ludwig et Bremicker, 2006]
Reference period	1961-1990	1961-2000	1961-1990	1971-2000
Future period(s)	2021-2050 and 2071-2100	2046-2065 and 2081-2099	2046-2065	2021-2050

2.1 Rheinblick2050

This project took place from 2008 to 2010 and was initiated by the International Commission for the Hydrology of the Rhine basin (CHR). The objective of the project was to estimate the impacts of future climate change on discharge of River Rhine and its major tributaries, namely the Main and Moselle and associated institutions from France, the Netherlands, Switzerland, Luxembourg and Germany [Görgen et al., 2010]. The set up of the modelling exercise is described in Tab. 1.

Climate projections indicated rising temperatures up to 2.5°C for the near future and from 2.5 to 5.0°C for the far future. Precipitation is not expected to significantly change in the near future but could increase in winter and decrease in summer for the far future. A slight mean discharge increase was projected for the near future while for the far future opposite changes in winter and summer, in agreement with the precipitation changes. Low flows were studied through the calculation of the mean annual minimum 7-day flow. The only noticeable changes occurred for low flows with an increase during near future winters and a decrease during summers at the end of the century. Estimations of high flows change based on a weather generator approach did not lead to significant conclusions due to important uncertainties.

2.2 VULNAR

This project was funded by the French National Research Agency (ANR) from 2008 to 2012 (<http://www.cnrn.meteo.fr/spip.php?article278>). It aimed at studying the effects of climate change on the Upper Rhine aquifer layer (see its configuration in Tab. 1).

An increase of air temperature of 3°C and a decrease of precipitation around 15% are expected by the end of the century compared with a reference period set up as 1961-1990. Snow falls were projected to

drastically decrease by 70%. The impacts of climate change on aquifer levels was low, with a slight increase of the relative infiltration from rivers and a decrease of the average water recharge.

2.3 Explore 2070

Within this project, ordered by the French Environment Ministry, the possible impact of climate changes on discharge was assessed over France [Chauveau et al., 2013]. The Isba-Modcou and GR4J models were applied on a common set of 543 gauge stations over France first forced by observed meteorological data and in a second step by climate projections [Chauveau et al., 2013] as detailed in Tab. 1.

An increase of air temperature in the future from 1.4°C to 3°C is the most robust result of climate projections over France. Regarding precipitation, trends are uncertain but may show a decrease of around 20% in summer with no significant trend for the rest of the year. Both hydrological models predict a decrease of around 10% of annual discharge for the Moselle and Sarre, and around 20% for the other Rhine tributaries. Low flows are predicted to significantly decrease for all Rhine tributaries but the intensities of this decrease are different for both models: around 20% for Isba-Modcou, and up to 60% for GR4J. This discrepancy shows the utility of adopting a multi-model approach to better quantify the uncertainty coming from hydrological modelling. Finally, the sign of flood change was shown to be very uncertain, which can be directly linked with the possible change of rainfall totals during the rainy season (winter). Floods may slightly increase for the Moselle and Sarre basins and slightly decrease for the other Rhine tributaries, but the dispersion of results is high. The metrics used were mainly based on discharge regimes, lowest monthly flow with a return period of 5 years, and maximum daily discharge with a return period of 10 years.

2.4 FLOW MS

The “Floods and low flows management of the Moselle and Sarre basins – FLOW MS” project started in 2009 and was coordinated by the International Commissions for protection of Moselle and Sarre (Commissions Internationales pour la protection de la Moselle et de la Sarre, CIPMS). FLOW MS associated French, Luxemburgish and German institutions. The modelling chain was set up as described in Tab. 1. The study area of the FLOW MS project covered the Moselle basin, i.e. including its Luxemburgish and German parts, up to its confluence with the Rhine in Koblenz. A sample of 37 discharge gauging stations was selected.

Climate projections indicated an increase of air temperature of 1°C until 2050 for the basin, with drier summers and wetter winters. From a hydrological perspective, the FLOW MS project concluded to likely lower flows during summers and to significantly higher floods for winter. The metrics used to obtain these conclusions were mainly based on monthly mean, minimal and maximal discharge values.

As it is clear from the reading of Tab. 1, the hydrological projections realized within these four projects highly differ in their settings: GCMs, downscaling methods, hydrological models and time periods differ. Moreover, the indices used to assess the evolution of discharge differed a lot between the models. As a consequence, the MOSARH21 project is necessary to update and homogenise the hydrological projections and indices. This MOSARH21 project is presented in the following.

3. THE HYDROLOGICAL MODELS USED IN THE MOSARH21 PROJECT

3.1 The conceptual semi-distributed GR5J daily model (GRSD)

GRSD is a daily semi-distributed rainfall-runoff model. Based on the daily lumped GR5J model, mainly composed of a production reservoir and of a routing reservoir [Le Moine, 2008], the GRSD model was developed by Lobligeois [2014]. Instead of considering input data as uniform on each basin like in GR5J, GRSD spatially distributes input data on sub-basins (Fig. 1). The sub-basin discretization scheme can either correspond to the gauging stations network or to the modeller’s choice based for example on a specific contributing area. Therefore this semi-distributed model explicitly accounts for the heterogeneity of basins characteristics. In addition to the five free parameters of GR5J calibrated against observed discharge series, the GRSD model, has two parameters related to its hydraulic propagation scheme. Within the MOSARH21 project, the GRSD discretization corresponds to the selected gauged sub-basins. Hence, the rainfall-runoff GRSD model is calibrated at each gauge station.

A degree-day snow accumulation and melting module, CemaNeige [Valéry et al., 2004], is used together with GRSD. Each sub-basin is cut into five zones of equal area on which the snow accumulation and melting are calculated. Its two parameters are set up at fixed median values for this project.

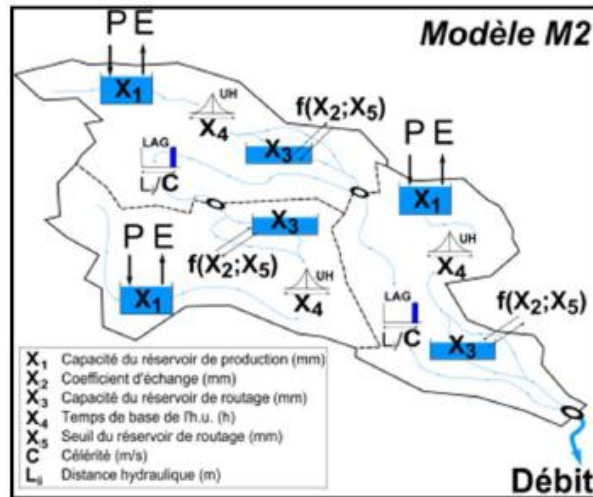
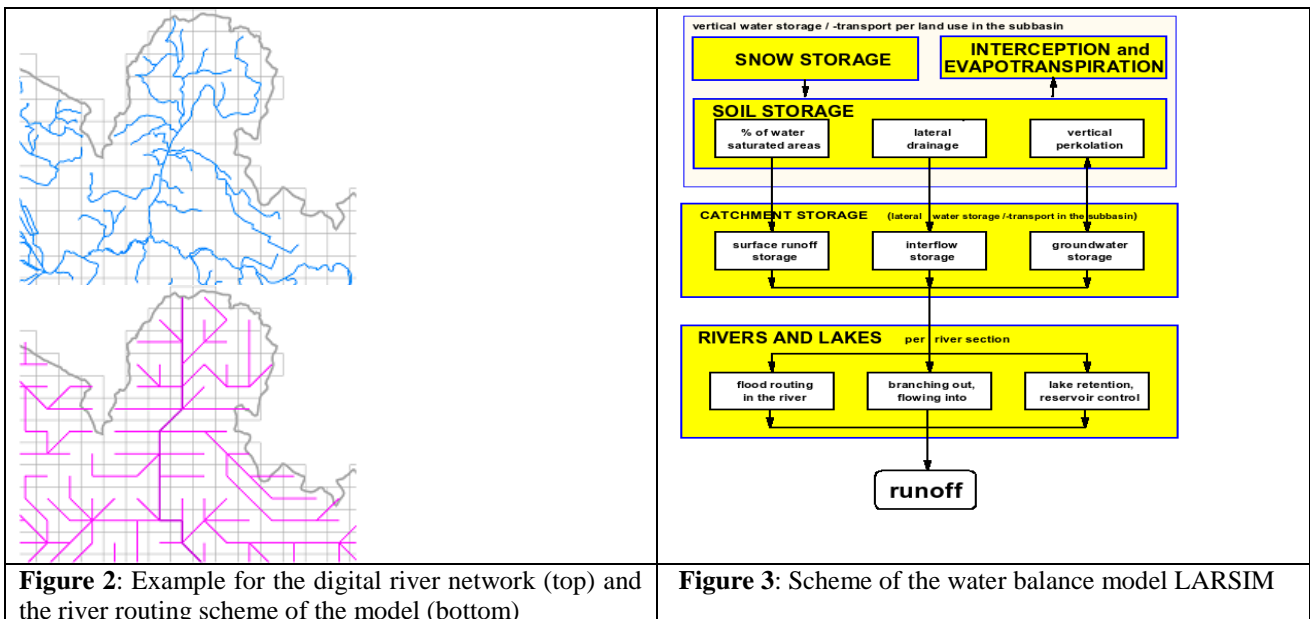


Figure 1: Spatial discretization and functioning of the GRSD model. From Lobligeois [2014].

3.2 The physically-based distributed LARSIM daily model

The water balance model LARSIM (Large Area Runoff Simulation Model; [Ludwig & Bremicker, 2006]) is a so-called conceptual model, i.e. the complex processes in the natural system are reproduced by simplified model concepts. The water balance model LARSIM is often applied with quadratic grid cells (subareas) of 1 km for the representation of the real river network (Fig. 2). Alternatively, a subarea structure based on hydrologic subareas is possible as well as any other grid cell size than 1km². For each cell, the following hydrological sub-processes are described: interception, evapotranspiration, snow accumulation, compaction and thaw, soil water replenishment, storage and lateral water transport within the area and translation and retention in channels and lakes (see LARSIM model schema, Figure 3).



LARSIM combines well-tried deterministic hydrological model components, which are as far as possible generally applicable and are based on accessible system data for the land surface. Emphasis is laid on the reliable determination of evapotranspiration by using the Penman-Monteith equations. Evapotranspiration and the soil water budget are calculated separately for different land uses and field capacities of the soils. Up to 16 different land uses with their specific evapotranspiration and run-off properties were recorded for each individual grid area (Figure 3).

The LARSIM model is widely used for operational flood forecast centres [Bremicker et al. 2013]. LARSIM has originally been developed for the French Rhine tributaries for hourly calculation time intervals for the purpose of flood forecast for the operational flood warning centres Meuse-Moselle and Rhin-Sarre of the French DREALs Lorraine and Alsace. LARSIM has hence been adapted to daily calculation time intervals for the purpose of the FLOW MS and MOSARH21 projects.

4. STUDY AREA AND DATA OF THE MOSARH21 PROJECT

4.1 The study area and monitoring discharge stations network

The French part of the Rhine basin represents 12 % of the whole Rhine catchment basin and includes two sub-districts: the River Rhine itself and its direct tributaries located in the Alsace region (8,160 km²) and the Moselle and Saar basins up to the German and Luxemburgish borders (15,360 km²). The eastern part comprises the 214 kilometres of the Rhine from the Swiss border at Basel (South) to the German border at Lauterbourg (North) and its tributaries which take their spring in the Vosges mountains (Moder, Sauer, Lauter, Bruche, Zorn, Lauch, Doller) and in the Jura mountains (River Ill). The western part comprises the Rivers Moselle and Saar and their tributaries. The Moselle has a total length of 313 kilometres between its spring in the Vosges Mountains and the border with Germany and Luxembourg. Its main tributary is the River Meurthe. The Saar also has its spring in the Vosges Mountain and flows to Germany at Saarguemines.

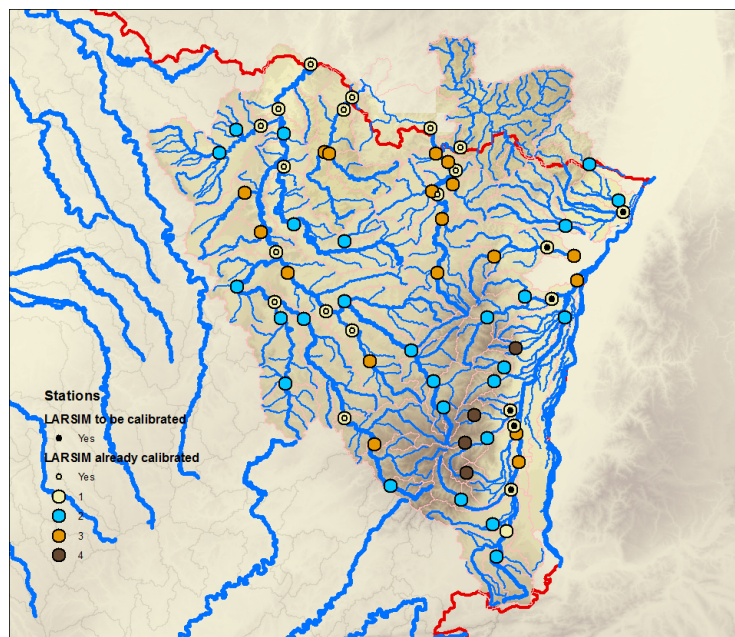


Figure 4: Monitoring discharge stations selected for the MOSARH21 project.

In order to make a link between past projects carried out on the French part or on the German part of the Rhine basin, and new projections produced in the MOSARH21 project, a set of gauging stations has been selected (Fig. 4). Calibrating the LARSIM model being time consuming, gauging stations where LARSIM was already calibrated and applied within the FLOW MS project are included. For Alsace, six other stations will be calibrated in the MOSARH21 project. All these stations were included in a priority dataset (category 1 in Fig. 2; 22 stations) for which a thorough analysis of models performance and hydrological projections will be made. A second category (category 2; 26 stations) for which only GRSD will be calibrated was established for operational purposes. These stations are of interest for the Water Public Authorities for either flood or drought monitoring. A detailed analysis of results for the LARSIM simulations will not be done on them. Finally, additional stations included within the Explore 2070 project were put in categories 3 and 4 (22 stations). These stations will not necessarily be subject to a detailed analysis. The quality of measured discharge and the potential anthropic influence was taken into consideration for selecting the stations.

4.2 Climate data

Daily meteorological data are derived from the gridded SAFRAN reanalysis [Vidal et al., 2010]. This dataset consists in an 8-km resolution atmospheric reanalysis over France from 1958. The atmospheric variables are rainfall, snowfall, mean air temperature, specific humidity, wind speed, solar and infrared radiation. The hydrological data were retrieved from the Banque Hydro website: <http://www.hydro.eaufrance.fr>. For the four stations located in Germany, data were obtained from the German federal Waterways and Shipping Administration (Wasser- und Schifffahrtsamt) and the Institute for Environment and Workerprotection in Saarland (Landesamt für Umwelt- und Arbeitsschutz Saarland).

The climate projections from four distinct sources will be used (Tab. 2). First, the projections from the two former projects, Explore 2070 and FLOW MS will be used. To have homogeneous hydrological projections, the Explore 2070 climate projections will be used as an input of both LARSIM and GRSD (a lumped version of the model, GR4J, was indeed used in Explore 2070). Regarding FLOW MS, GRSD will be fed with its projections, while LARSIM will be fed only on the Alsace region (projections were made only on the rest of the basin during the FLOW MS project).

New projections based on AR5 are made available by National Meteorological Services. Due to the availability of downscaled projections for most of the variables required by LARSIM, our choice should move towards the datasets available on the French Drias portal (<http://drias-climat.fr/>). Two climate modelling procedures are available. The first one, produced by the Pierre-Simon Laplace Institute (IPSL) is based on the IPSL-CM4 GCM, dynamically downscaled by the WRF model and corrected using a Cumulative Distribution Function – Transform (<http://drias-climat.fr/accompagnement/section/180>) for RCPs 4.5 and 8.5. The second climate modelling procedure, produced by the National Centre for Meteorological Research (CNRM), is based on the Arpege-global GCM, dynamically downscaled using the Aladin-Climat limited area model (LAM) and corrected with the quantile-quantile method (<http://drias-climat.fr/accompagnement/section/179>) for RCPs 2.6, 4.5 and 8.5. Both chains provide data at an 8-km resolution respectively from 1971-2100 and 1950-2100. Due to time computation, the MOSARH21 project will mainly focus on the 2021-2050 period and the RCP8.5 scenario for producing LARSIM simulations, while the other RCPs and the period 2071-2100 will be considered only for GRSD simulations.

Additional downscaled climate projections will be available for the project. They come from the use of a simpler statistical method, the Advanced Delta-Change [van Pelt et al., 2012]. These projections will be made available for the climate forcing of the GRSD rainfall-runoff model.

Origin of projection dataset	Explore 2070	Flow MS	Drias	Advanced-Delta Change
Scenario	AR4 A1B	AR4 A1B	RCP8.5 (RCPs2.6 & 4.5)	RCPs2.6, 4.5, 6.0 & 8.5
Number of GCMs	7	1	2	~ 20
Downscaling	Statistical weather-type approach	Regional Climate Model	Regional Climate Models	Statistical method
Unbiasing	-	Linear scaling & quantile-quantile	Distribution fitting-based methods	-
Spatial resolution	8x8 km	7x7 km	8x8 km	8x8 km
Hydrological models to run	LARSIM & GRSD	GRSD	LARSIM & GRSD (GRSD)	GRSD

Table 2: Overview of the hydrological projections realized within MOSARH21.

5. METHODS

5.1 Incorporating sources of uncertainties in the hydrological projections through an ensemblist approach

Throughout the complete modelling chain set up to produce hydrological projections, many sources of uncertainties exist: greenhouse gaz emission scenarios, GCMs, downscaling methods and hydrological models. In order to take into account these sources of uncertainties, the common approach is to consider several options for each step. This approach was adopted in MOSARH21: several RCPs will be considered, as well as two GCMs, two downscaling methods and two hydrological models.

An additional specific task will be performed on the parameterization of the hydrological models. Indeed, this source of uncertainty is most of the time neglected in climate change impact studies. However, it

has been shown that hydrological model errors, regardless to the mathematical structure of the model, suffer from a dependence to the hydro-climatic conditions of the calibration period [Coron et al., 2012; Thirel et al., 2015]. In order to take into account this source of uncertainty, the models will be calibrated over different climate contrasted periods.

5.2 Making projections of very high flows

Stochastic rainfall generators are tools that use statistical properties of rainfall observations time series (typically mean, variability, extreme values, spatial and temporal correlations, and successions of days with or without rainfall [Aubert, 2011]) to produce very long rainfall time series. These tools are largely used for providing inputs to hydrological models when the knowledge of extreme flows is necessary (e.g. dams design).

Within MOSARH21, a stochastic rainfall generator will be calibrated first on the statistical properties of rainfall time series obtained from a reference period, and then on the statistical properties of future conditions coming from the climate projections. Up to a thousand years of rainfall will be produced for both periods. Two sets of runs of GRSD (and possibly LARSIM) will then be produced. From the discharge time series obtained for both runs, 100-yr and 1000-yr floods will be calculated.

5.3 Common metrics to assess floods and low flows occurrences

One of the aims of the project is to assess the discharge projections on our study area on a common basis for both AR4 projections and the new AR5 projections. This needs to design an experimental protocol relying on a selected set of metrics. A review of the metrics used in the past studies will be made in order to select the adequate metrics for this project. A particular focus will be made on high and low flows, as well as on regimes.

6. CONCLUSIONS

The MOSARH21 project is still in its preliminary phase, therefore only the general context and the methodological aspects were presented here. Its aim is to update the expected discharge modifications related to future climate change on a set of stations with interests for the operational stakeholders, and to homogenise the indices used to assess the future changes taking into account the uncertainties of the modelling chain.

ACKNOWLEDGEMENTS

The authors gratefully thank Météo-France for providing the observed and future meteorological data, DREAL Lorraine and DREAL Alsace for LARSIM on the Moselle, Sarre and Rhine tributaries, KIT Institute of Meteorology and Climate Research - Troposphere Research (IMK-TRO) for COSMO-CLM data. Observed discharge data were downloaded from the Banque Hydro database (<http://www.hydro.eaufrance.fr/>) for the French stations and from the German federal Waterways and Shipping Administration (Wasser- und Schifffahrtsamt) and the Institute for Environment and Workerprotection in Saarland (Landesamt für Umwelt- und Arbeitsschutz Saarland) for the German stations. This project is funded by the Rhine-Meuse Water Agency (Agence de l'Eau Rhin-Meuse) in France.

REFERENCES

- Aubert, Y. (2011). Estimation des valeurs extrêmes de débit par la méthode Shyreg, *Thèse de doctorat*, Irstea (Aix), 316 pp.
- Bauwens A., Degré A, Deraedt D., Döring R., Drogue G., Huber N.P., Vanneuville W., Sinaba B., Fournier M.(2014). International assessment of future low-flow regimes and their impact on three water-related sectors in the Meuse basin – a collaborative approach. *International Journal of River Basin Management*, **13**, **1**. <http://dx.doi.org/10.1080/15715124.2014.983523>.
- Bergström, S., 1976. Development and application of a conceptual runoff model for Scandinavian catchments, *SMHI Report RHO 7*, Norrköping, 134 pp.
- Boé J., Terray L., Habets F., Martin E. (2007). Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. *International Journal of Climatology*. **27(12)** : 1643-1655.

- Bremicker M., Brahmner G., Demuth N., Holle F.-K. & Haag I. (2013): Räumlich hoch aufgelöste LARSIM Wasserhaushaltsmodelle für die Hochwasservorhersage und weitere Anwendung. – KW Korrespondenz Wasserwirtschaft 6 (9), 509–519
- Chauveau, M., S. Chazot, C. Perrin, P.-Y. Bourgin, E. Sauquet, J.-P. Vidal, N. Rouchy, E. Martin, J. David, T. Norotte, P. Maugis et X. De Lacaze (2013). Quels impacts des changements climatiques sur les eaux de surface en France à l'horizon 2070 ? (What will be the impacts of climate change on surface hydrology in France by 2070?). *La Houille Blanche*, (4): 5-15, doi:10.1051/lhb/2013027.
- CIPMS (2013). Identification des impacts possibles du changement climatique dans le bassin versant de la Moselle et de la Sarre, Commissions internationales pour la protection de la Moselle et de la Sarre, 13 p.
- Coron, L., Andréassian, V., Bourqui, M., Perrin, C. and Hendrickx, F. (2011). Pathologies of hydrological models used in changing climatic conditions: a review, *Hydro-Climatology: Variability and Change. IAHS Publication*. **344**, pp. 39-44.
- Forzieri, G., Feyen, L., Rojas, R., Flörke, M., Wimmer, F., and Bianchi, A. (2014). Ensemble projections of future streamflow droughts in Europe, *Hydrol. Earth Syst. Sci.*, **18**, 85-108, doi:10.5194/hess-18-85-2014.
- Görgen, K., Beersma, J., Brahmner, G., Buiteveld, H., Carambia, M., de Keizer, O., Krahe, P., Nilson, E., Lammersen, R., Perrin, C. and Volken, D. (2010). Assessment of Climate Change Impacts on Discharge in the Rhine River Basin: Results of the RheinBlick2050 Project, *CHR report*, **I-23**, 229 pp., Lelystad, ISBN 978-90-70980-35-1.
- Habets, F., et al. (2008), The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France, *J. Geophys. Res.*, **113**, D06113, doi:10.1029/2007JD008548.
- IPCC (2007). – IPCC AR4 SYR . Core Writing Team; Pachauri, R.K; and Reisinger, A., ed., *Climate Change 2007: Synthesis Report, Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, IPCC, ISBN 92-9169-122-4.
- IPCC (2014). – Summary for Policymakers. In: *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1-32.
- Lecluse, S. (2014). Modélisation de l'influence du changement climatique sur la nappe phréatique du Rhin Supérieur. *Thèse de doctorat*. Univ. de Strasbourg. Laboratoire d'hydrologie et de géochimie de Strasbourg. 292 p.
- Ledoux, E., G. Girard, G. De Marsily, and J. Deschenes (1989), Spatially distributed modeling: Conceptual approach, coupling surface water and ground-water, in *Unsaturated Flow Hydrologic Modeling: Theory and Practice*, NATO ASI Series C, vol. **275**, edited by H. J. Morel-Seytoux, pp. 435–454, Kluwer Acad., Norwell, Mass.
- Le Moine, N. (2008) Le bassin versant de surface vu par le souterrain : une voie d'amélioration des performances et du réalisme des modèles pluie-débit ? *Thèse de Doctorat*, Université Pierre et Marie Curie (Paris), Cemagref (Antony), 324 pp.
- Lobligeois, F. (2014). Mieux connaître la distribution spatiale des pluies améliore-t-il la modélisation des crues ? Diagnostic sur 181 bassins versants français. *Thèse de Doctorat*, Irstea (Antony), AgroParisTech (Paris), 312 pp.
- Ludwig, K. et M. Bremicker (2006). The Water Balance Model LARSIM - Design, Content and Applications, Institut für Hydrologie der Universität Freiburg i. Br., 130 p
- Perrin, C., Michel, C. and Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, **279** : 275-289, DOI: 10.1016/S0022-1694(03)00225-7.
- Renard, B., M. Lang, P. Bois, A. Dupeyrat, O. Mestre, et al. (2006). Evolution des extrêmes hydrométriques en France à partir de données observées. *La Houille Blanche*, **6**, p. 48 - p. 54.
- Roudier, P., Ducharne, A., and Feyen, L. (2014). Climate change impacts on runoff in West Africa: a review, *Hydrol. Earth Syst. Sci.*, **18**, 2789-2801, doi:10.5194/hess-18-2789-2014.
- Thirel, G., Andréassian, V., Perrin, C. (2015). Editorial: On the need to test hydrological models under changing conditions. *Hydrological Sciences Journal*, under review.
- van Pelt, S. C., J. J. Beersma, T. A. Buishand, B. J. J. M. van den Hurk et P. Kabat (2012). Future changes in extreme precipitation in the Rhine basin based on global and regional climate model simulations. *Hydrology and Earth System Sciences* **16(12)**: 4517-4530.
- Valéry, A., Andréassian, V., Perrin, C., 2014. 'As simple as possible but not simpler': What is useful in a temperature-based snow-accounting routine? Part 2 – Sensitivity analysis of the Cemaneige snow accounting routine on 380 catchments, *Journal of Hydrology*, **517(0)**: 1176-1187, doi: 10.1016/j.jhydrol.2014.04.058.
- Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., Soubeyroux, J.-M. (2010), A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *International Journal of Climatology*, **30 (11)**, pp. 1627-1644.



Spatial variability of the parameters of a semi-distributed hydrological model

Alban de Lavenne, Guillaume Thirel, Vazken Andréassian, Charles Perrin, and Maria-Helena Ramos

Irstea, Hydrosystems and Bioprocesses Research Unit (HBAN), 1, rue Pierre-Gilles de Gennes, CS 10030, 92761 Antony Cedex, France

Correspondence to: Alban de Lavenne (alban.de-lavenne@irstea.fr)

Published: 12 May 2016

Abstract. Ideally, semi-distributed hydrologic models should provide better streamflow simulations than lumped models, along with spatially-relevant water resources management solutions. However, the spatial distribution of model parameters raises issues related to the calibration strategy and to the identifiability of the parameters. To analyse these issues, we propose to base the evaluation of a semi-distributed model not only on its performance at streamflow gauging stations, but also on the spatial and temporal pattern of the optimised value of its parameters. We implemented calibration over 21 rolling periods and 64 catchments, and we analysed how well each parameter is identified in time and space. Performance and parameter identifiability are analysed comparatively to the calibration of the lumped version of the same model. We show that the semi-distributed model faces more difficulties to identify stable optimal parameter sets. The main difficulty lies in the identification of the parameters responsible for the closure of the water balance (i.e. for the particular model investigated, the intercatchment groundwater flow parameter).

1 Introduction

1.1 What hydrological good sense suggests

Developing modelling tools that help to understand the spatial distribution of water resources is a key issue for better management. The dynamics of streamflow depends on (i) the spatial variability of precipitation (which, a priori, should be better handled by a semi-distributed hydrological model), (ii) the heterogeneity of catchment behavior (which can be dealt explicitly with by spatially-variable model parameters), and, increasingly, (iii) localized human regulations (for instance, water reservoirs). Since calibration is generally based on discharge measurements at the outlet of the catchment only, and gauging stations are not available everywhere, semi-distributed hydrological models are often difficult to parameterize. As argued by Pokhrel and Gupta (2011), difficulties are due to the smoothing effect of catchments and to the dispersive effect of flow routing combined with numerical issues and measurement uncertainty. The authors state that the impact of spatial variability could become “virtually

non-detectable by conventional performance measures by the time the water reaches the catchment outlet”.

This raises the need to better understand how well parameters are identified in a semi-distributed model compared to a lumped model. The variability of catchment model parameters calibrated over different periods (“time variability”) is one way of approaching this question. Indeed, as reminded by Merz et al. (2011), parameters of a rainfall–runoff model are supposed to represent stable catchment conditions, while the time-varying conditions are supposed to be triggered by the time-series of meteorological inputs. Thus, optimised parameter values should not be overly sensitive to changes of climatic conditions, and one would expect a semi-distributed model to be more stable than a lumped one (because the parameters of the lumped model would have to account implicitly for changing spatial precipitation patterns).

1.2 What the literature says

However, literature provides many examples showing that this assumption is hardly satisfied. Merz et al. (2011) raise two main difficulties. First, problems in model structure and

data measurements tend to be compensated by calibration. For instance, Wagener et al. (2003) identify inconsistencies in the structure of a rainfall–runoff model by highlighting instabilities of the optimal values of the parameters between periods with and without rainfall events. In contrast, Juston et al. (2009) found that data subset of their input data (from daily to quarterly sampling intervals) can provide very similar constraints on model calibration and parameter identification.

Secondly, conditions of the catchment itself may change over time, which consequently, and understandably, shifts optimal model parameters, and justifies rigorous evaluations of the model robustness (Thirel et al., 2015). For instance, changes of land use might directly impact optimal parameter sets (Andréassian et al., 2003; Brown et al., 2005; Versteegen et al., 2016). Trends on parameters can also be related to changes in climatic conditions (Merz et al., 2006; Merz and Blöschl, 2009). On a study based on 273 Austrian catchments, Merz et al. (2011) attribute a doubling of the parameter that controls runoff generation in their model to be related to hydrological changes (such as higher evapotranspiration and drier catchment conditions) rather than calibration artifacts. For about one third of 17 African catchments, Niel et al. (2003) find their model parameters to be unstable, but they did not identify any climate-related reason. Wilby (2005) analyses this question in the context of climate change impact assessment for the River Thames. The author finds model parameters to be highly sensitive to training periods, and recommends the quantification of those large uncertainties due to parameter instability, identifiability and non-uniqueness. Similarly, Brigode et al. (2013) found that the uncertainty due to the climate characteristics of the calibration period is larger than the uncertainty in the estimation of parameters that is often quantified on the basis of Bayesian inference. They attribute this finding to a lack of robustness of the model and recommend more efforts to be put into this aspect.

1.3 Scope of the paper

This paper investigates the procedure of parameter identifiability in a semi-distributed model by comparing model calibration schemes and results with a lumped model on which it is based. From this comparison, we address two main questions: (1) Does spatial distribution of parameters interfere with parameters identifiability? Indeed, one could hope that applying parameters to a more geographically-limited area tends to facilitate their identification. (2) What are the parameters that are the most variable in the lumped and in the semi-distributed models? In this way, we aim to diagnose which components of the model are the least robust, in the sense that their parameterisation is difficult to transpose in time and space.

2 Material and methods

2.1 Study area and hydro-meteorological data

The model is implemented in Eastern France, close to the border with Germany, over 64 sub-catchments of French tributaries of the River Rhine (Fig. 1), namely the rivers Moselle, Sarre and the smaller tributaries of the Rhine located in the Vosges massif. Catchments size vary from 27 to 770 km² and this represents a total area of about 4340 km². Climate is predominantly oceanic with continental influence. Annual precipitation (P) varies from about 700 mm in the lowland to about 1600 mm in the Vosges massif. Average daily temperature (T) and potential evapotranspiration (PE) in the catchment vary from 7 to 10 °C and 540 to 690 mm, respectively.

The two hydrological models implemented (lumped and semi-distributed) require daily time series of P and PE as input data. We used climate data from the SAFRAN meteorological reanalysis of Météo-France (Vidal et al., 2010), which is provided on a square grid of 8 km × 8 km. Discharge data were extracted from the French Hydro database (Leleu et al., 2014) at the daily time step. They were used to perform the calibration and the evaluation of the models. The study period is 1971–2000.

2.2 The GR5J lumped and GRSD semi-distributed rainfall–runoff models

The GRSD semi-distributed rainfall–runoff model was developed by Lerat (2009) and Lobligeois (2014). It is based on the GR5J lumped model (Fig. 2) proposed by Le Moine (2008), which has five free parameters to calibrate (Table 1). The main components of the model are two stores: a production store (with maximum capacity X_1) and a routing store (capacity X_3), which is filled by the output of a unit hydrograph (of time base X_4). Two other parameters, X_2 and X_5 , are used to quantify the intercatchment groundwater flows (IGF). In order to account for snow accumulation and melt, the model is combined with a degree-day snow module (Valéry et al., 2014), which contains two additional parameters (C_{TG} and K_f). However, in our study, these parameters were not calibrated and fixed at their default values, respectively set at 0.2 and 4.5 mm °C⁻¹, as proposed by previous studies in France (Valéry et al., 2014).

The semi-distributed model is applied on sub-catchments. The delineation of sub-catchments is performed only at gauging stations, which means that discharge measurements are available for every hydrological units of the model. The lumped GR5J model is applied on hydrological units composed of upstream catchments (headwater catchments) or intermediate sub-catchments (drained area between downstream and upstream stations). In that way, each hydrological unit receives its own meteorological inputs (P and PE) and uses a distinct parameter set (see Sect. 2.3.2).

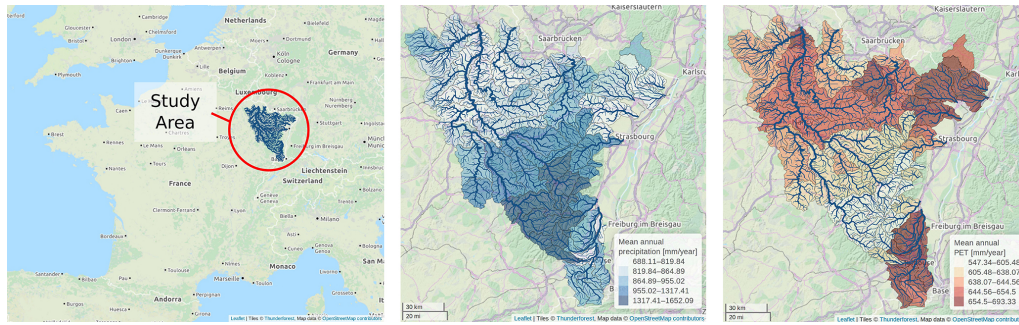


Figure 1. Location of the study area (left), and average annual precipitation (centre) and annual potential evapotranspiration (right) for each sub-catchment (climate data are estimated from the 1971–2000 SAFRAN database).

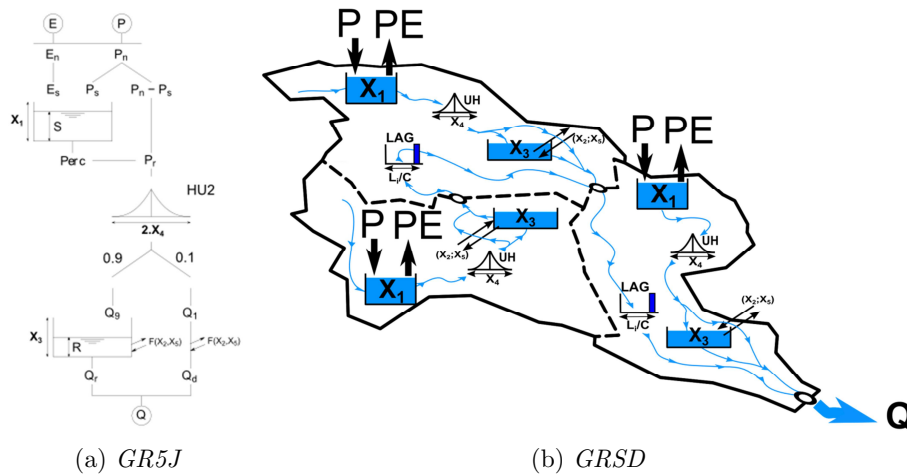


Figure 2. Schematic representation of the GR5J and GRSD semi-distributed model (from Lobligeois et al., 2014).

Table 1. List of the parameters for the semi-distributed conceptual rainfall–runoff model GRSD.

Model parameter	GR5J	GRSD	Description
X1	Free	Free	Production store capacity [mm]
X2	Free	Free	Groundwater exchange coefficient [mm d ⁻¹]
X3	Free	Free	Routing store capacity [mm]
X4	Free	Free	Time base of the unit hydrograph [d]
X5	Free	Free	Threshold for groundwater exchange [–]
C	–	Free	Average streamflow velocity [ms ⁻¹]
C _{TG}	Fixed	Fixed	Ponderation coefficient of the snow thermic state [–]
K _f	Fixed	Fixed	Degree-day factor [mm °C ⁻¹]

The outflow of each GR5J model is finally routed to its downstream catchment using a linear lag propagation model (Bentura and Michel, 1997). Previous studies have shown that this propagation model gives a satisfactory level of efficiency compared to more sophisticated channel routing methods (Lobligeois et al., 2014). This routing functionality implies an additional free parameter (compared to the lumped model GR5J) that needs to be calibrated on each hydrological unit: the average flow velocity C .

2.3 Methodology

2.3.1 Goodness of fit criteria

In order to quantify the agreement between simulations (S) and observations (O), we used the Kling–Gupta Efficiency (KGE) (Gupta et al., 2009), which is based on a decomposition of the Nash–Sutcliffe efficiency (Nash and Sutcliffe, 1970). Moreover, in order to evaluate performances on high and low flows, we used an objective function KGE_m composed of two criteria (Eq. 1): a KGE applied on discharge values to emphasize high flows and a KGE applied on inverse discharge values to emphasize low flows. Both criteria

are applied on the selected discharge time series. Similarly to the KGE criteria, KGE_m varies between $-\infty$ and 1, which is its optimal value.

$$KGE_m(S, O) = 0.5 \cdot \left(KGE(S, O) + KGE\left(\frac{1}{S+\epsilon}, \frac{1}{O+\epsilon}\right) \right), \quad (1)$$

where O and S are the observed and simulated discharges. In order to face numerical problems in case of zero discharge when using the inverse transformation, an ϵ constant is used and set to 1 % of the mean value of O (Pushpalatha et al., 2012).

2.3.2 Calibration strategy of the semi-distributed model

Following Lerat et al. (2012), we performed a multi-site calibration of the GRSD semi-distributed model. Streamflow data at interior points are used to calibrate the model at one outlet. Each intermediate catchment is allowed to have a different parameter set. This is done sequentially, from upstream to downstream points: once the upstream catchment is calibrated, its parameters remain fixed during the calibration of the downstream intermediate catchment.

Sequential calibration is a common strategy for semi-distributed models (see e.g., Andersen et al., 2001; Moussa et al., 2007). It needs as much calibration runs as there are interior points. However, it only uses successive single objective functions, rather than using multi-response objective function to optimize every interior points simultaneously.

2.3.3 Rolling calibration periods

Similarly to Coron et al. (2012), we calibrated the parameters using 10-year long consecutive periods between 1971 and 2000, and used the rest of the time series (20 years) for validation (Fig. 3). This is equivalent to 21 split-sample tests (Klemeš, 1986) performed every year.

Following the work of Merz and Blöschl (2009) and Merz et al. (2011), this enables to provide 21 parameter sets for each of the 64 catchments in order to analyse the temporal and spatial changes of the calibrated parameters. This testing strategy is applied to both models, the lumped model and the semi-distributed model. Parameter variability can thus be compared between the two modelling strategies.

3 Results and discussion

3.1 Performance of the streamflow simulations

The comparison of the goodness-of-fit between the lumped GR5J model and the semi-distributed GRSD model shows slightly better results of the lumped model during calibration and identical results during validation (Fig. 4). One would expect higher performances of the semi-distributed model because it accounts for spatial heterogeneities and nonlinearities that can influence the response of the system. How-

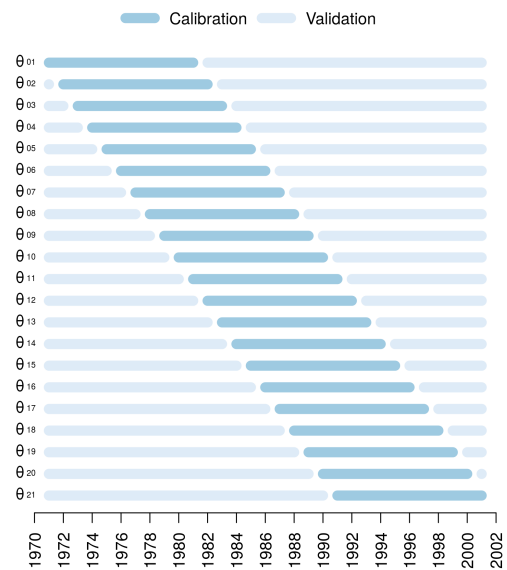


Figure 3. Illustration of the rolling calibration period methodology: 21 parameter sets θ can be identified for each catchment.

ever, literature provides numerous examples of similar results (e.g., Reed et al., 2004), where lumped models perform better. To explain such behaviour, calibration strategies, which are not as well defined for semi-distributed models as for lumped models, are often pointed out (Pokhrel and Gupta, 2011).

As expected, performance on upstream catchments are similar between the lumped and the semi-distributed catchments. Indeed, for those catchments, models are strictly identical (Sect. 2.2). Minor differences can be explained by calibration artifacts and by the fact that both models are implemented in two different modelling environments.

We did not detect any significant performance trends in time. Calibration performances are rather stable, whereas validation performances are subject to more fluctuations. These are similar between the lumped and the semi-distributed models. Results illustrate that both models are potentially able to produce stable efficiency (KGE_m) all along 1971–2000 period, but each calibration period does not provide the same robustness (as observed by validation).

3.2 Temporal trends and variability of parameters

Here, we compare the temporal trends and the variability of parameters, and differences between the two models. The distribution of parameter values according to the 21 calibration periods is given in Fig. 5.

As expected from the structure of the models (Sect. 2.2), upstream catchments have similar parameter sets for both models, whereas different optimum parameter values are obtained for downstream catchments. The main differences concern the parameter X1 (capacity of production store), which is higher for the lumped model than for the semi-

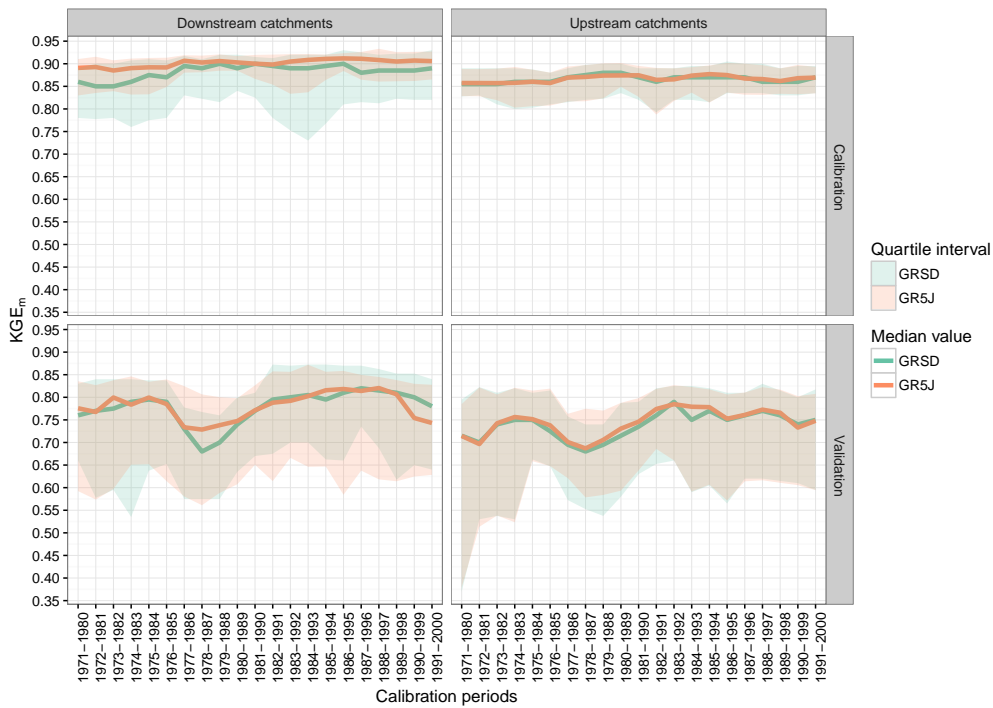


Figure 4. Quantile values of the goodness-of-fit (KGE_m) during calibration and validation over the 64 catchments for the lumped model (GR5J) and the semi-distributed model (GRSD) along the 21 calibration periods. Upstream catchments are all headwater hydrological units, and downstream catchments are all the others.

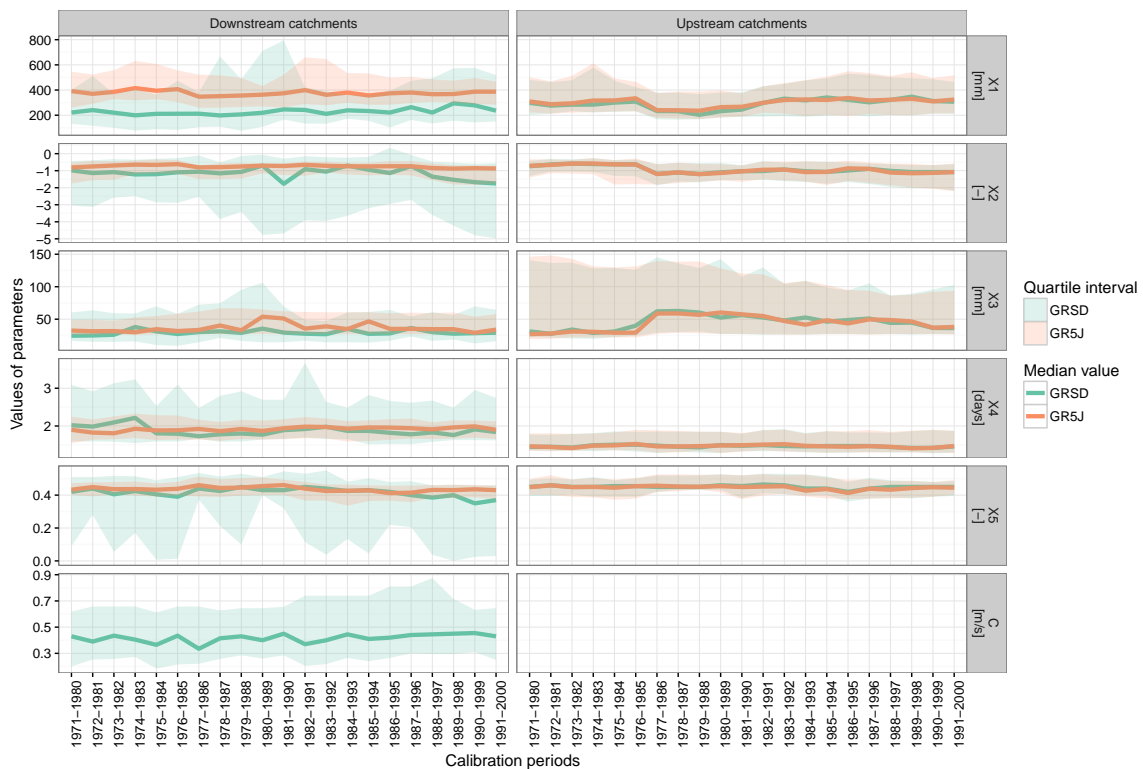


Figure 5. Quantile values of optimised parameter values according to the calibration periods for the 64 catchments. Upstream catchments are all headwater hydrological units, and downstream catchments are all the others.

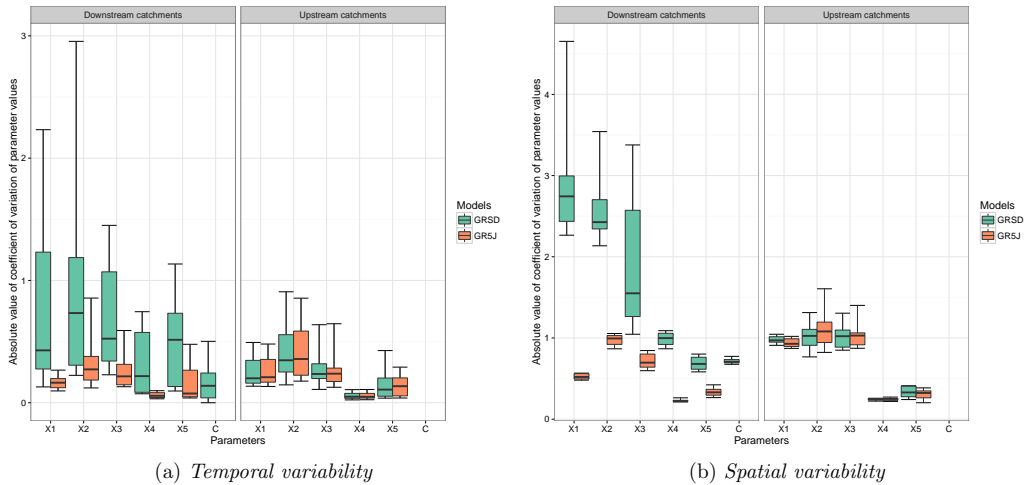


Figure 6. Variability of parameters values among calibration periods within one catchment (64 catchments summarized by boxplots, **a**) and variability of parameters values among catchments within one calibration period (21 calibration periods summarized by boxplots, **b**). Boxplot limits describe the 10th, 25th, 50th, 75th and 90th quantiles.

distributed model. This means a more dynamic response of the downstream hydrological units of the GRSD model and lower evapotranspiration losses. The smoothing of the hydrographs may, in fact, be achieved by the succession of responses of the sub-catchments, from upstream to downstream.

The smaller production store in GRSD appears to be compensated by X2 and X5. Indeed, those parameters aim to quantify intercachment groundwater flows (IGF), which is the amount of water that daily gets out/in of the catchment to fill/empty the routing store and the direct flow component. The X2 parameter quantifies IGF according to a linear relation with the routing store rate ($\frac{S_{\text{rout}}(t)}{X3}$), whereas X5 allows changing the sign of IGF during the year (Eq. 2). In GRSD, where we observe greater negative values of X2 and smaller values of X5 (Fig. 5), IGF may be higher to compensate lower evapotranspiration losses in the production store.

$$\text{IGF}(t) = X2 \cdot \left(\frac{S_{\text{rout}}(t)}{X3} - X5 \right) \quad (2)$$

By looking at parameter values according to the calibration periods, a relative stability of the median value appears among catchments. Only the parameters X2 and X5 in GRSD tend to slowly decrease. These trends are not observed in the lumped model, whose parameters appear more stable for downstream catchments.

From the relative stability in Fig. 5, one could conclude about overall relative robustness of the calibration. However, this does not evaluate parameter stability for each catchment. Therefore, for each hydrological unit of both models, we calculated the coefficient of variation of the parameter values using the 21 parameter sets (Fig. 6).

First, results clearly show a higher temporal variability of the parameters of the semi-distributed model, comparatively

to the variability observed with the lumped model (Fig. 6). It appears that limiting the geographical extent of the area on which the parameter set is applied does not facilitate its identification. One reason for this can come from the calibration strategy, in which one sub-catchment receives simulated outflows from upstream catchments. This upstream volume of water can already represent most of the observed downstream hydrograph. Therefore, parameters applied just on a downstream intermediate sub-catchment might have a minor impact, with problems of sensitivity. Thus calibration may converge to more unstable values, which brings only small improvements to downstream simulations.

Second, it is shown that the most important temporal instability of parameter values is related to the parameter X2, followed by X3 and X5 parameters. All three parameters are used to quantify IGF (Eq. 2). If parameter X2 is showing the most important temporal instability for both models, it is not the case of parameter X5 with the lumped model, where it appears to be one of the least variable parameter (just after parameter X4). This result highlights the problems encountered in quantifying IGF. It seems to be the least identifiable parameter in the lumped model (with high variability of parameter X2). The problem of IGF parameter identifiability is even exacerbated with the semi-distributed model. This might be due to the high inter-dependency of parameters in the formulation of IGF (Eq. 2) in the models. Future improvements of the GRSD semi-distributed model should focus on this issue.

The most stable parameter appears to be X4 (time base of unit hydrograph), which is consistent with previous works (Lobligeois, 2014), where it has been shown that this parameter can easily be related to catchment physical characteristics, such as catchment size.

3.3 Spatial variability of parameters

We also analysed the spatial variability of the parameter values, considering variability between parameters and between models. To this end, we estimated the coefficient of variation of the parameter values among catchments (one performed by calibration period). The aim is to quantify how much parameters can be different between catchments.

Similarly to the temporal variability of parameters, spatial variability appears to be higher with the semi-distributed model than with the lumped model. However, contrary to the temporal variability, spatial variability is more expected here, as one of the objectives of a semi-distributed model is precisely to consider those spatial heterogeneities of the hydrological response. However, we noticed again that parameters X1 and X2 are the most variable parameters for the semi-distributed model. Therefore they appear among the most variable parameters for both analyses, the time variability and the space variability analyses. These two parameters control water balance. Similarly to the temporal variability, they are expected to be highly variable in space in order to get along with sequential observations at each downstream station during calibration.

Spatial variabilities are not constant over time (as observed by the boxplot widths on Fig. 6). Particularly, the spatial variability of parameter X3 (capacity of routing store) is stable for the lumped model, whereas it appears to be very dependent on the calibration period for the semi-distributed model. For instance, it has the highest spatial variability during the “1984–1994” calibration period, and is among the lowest variability just 3 years after (not shown). Once again, a robustness problem of GRSD is identified and needs to be addressed in further investigations.

4 Conclusion

In this paper, we compared the spatio-temporal variability of the parameters of a semi-distributed model (GRSD) and a lumped model (GR5J) on which it is based. We applied a rolling calibration strategy over 21 periods and 64 French catchments.

A classical evaluation of discharge simulations using goodness-of-fit criteria was applied to the outputs of both models. It illustrates a slightly better performance of the lumped model during calibration, and similar performance of the models during validation. However, further investigation on parameter identifiability highlighted much higher temporal variabilities of the semi-distributed model. This study also showed that it is more difficult to identify catchment’s specific parameter sets with the semi-distributed model than with the lumped model.

The methodology applied also enabled to identify the more unstable parameters. Results showed that the parameters related to the quantification of intercatchment groundwater flows (IGF) are the most unstable. We conclude that further

modelling efforts should focus on the model structure in order to better quantify IGF.

This work also emphasizes the fact that the calibration strategy and the evaluation approach of a semi-distributed model should not focus only on goodness-of-fit performance, but also on parameter identifiability, especially if the model aims to be used to explore future scenarios in a changing world. Such an approach would also facilitate the application of the model at ungauged locations, since parameters that depict high variability in time and space might be more difficult to regionalize.

Data availability

The streamflow data used in this study are freely available at: <http://www.hydro.eaufrance.fr>. The climatic data are produced by Météo-France. Information can be found at: <http://publitheque.meteo.fr>.

Acknowledgements. The first author was partially funded by the French Waterboard “Agence de l’Eau Rhin-Meuse” within the MOSARH21 project (project #15C92002). Météo-France and SCHAPI are thanked for making climatic and hydrological data available for this study. The authors thank the anonymous reviewer for his comments on the manuscript.

References

- Andersen, J., Refsgaard, J. C., and Jensen, K. H.: Distributed hydrological modelling of the Senegal River Basin – model construction and validation, *J. Hydrol.*, 247, 200–214, doi:10.1016/S0022-1694(01)00384-5, 2001.
- Andréassian, V., Parent, E., and Michel, C.: A distribution-free test to detect gradual changes in watershed behavior, *Water Resour. Res.*, 39, 1252, doi:10.1029/2003WR002081, 2003.
- Bentura, P. L. and Michel, C.: Flood routing in a wide channel with a quadratic lag-and-route method, *Hydrolog. Sci. J.*, 42, 169–189, doi:10.1080/02626669709492018, 1997.
- Brigode, P., Oudin, L., and Perrin, C.: Hydrological model parameter instability: A source of additional uncertainty in estimating the hydrological impacts of climate change?, *J. Hydrol.*, 476, 410–425, doi:10.1016/j.jhydrol.2012.11.012, 2013.
- Brown, A. E., Zhang, L., McMahon, T. A., Western, A. W., and Vertessy, R. A.: A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation, *J. Hydrol.*, 310, 28–61, doi:10.1016/j.jhydrol.2004.12.010, 2005.
- Coron, L., Andréassian, V., Perrin, C., Lerat, J., Vaze, J., Bourqui, M., and Hendrickx, F.: Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments, *Water Resour. Res.*, 48, W05552, doi:10.1029/2011WR011721, 2012.
- Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *J. Hydrol.*, 377, 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.

- HYDRO: <http://www.hydro.eaufrance.fr>, last access: May 2016.
- Juston, J., Seibert, J., and Johansson, P.-O.: Temporal sampling strategies and uncertainty in calibrating a conceptual hydrological model for a small boreal catchment, *Hydrol. Process.*, 23, 3093–3109, doi:10.1002/hyp.7421, 2009.
- Klemeš, V.: Operational testing of hydrological simulation models, *Hydrolog. Sci. J.*, 31, 13–24, doi:10.1080/02626668609491024, 1986.
- Leleu, I., Tonnelier, I., Puechberty, R., Gouin, P., Viquendi, I., Cobos, L., Foray, A., Baillon, M., and Ndima, P.-O.: La refonte du système d'information national pour la gestion et la mise à disposition des données hydrométriques, *La Houille Blanche*, 25–32, doi:10.1051/lhb/2014004, 2014.
- Le Moine, N.: Le bassin versant de surface vu par le souterrain: une voie d'amélioration des performances et du réalisme des modèles pluie-débit?, PhD thesis, Université Pierre et Marie Curie, Paris, France, 348 pp., 2008.
- Lerat, J.: Quels apports hydrologiques pour les modèles hydrauliques? Vers un modèle intégré de simulation des crues, PhD thesis, University of Pierre et Marie Curie, Paris, France, 390 pp., 2009.
- Lerat, J., Andréassian, V., Perrin, C., Vaze, J., Perraud, J. M., Ribstein, P., and Loumagne, C.: Do internal flow measurements improve the calibration of rainfall-runoff models?, *Water Resour. Res.*, 48, W02511, doi:10.1029/2010WR010179, 2012.
- Lobligeois, F.: Mieux connaître la distribution spatiale des pluies améliorées-t-il la modélisation des crues?, Diagnostic sur 181 bassins versants français, PhD thesis, AgroParisTech, Paris, France, 310 pp., 2014.
- Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., and Loumagne, C.: When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events, *Hydrol. Earth Syst. Sci.*, 18, 575–594, doi:10.5194/hess-18-575-2014, 2014.
- Merz, R. and Blöschl, G.: A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria, *Water Resour. Res.*, 45, W01405, doi:10.1029/2008WR007163, 2009.
- Merz, R., Blöschl, G., and Parajka, J.: Spatio-temporal variability of event runoff coefficients, *J. Hydrol.*, 331, 591–604, doi:10.1016/j.jhydrol.2006.06.008, 2006.
- Merz, R., Parajka, J., and Blöschl, G.: Time stability of catchment model parameters: Implications for climate impact analyses, *Water Resour. Res.*, 47, W02531, doi:10.1029/2010WR009505, 2011.
- Météo-France: <http://publitheque.meteo.fr>, last access: May 2016.
- Moussa, R., Chahinian, N., and Bocquillon, C.: Distributed hydrological modelling of a Mediterranean mountainous catchment – Model construction and multi-site validation, *J. Hydrol.*, 337, 35–51, doi:10.1016/j.jhydrol.2007.01.028, 2007.
- Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models part I – A discussion of principles, *J. Hydrol.*, 10, 282–290, doi:10.1016/0022-1694(70)90255-6, 1970.
- Niel, H., Paturel, J.-E., and Servat, E.: Study of parameter stability of a lumped hydrologic model in a context of climatic variability, *J. Hydrol.*, 278, 213–230, doi:10.1016/S0022-1694(03)00158-6, 2003.
- Pokhrel, P. and Gupta, H. V.: On the ability to infer spatial catchment variability using streamflow hydrographs, *Water Resour. Res.*, 47, W08534, doi:10.1029/2010WR009873, 2011.
- Pushpalatha, R., Perrin, C., Le Moine, N., and Andréassian, V.: A review of efficiency criteria suitable for evaluating low-flow simulations, *J. Hydrol.*, 517, 1176–1187, doi:10.1016/j.jhydrol.2011.11.055, 2012.
- Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., and DMIP Participants: Overall distributed model intercomparison project results, *J. Hydrol.*, 298, 27–60, doi:10.1016/j.jhydrol.2004.03.031, 2004.
- Thirel, G., Andréassian, V., and Perrin, C.: On the need to test hydrological models under changing conditions, *Hydrolog. Sci. J.*, 60, 1165–1173, doi:10.1080/02626667.2015.1050027, 2015.
- Valéry, A., Andréassian, V., and Perrin, C.: 'As simple as possible but not simpler': What is useful in a temperature-based snow-accounting routine? Part 2 – Sensitivity analysis of the Cemanège snow accounting routine on 380 catchments, *J. Hydrol.*, 517, 1176–1187, doi:10.1016/j.jhydrol.2014.04.058, 2014.
- Verstegen, J. A., Karssen, D., van der Hilst, F., and Faaij, A. P.: Detecting systemic change in a land use system by Bayesian data assimilation, *Environ. Modell. Softw.*, 75, 424–438, doi:10.1016/j.envsoft.2015.02.013, 2016.
- Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., and Soubeyrou, J.-M.: A 50-year high-resolution atmospheric reanalysis over France with the Safran system, *Int. J. Climatol.*, 30, 1627–1644, doi:10.1002/joc.2003, 2010.
- Wagener, T., McIntyre, N., Lees, M. J., Wheeler, H. S., and Gupta, H. V.: Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis, *Hydrol. Process.*, 17, 455–476, doi:10.1002/hyp.1135, 2003.
- Wilby, R. L.: Uncertainty in water resource model parameters used for climate change impact assessment, *Hydrol. Process.*, 19, 3201–3219, doi:10.1002/hyp.5819, 2005.