



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

03/10/2017

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.

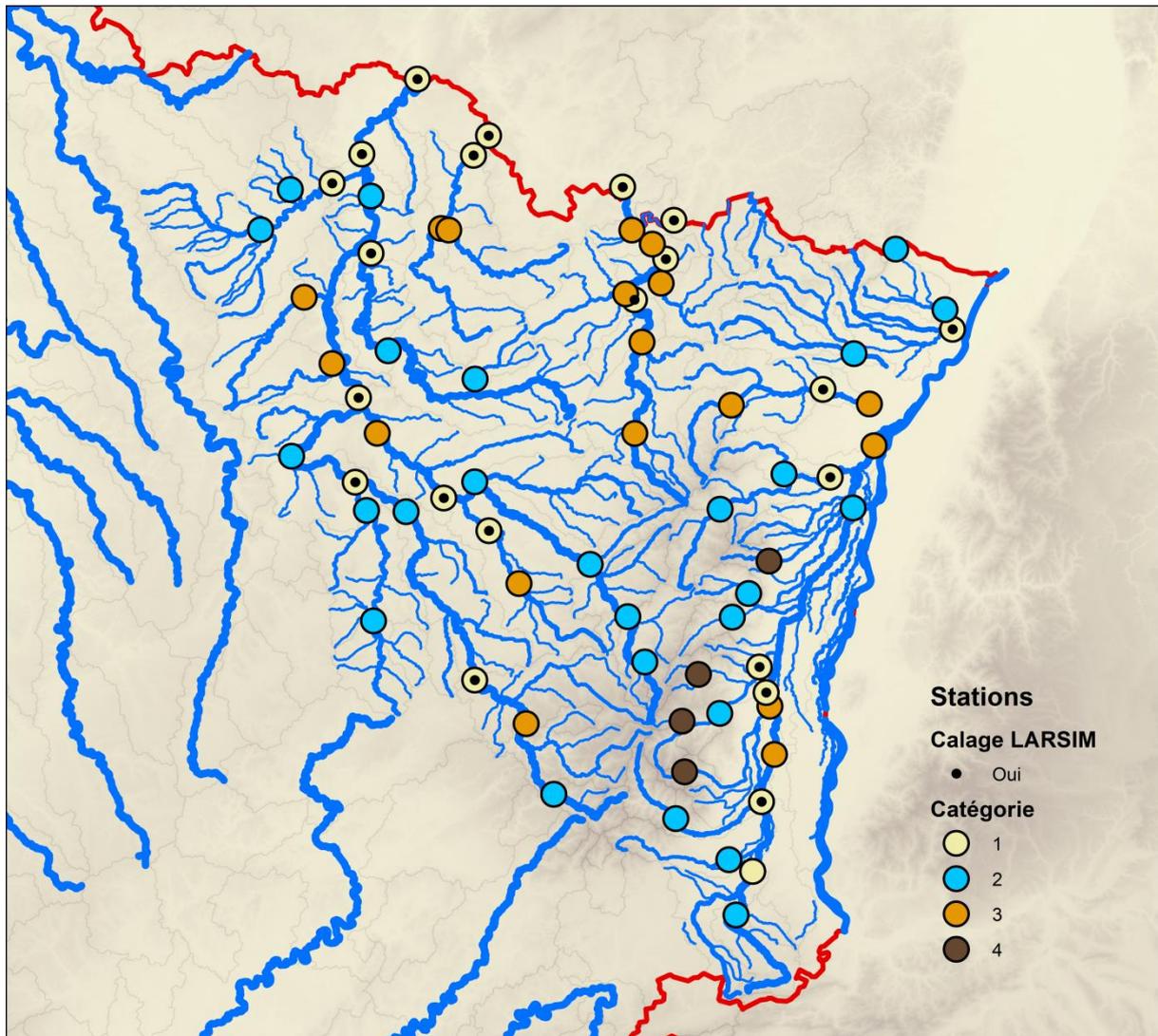


Figure 1 : Liste des stations. Toutes les stations sont calées avec GRSD dans le projet. Les stations calées de manière automatique pour LARSIM apparaissent avec un point noir.

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.

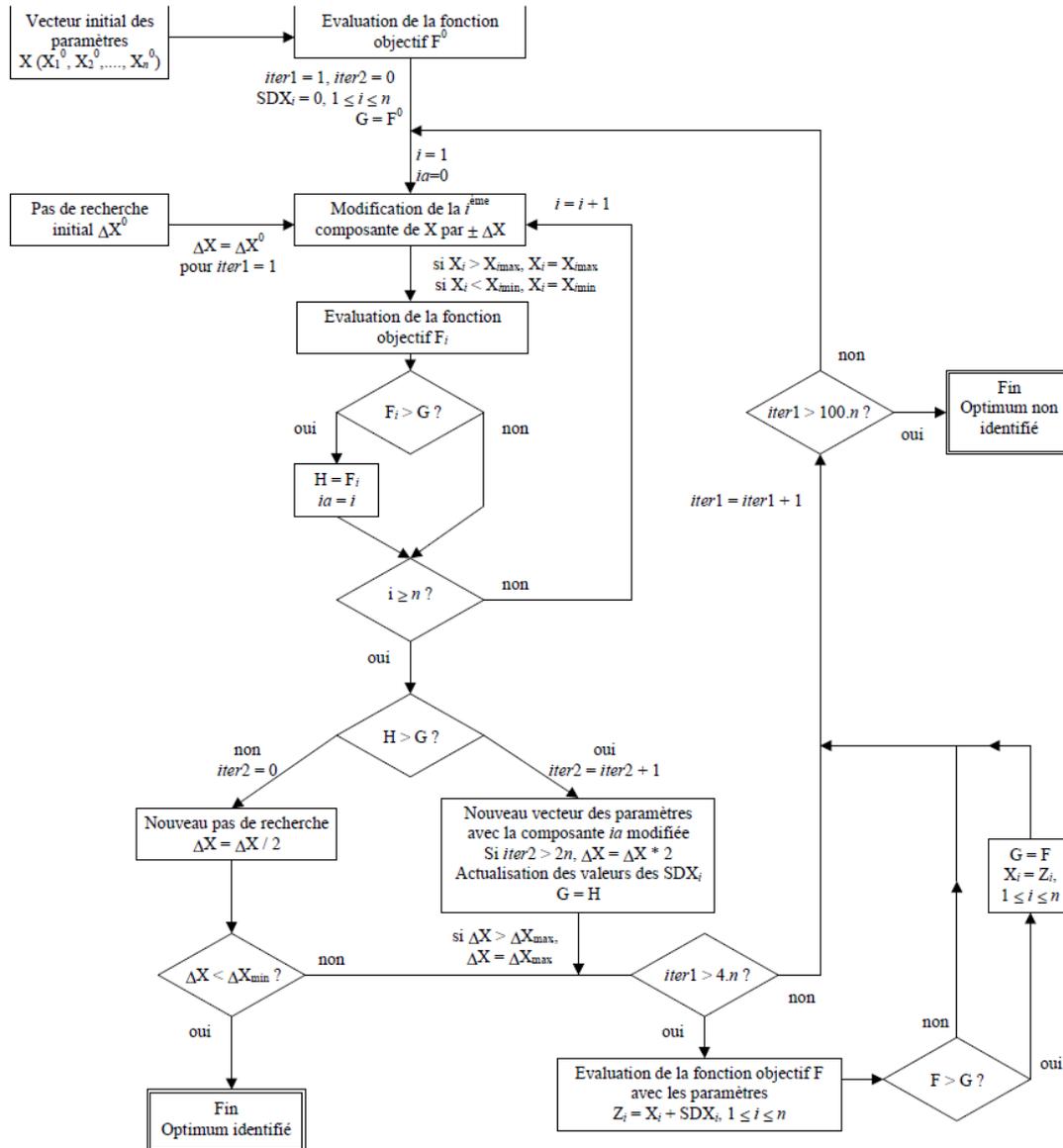


Figure 2 : Diagramme schématique de la méthode pas-à-pas (tiré de Perrin, 2000)

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.

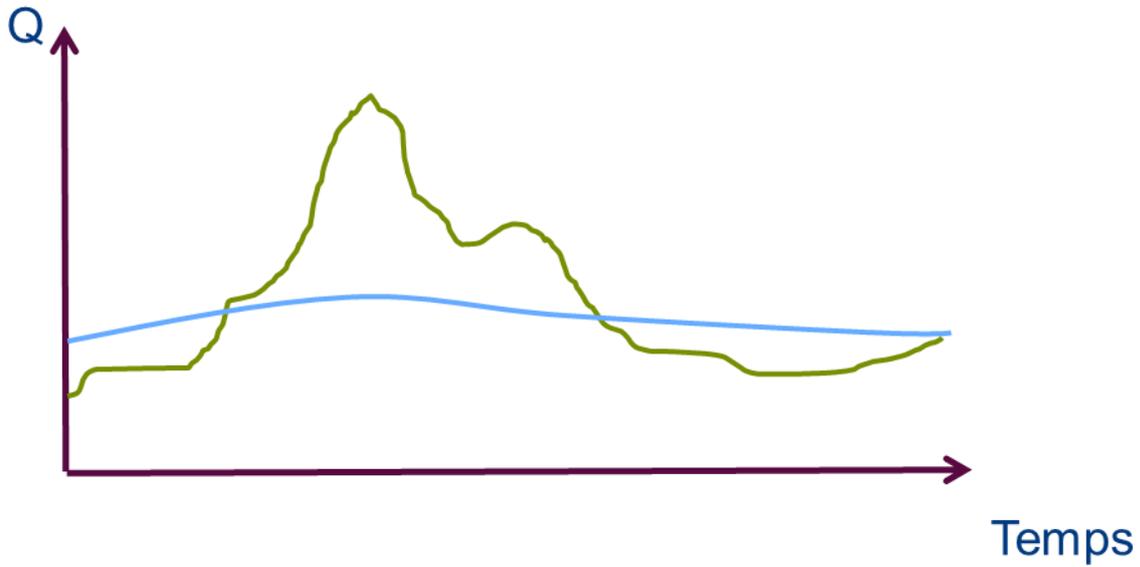


Figure 3 : Exemple d'un débit issu d'un calage complet (en vert) et d'un débit de base (en bleu) dont seul le coefficient KG a été calé

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.

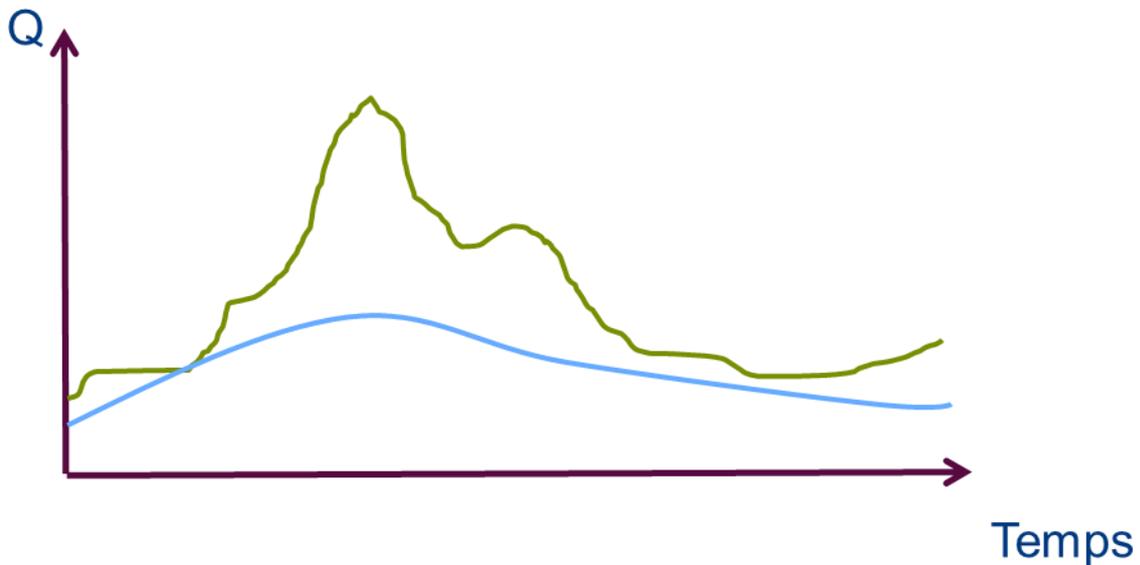


Figure 4 : Résultat du calage des coefficients beta et EQB qui impactent la composante long terme (débit de base) des débits. En vert le débit après calage complet, en bleu le débit de base après calage des deux coefficients

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.

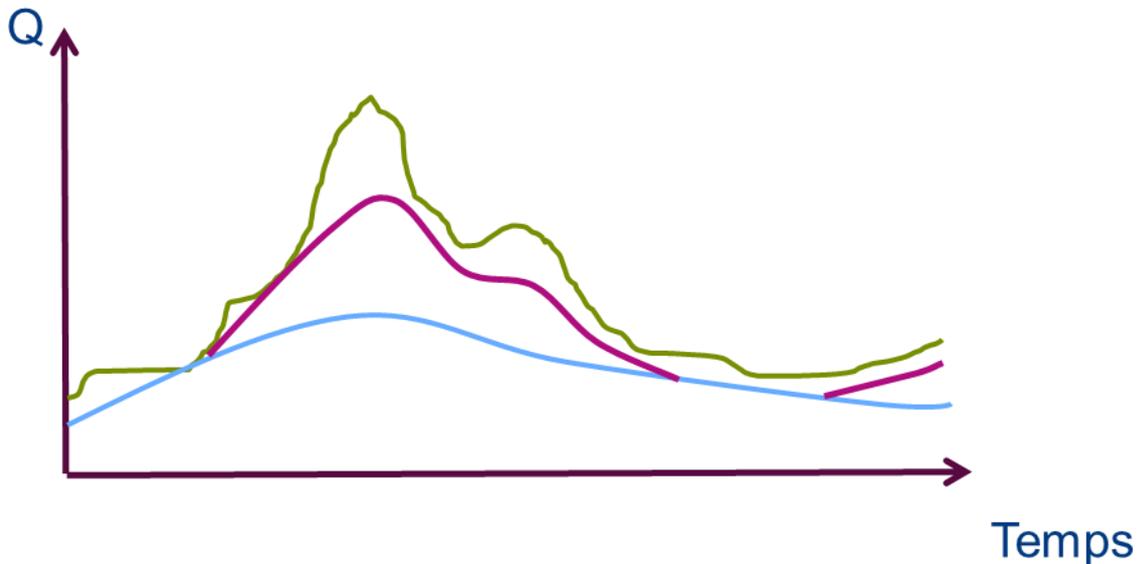


Figure 5 : Résultat du calage des coefficients Dmax et EQI qui impactent la composante intermédiaire des débits. En rose, le débit intermédiaire de LARSIM après calage des deux coefficients

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.

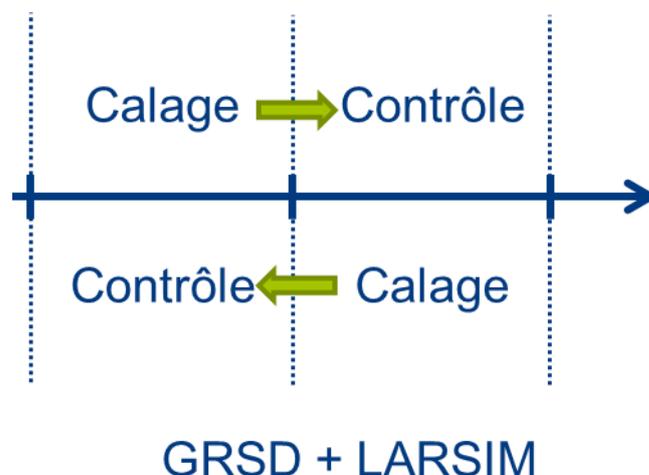


Figure 6 : Procédure de calage-contrôle utilisée pour GRSD et LARSIM

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.

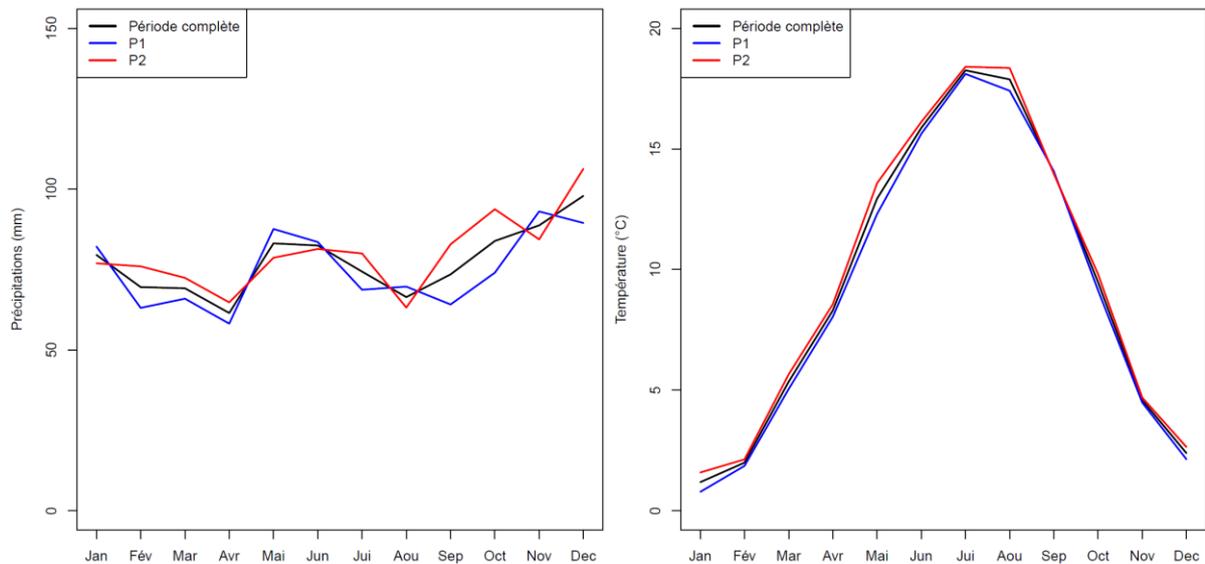


Figure 7 : Evolution mensuelle des précipitations et de la température sur les périodes d'étude (P1 : 1971-1985 ; P2 : 1986-2000) sur l'ensemble de la zone du projet MOSARH21. L'évapotranspiration potentielle suivant les mêmes évolutions que la température, elle ne figure pas sur ces graphes.

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.

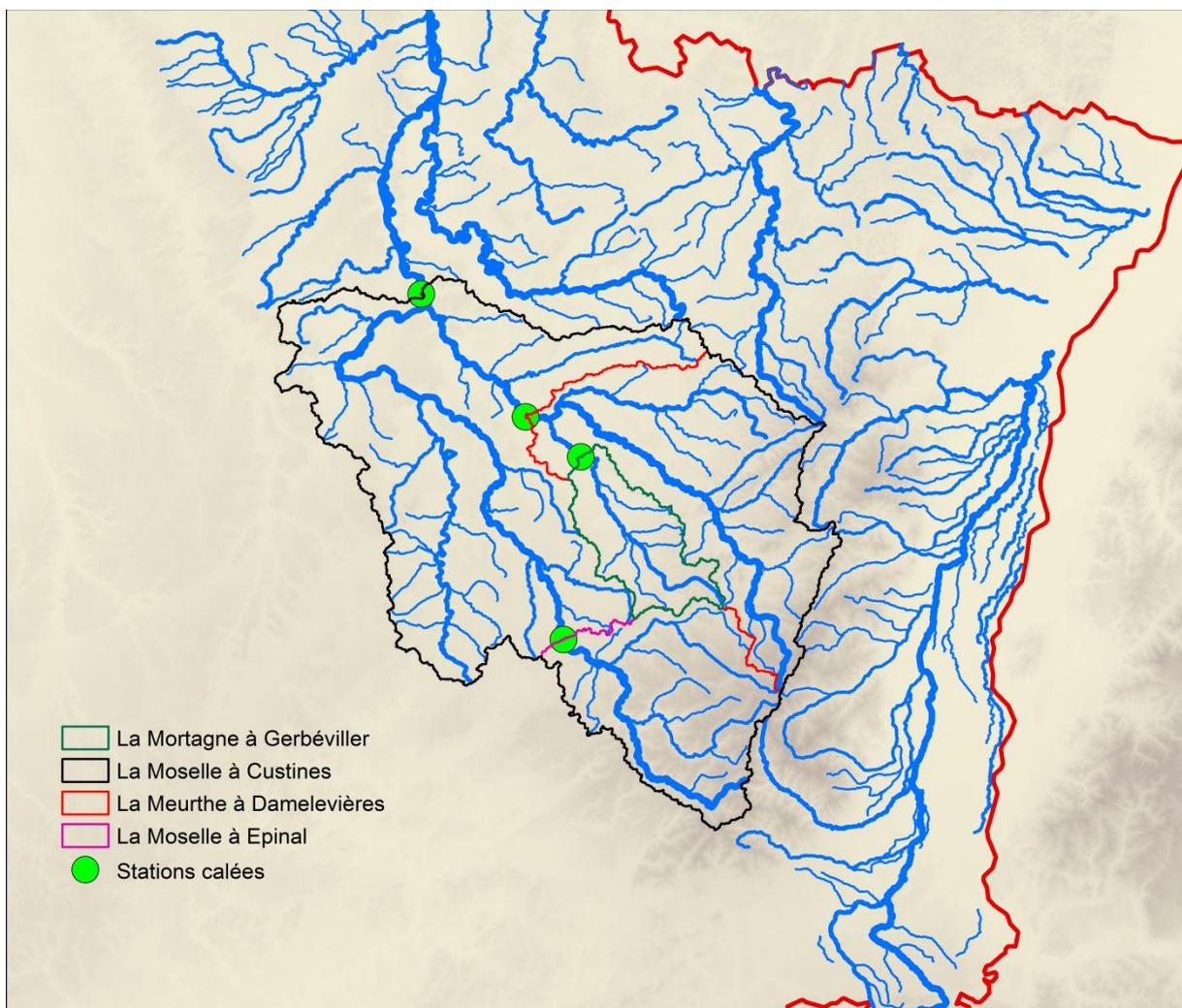


Figure 9 : Bassin versant de la Moselle à Custines. Le test de calage de LARSIM a été effectué sur ces quatre stations.

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.

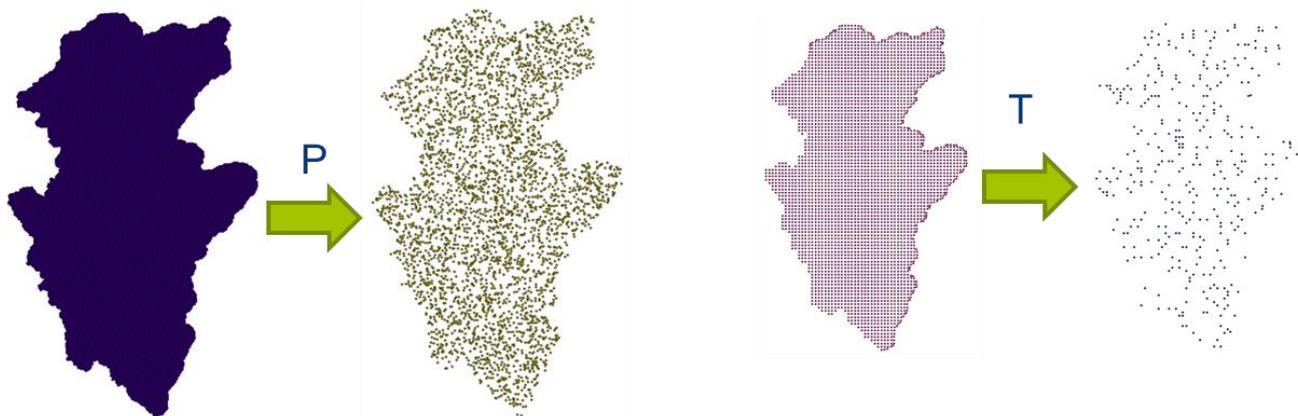


Figure 10 : Impact de la réduction des jeux de données sur les précipitations et les températures

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.

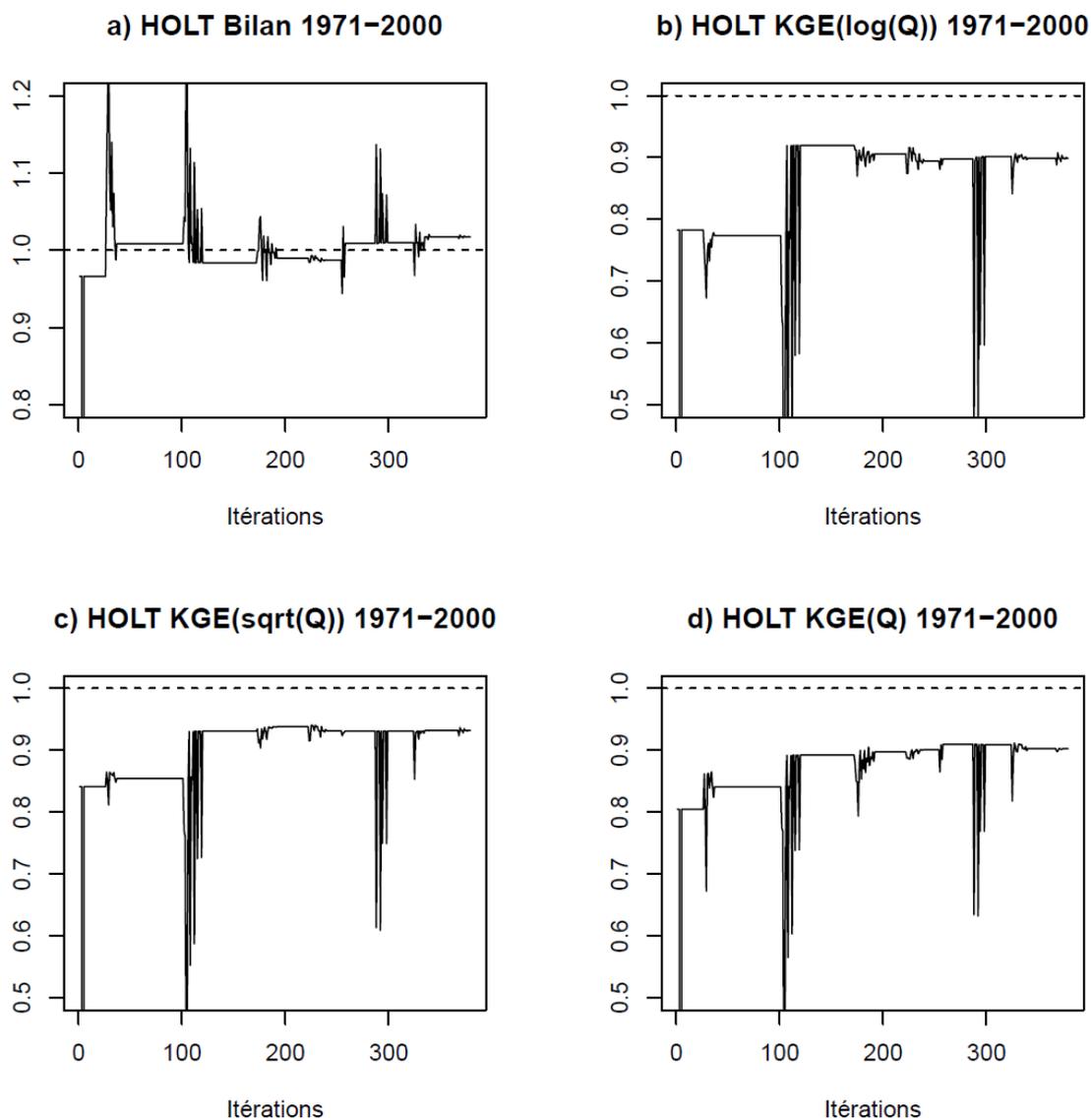


Figure 11 : Evolution des performances de LARSIM à chaque itération successive pour le bilan (a), pour le débit de base (b), pour le débit intermédiaire (c) et pour les hauts débits(d). Exemple de la station de la Bruche à Holtzheim pour un calage sur la période 1971-2000. Les valeurs des critères d'optimisation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

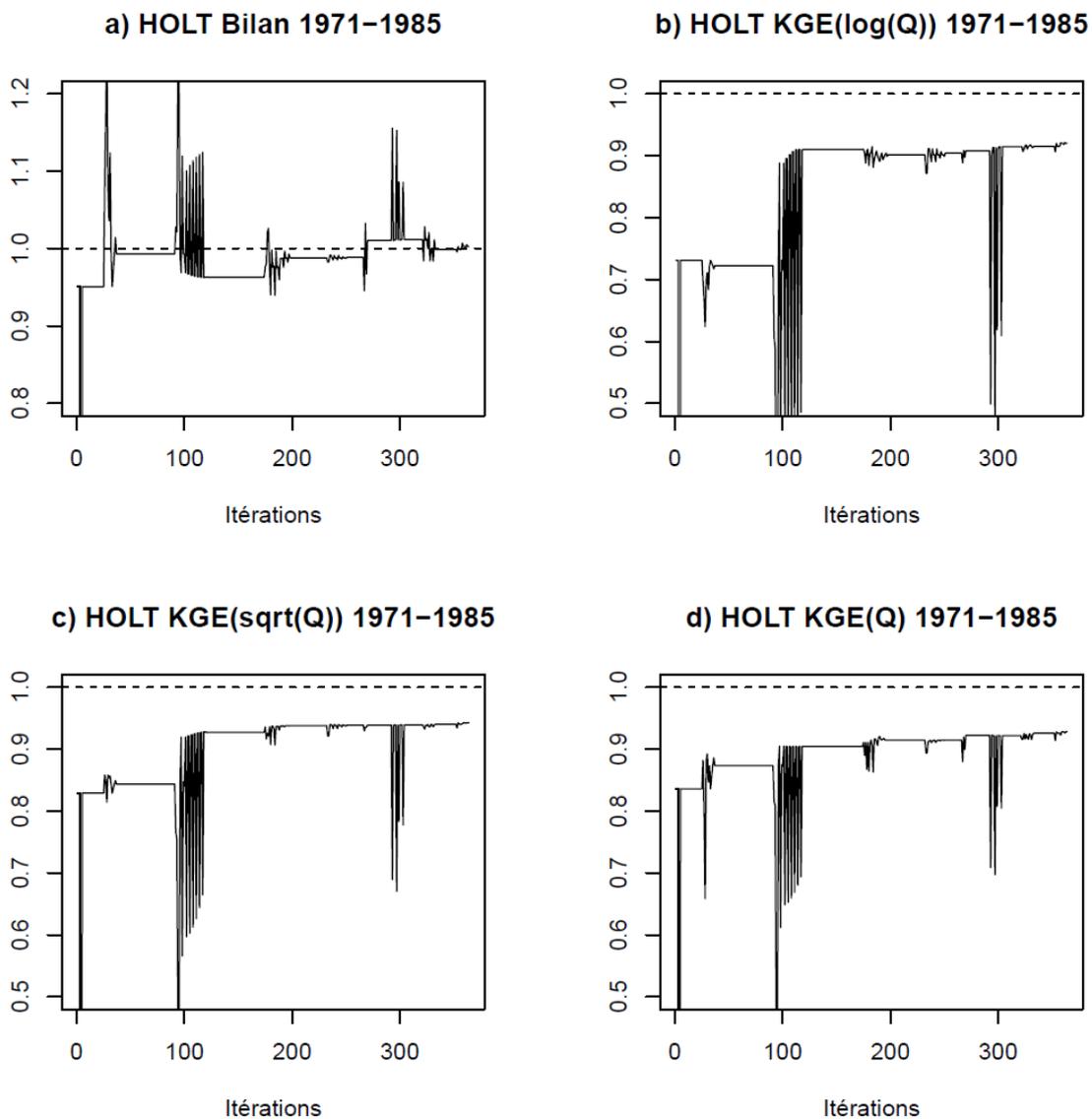


Figure 12 : Idem que la Figure 11, pour un calage et des critères d'optimisation calculés sur la période 1971-1985.

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.

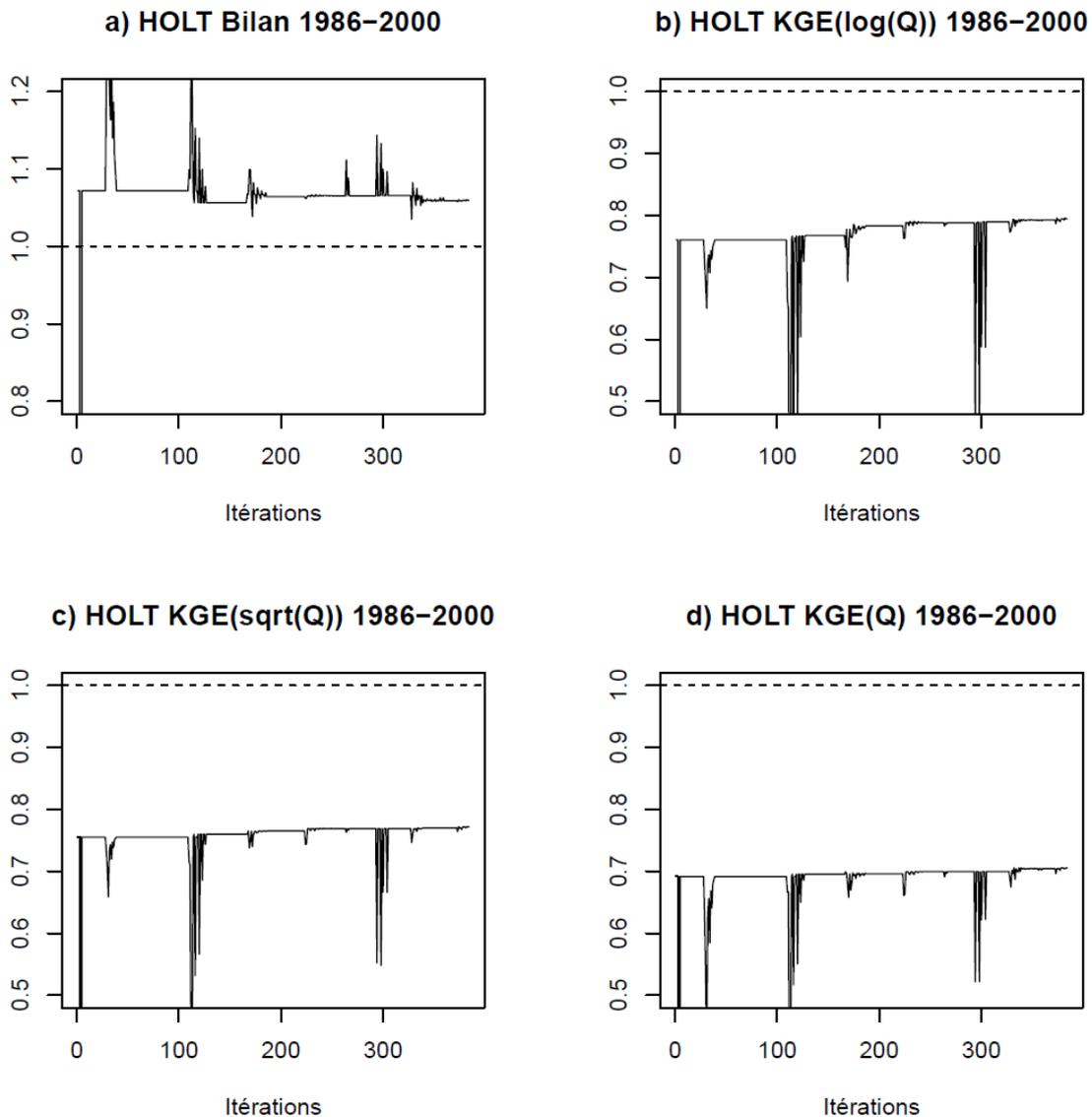


Figure 13 : Idem que la Figure 11, pour un calage et des critères d'optimisation sur la période 1986-2000.

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 (se reporter à l'Annexe E du rapport final). 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.

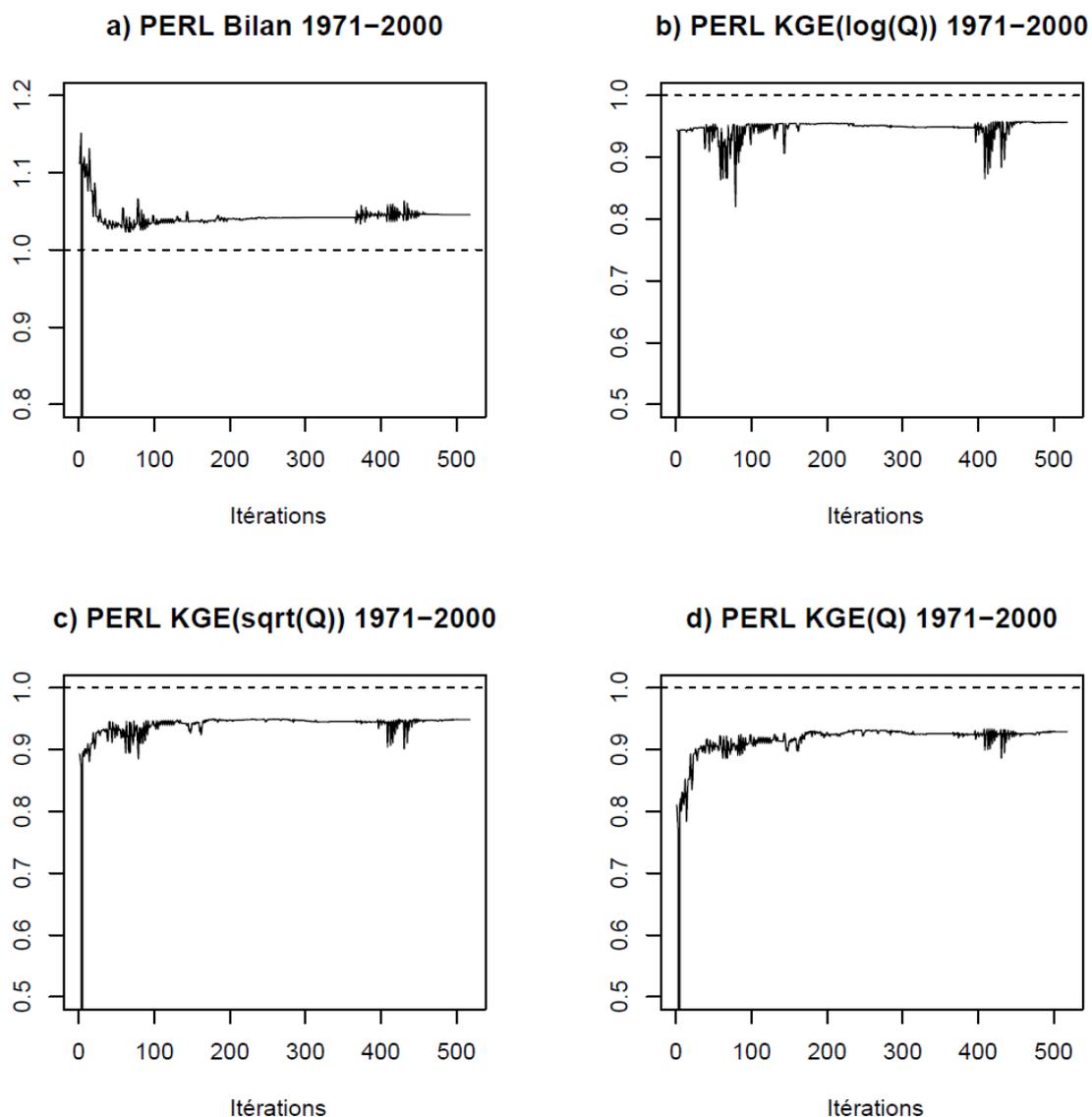


Figure 14 : Evolution des performances de LARSIM à chaque itération successive pour le bilan (a), pour le débit de base (b), pour les débits intermédiaires (c) et pour les hauts débits (d). Exemple de la station de la Moselle à Perl (DE) pour un calage sur la période 1971-2000. Les valeurs des critères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

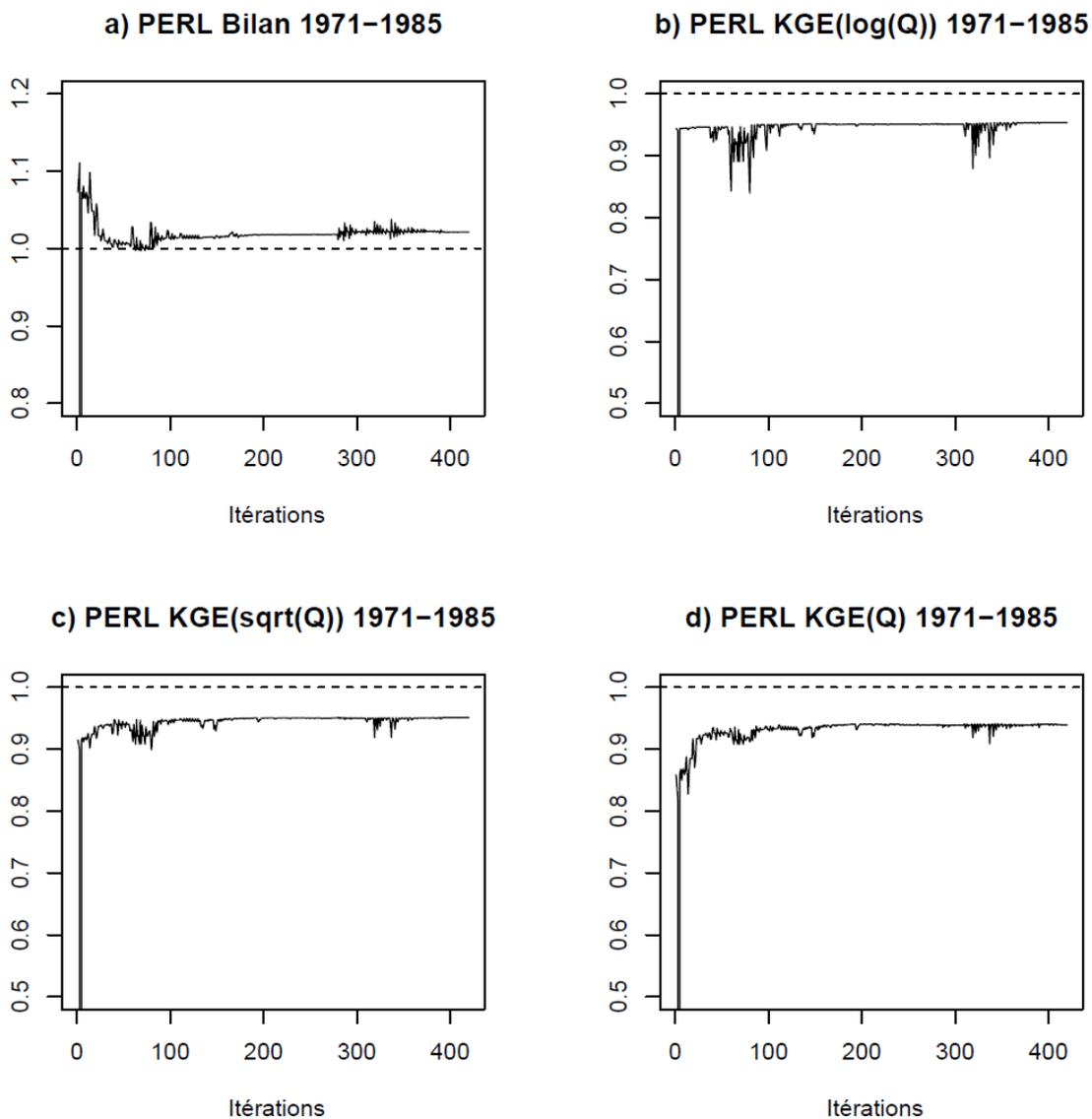


Figure 15 : Idem que la Figure 14, pour un calage et des critères d'optimisation sur la période 1971-1985.

L'évolution des valeurs des sept paramètres de LARSIM calés dans le cadre de MOSARH21 pour l'ensemble des calages effectués est disponible dans l'Annexe E du rapport final. Nous montrons l'un de ces graphes 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.

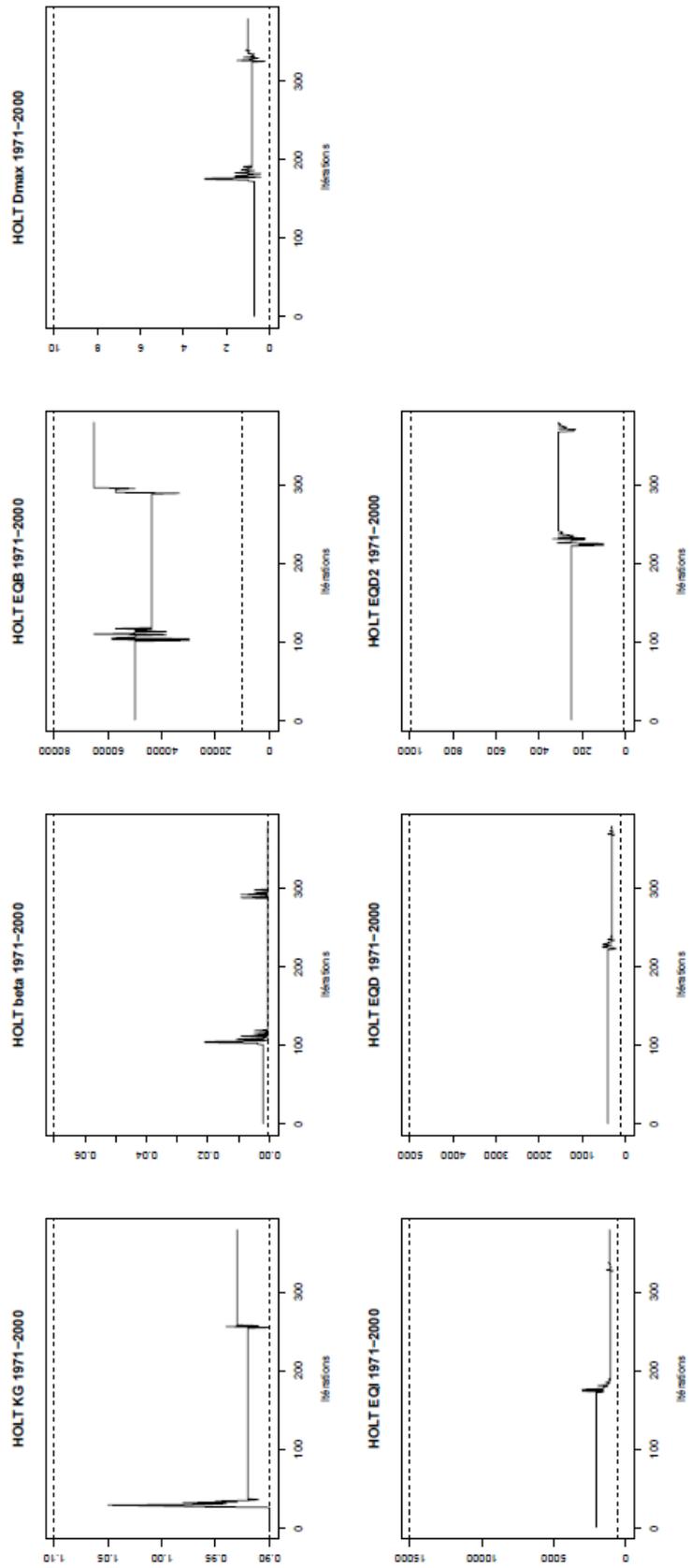


Figure 16 : Evolution des paramètres de LARSIM à chaque itération successive. Exemple de la station de la Bruche à Holtzheim pour un calage sur la période 1971-2000.

3.4. Cartes des paramètres des modèles LARSIM et GRSD

Des cartes de paramètres des deux modèles hydrologiques sont fournies avec ce rapport (voir Annexe F¹).

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.

¹ voir dans les dossiers « Compl_Per », « P01 » et « P02 » qui correspondent aux noms des périodes de calage, les sous-dossiers « map_parameter ». Trois fichiers sont présents à chaque fois : l'un pour le calage de GRSD, un autre pour le calage automatique de LARSIM et le dernier pour le calage manuel de LARSIM. Les valeurs de chacun des paramètres sont représentées sur des cartes différentes, pour les sous-bassins ayant fait l'objet d'un calage. Il est possible de zoomer et de cliquer sur les sous-bassins pour obtenir les valeurs des paramètres.

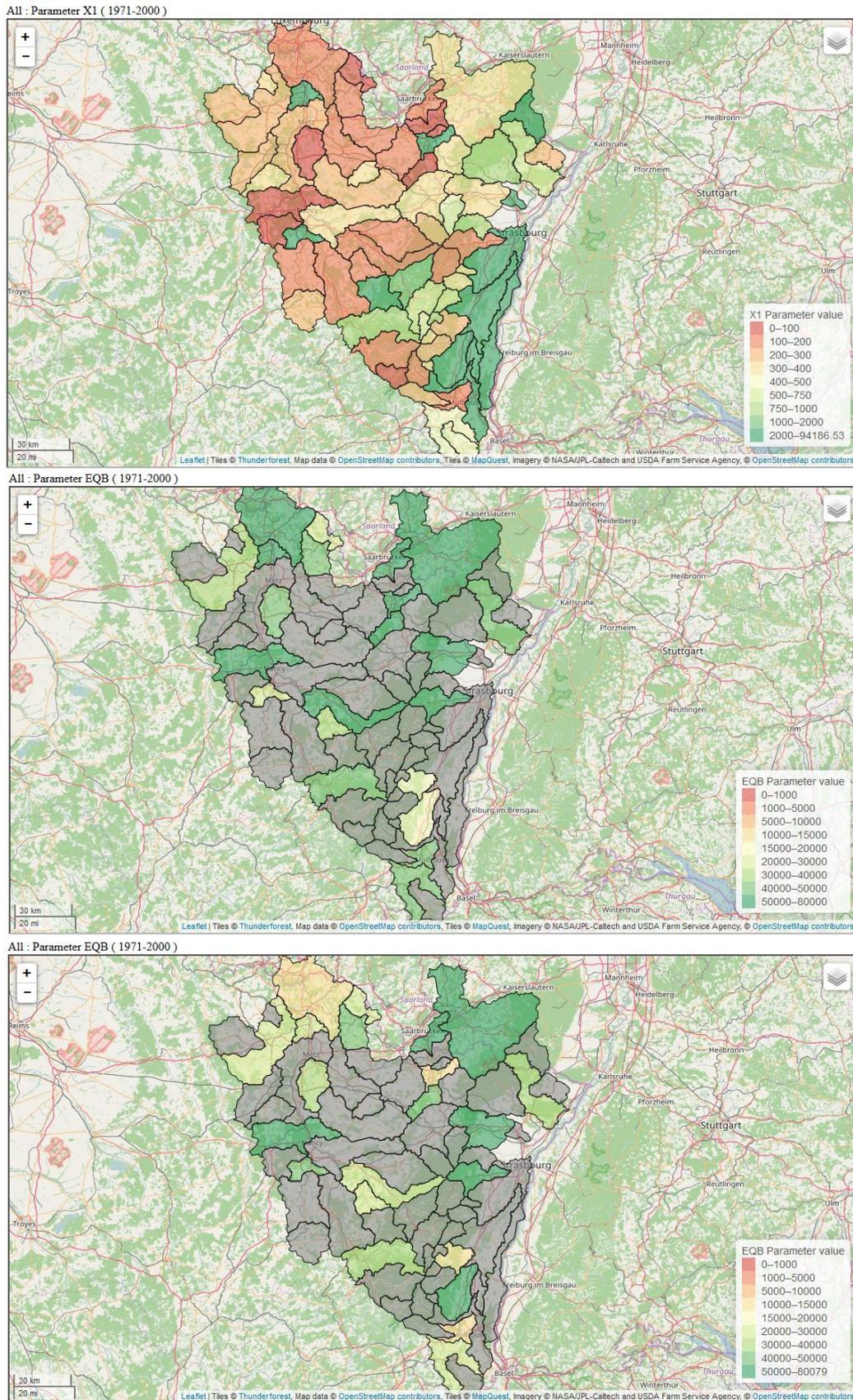


Figure 17 : Cartes des paramètres X1 (pour GRSD, en haut) et EQB (pour LARSIM calage manuel au milieu, LARSIM calage automatique en bas). Il faut noter que le maillage des paramètres de LARSIM ne correspond pas entièrement à celui de GRSD : certains sous-bassins autres que les 70 du projet comprennent 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 (voir Annexe G pour les différents hydrogrammes)². Ils semblent généralement bons, les différences entre les modèles et avec les observations de débit restant raisonnables.

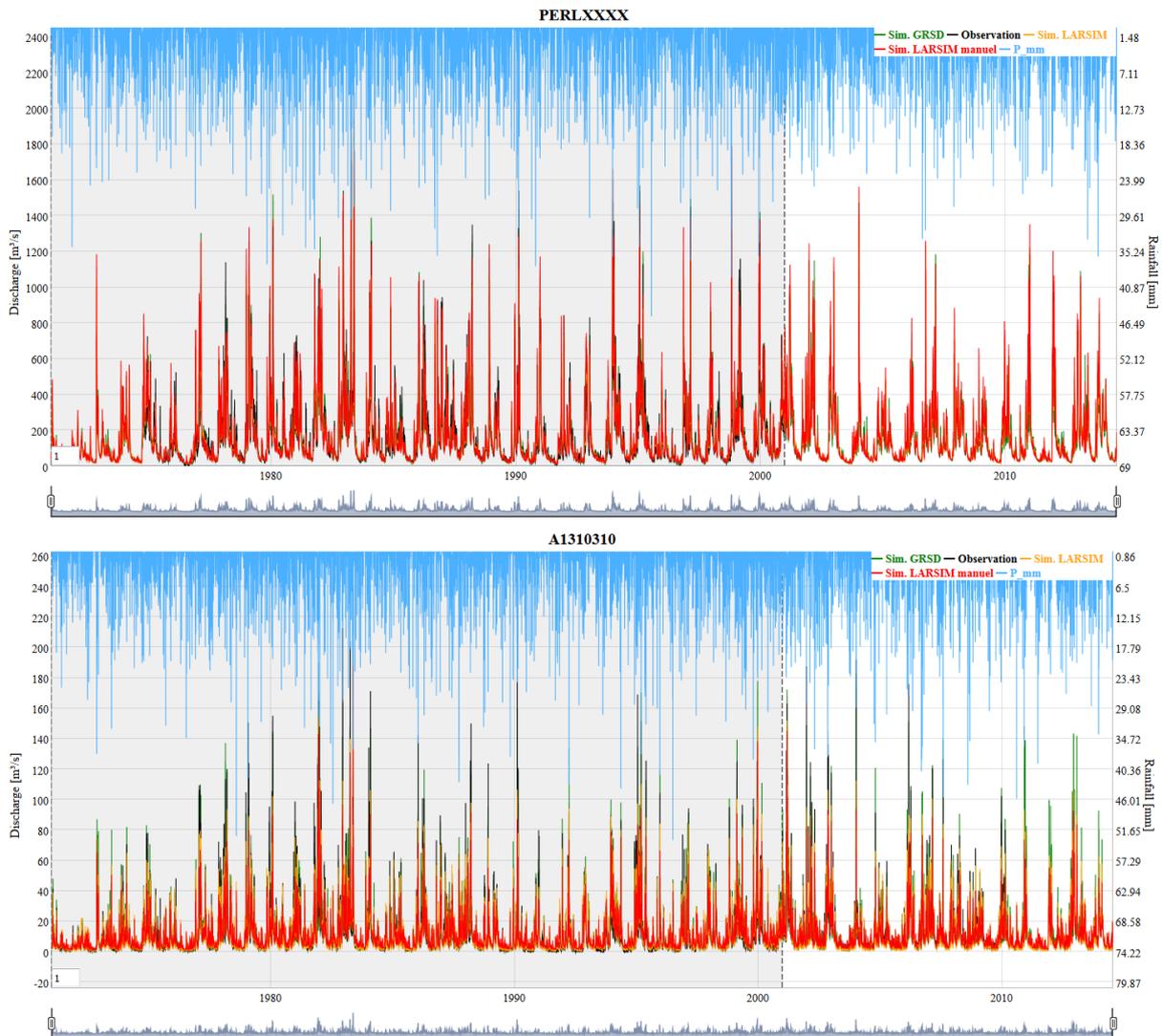


Figure 18 : Hydrogrammes de la Moselle à Perl (en haut) et de l'III à Ensisheim (en bas). En vert, les simulations de GRSD, en orange les simulations de LARSIM calé en automatique, en rouge les simulations de LARSIM calé en manuel, en noir les débits observés, en bleu les précipitations. Les deux modèles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

² Pour une analyse détaillée des performances des modèles en termes de simulations de chroniques de débits, des hydrogrammes, disponibles dans l'Annexe F, sont fournis sous format html (voir dans les dossiers « Compl_Per », « P01 » et « P02 » qui correspondent aux noms des périodes de calage, les sous-dossiers « hydrograph »). Il y a un fichier par station, et chaque fichier comprend les débits observés, les débits simulés par GRSD, les débits simulés par LARSIM calé automatiquement (si disponibles) et les débits simulés par LARSIM calé manuellement. Ces graphiques sous format html permettent d'effectuer des zooms.

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.

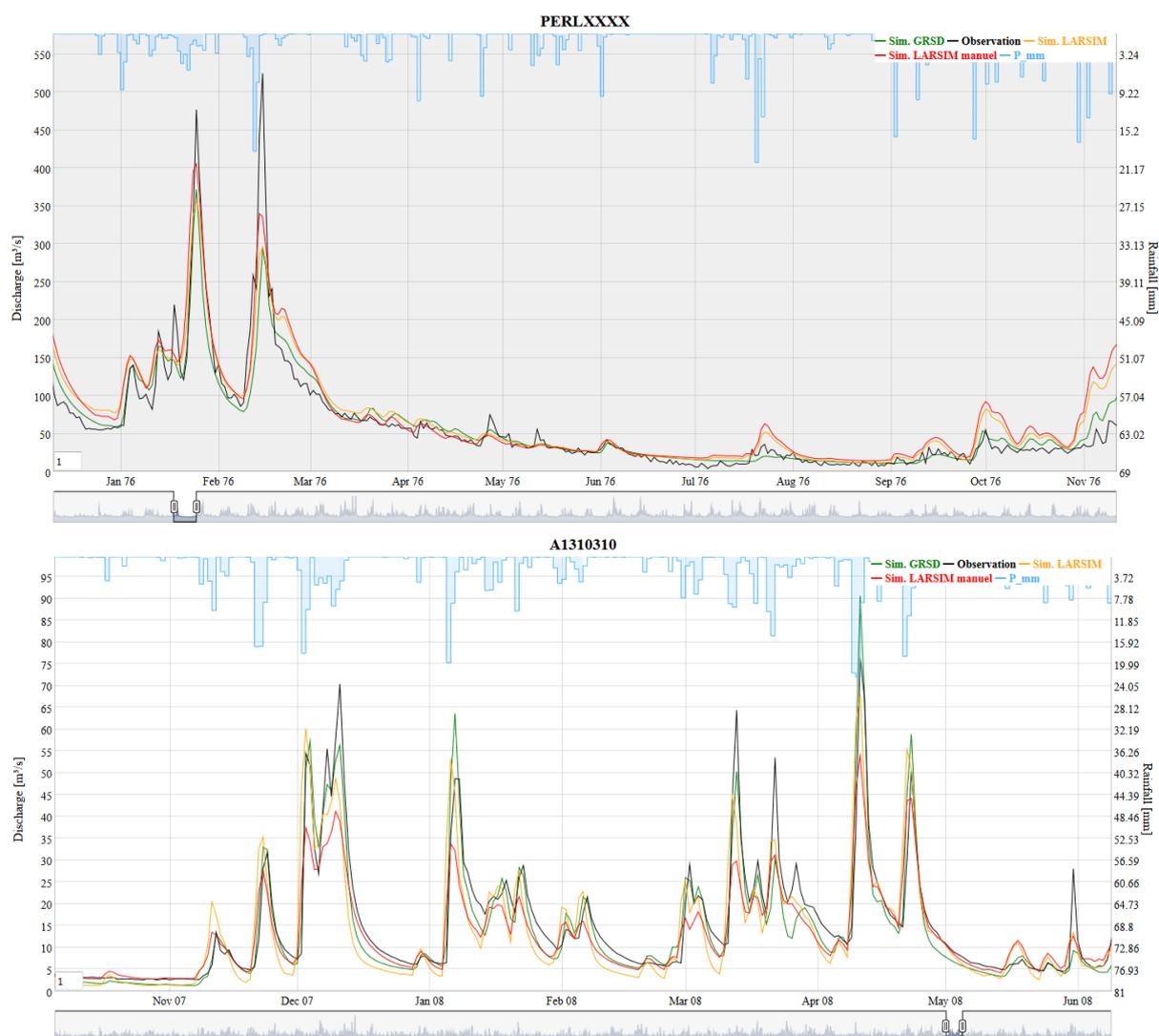


Figure 19 : Hydrogrammes de la Moselle à Perl (en haut) pendant la sécheresse de 1976 et de l'III à Ensisheim (en bas) pendant une période de crues de l'hiver 2007-2008. Ces courbes sont extraites de celles de la Figure 18.

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.

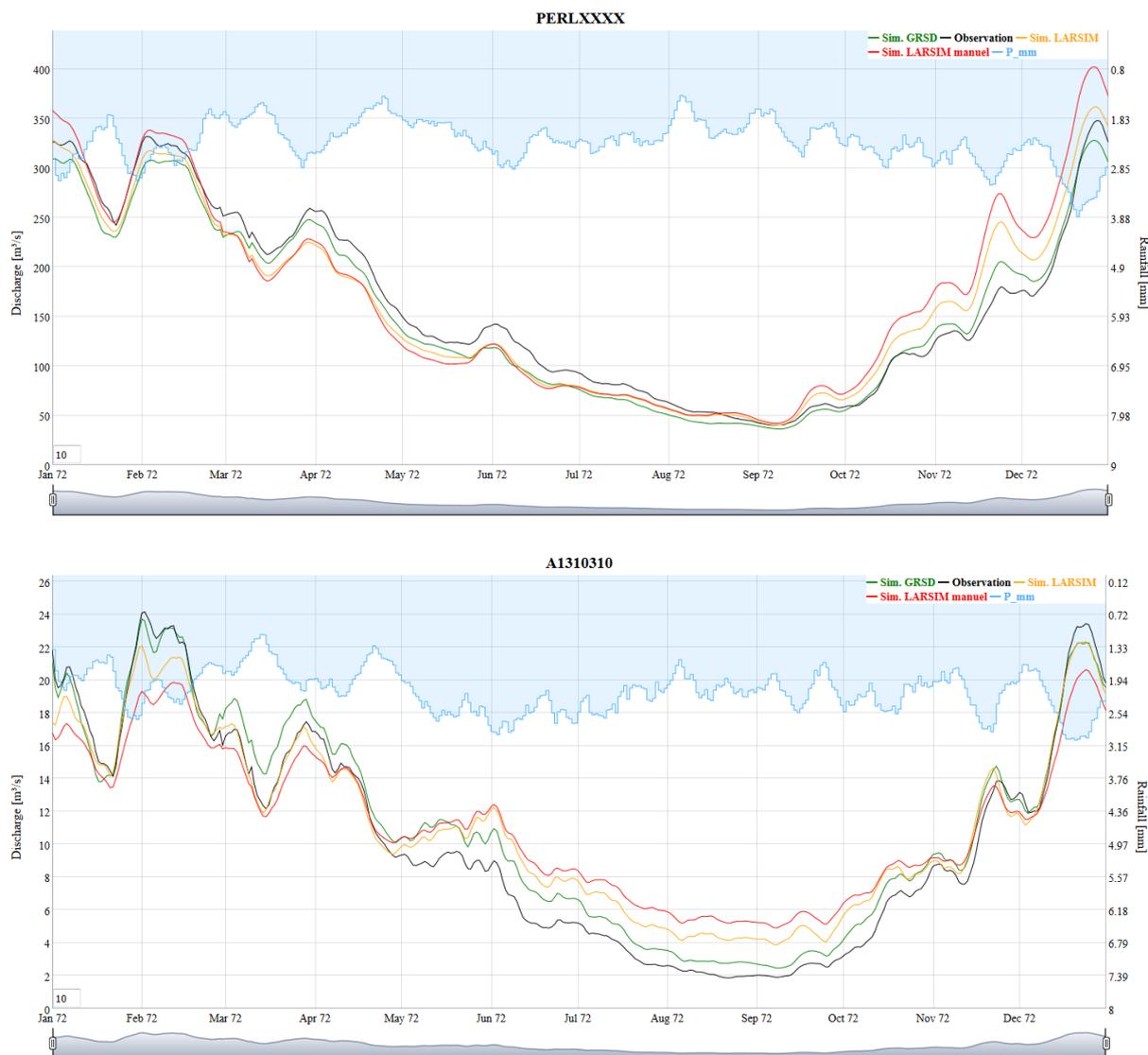


Figure 20 : Courbes de régimes journaliers interannuels pour la Moselle à Perl (en haut) et l'III à Ensisheim (en bas) sur la période 1971-2000.

Les courbes de régime sont fournies sous format html (Annexe H).

3.7. Cartographie des métriques d'erreur

De la même manière que pour les paramètres, des cartes de performances ont été produites (Annexe I). 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.

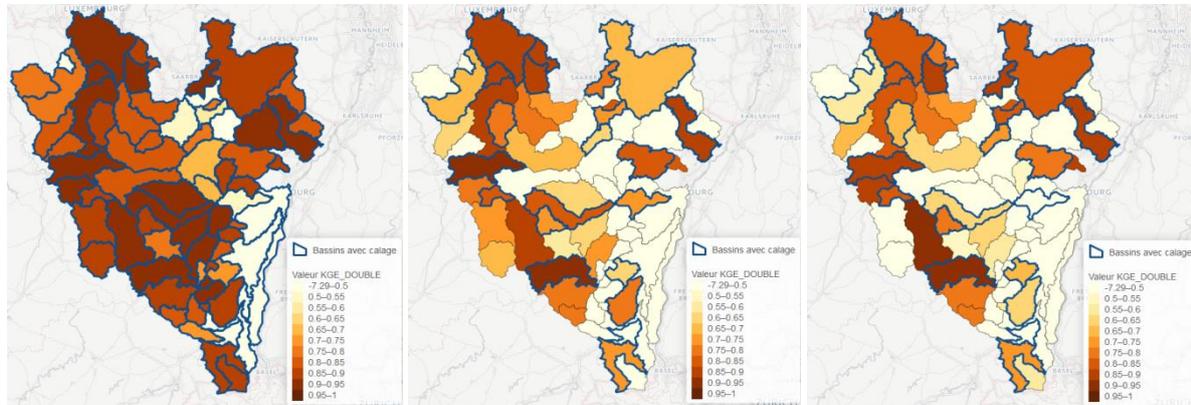


Figure 21 : Valeurs de KGE_DOUBLE (hauts et bas débits) calculées sur la période 1971-2000 pour les modèles calés sur la période 1971-2000 (à gauche GRSD, au centre LARSIM calage automatique, à droite LARSIM calage manuel). Un ton foncé indique les meilleures performances, un ton clair indique de mauvaises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

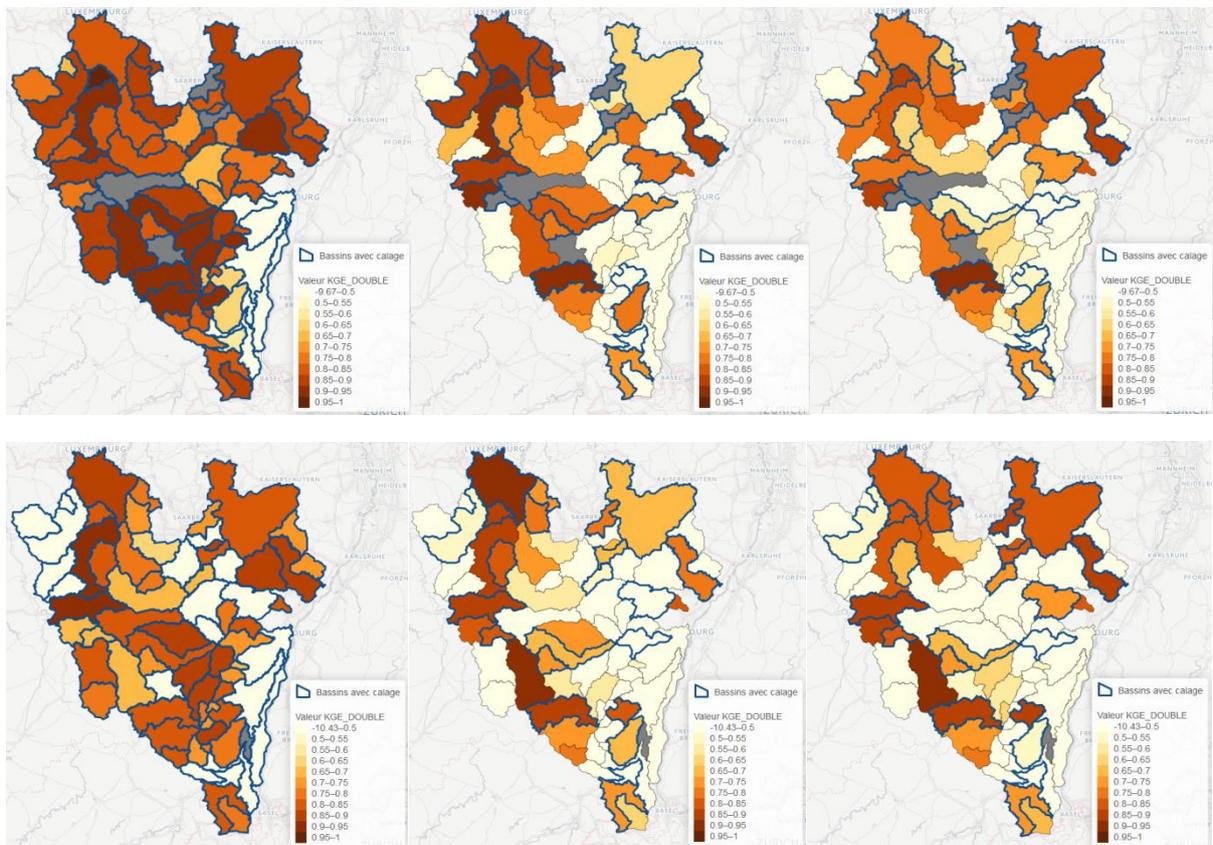


Figure 22 : Valeurs de KGE_DOUBLE calculées sur la période 1971-1985 (période de calage, en haut) et 1986-2000 (période de contrôle, en bas) pour les modèles calés sur la période 1971-1985 (à gauche GRSD, au centre LARSIM calage automatique, à droite LARSIM calage manuel). Un ton foncé indique les meilleures 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).

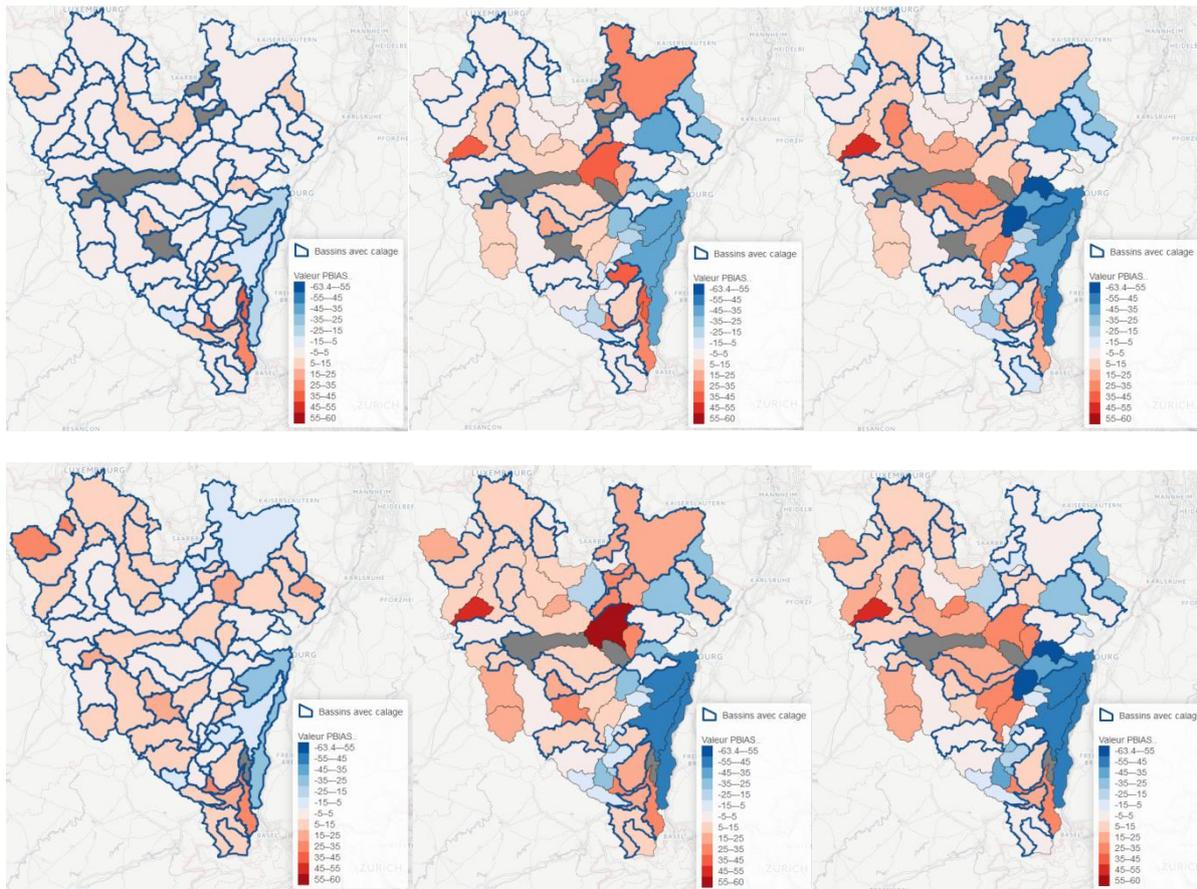


Figure 23 : Idem que Figure 22 mais pour le bilan (biais du débit moyen).

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.

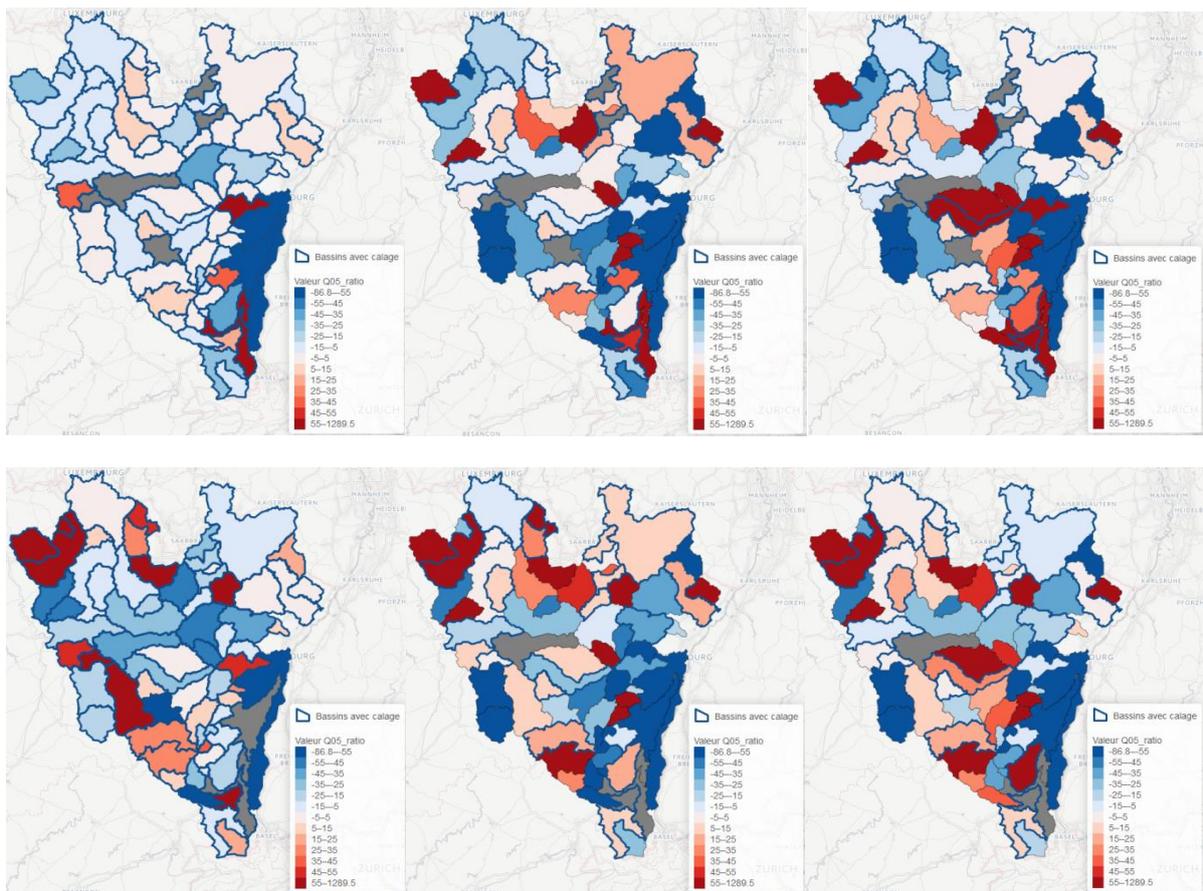
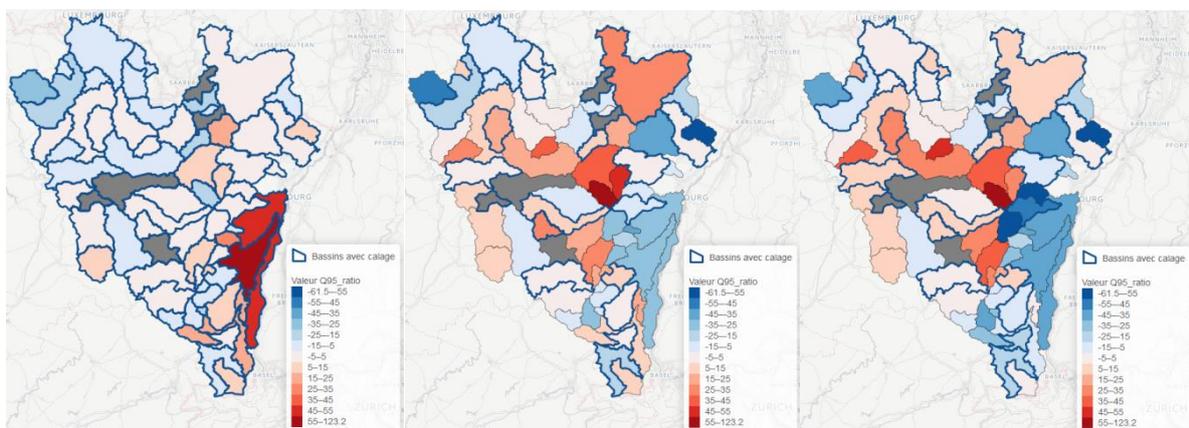


Figure 24 : Idem que Figure 22 mais pour le quantile 5 % (bas débits).

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.



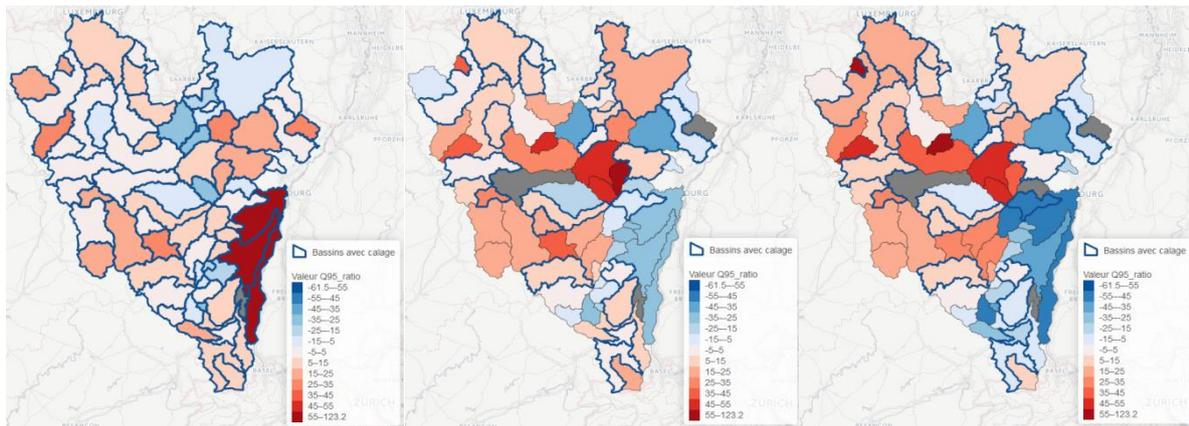


Figure 25 : Idem que Figure 22 mais pour le quantile 95 % (hauts débits).

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.

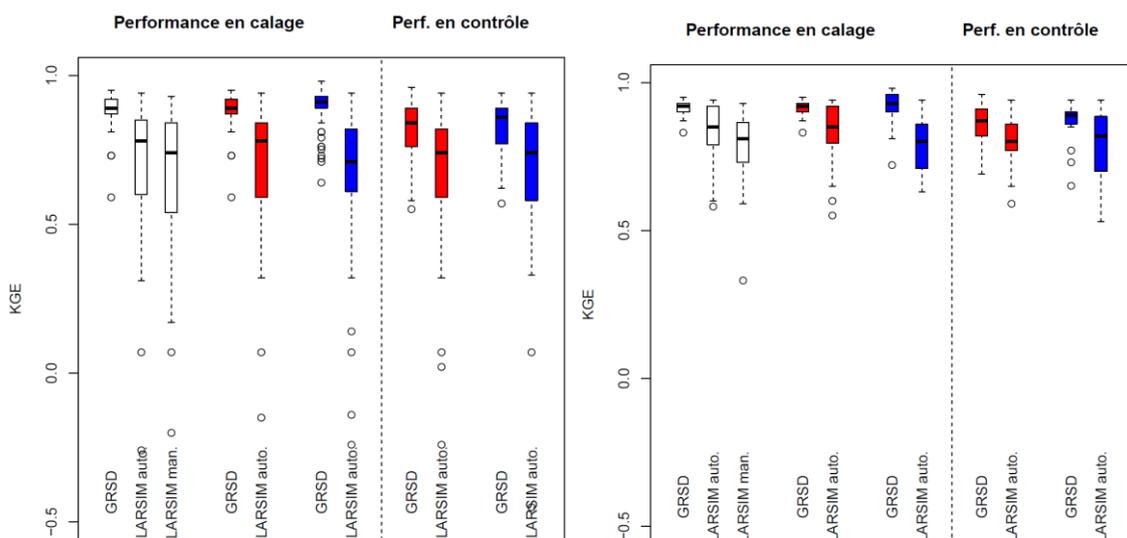


Figure 26 : Distributions des valeurs de KGE pour les différentes configurations des modèles en calage et en contrôle. En blanc les calages sur la période complète, en rouge les calages sur P1, en bleu les calages sur P2. Gauche : les 70 stations.

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.

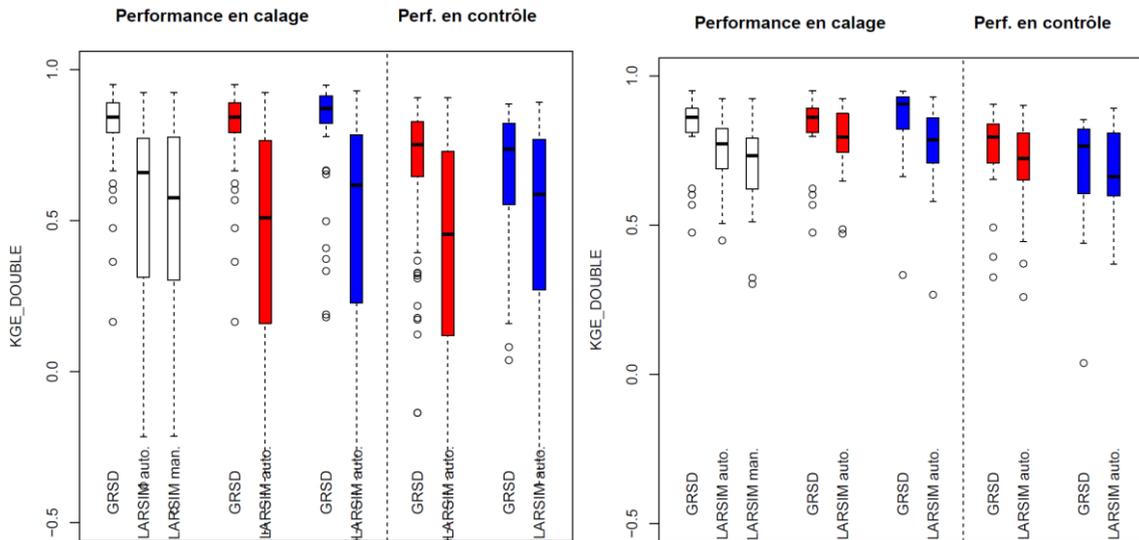


Figure 27 : Idem que Figure 26 pour le KGE_DOUBLE (hauts et bas débits). Gauche : les 70 stations. Droite : uniquement pour les stations calées dans LARSIM.

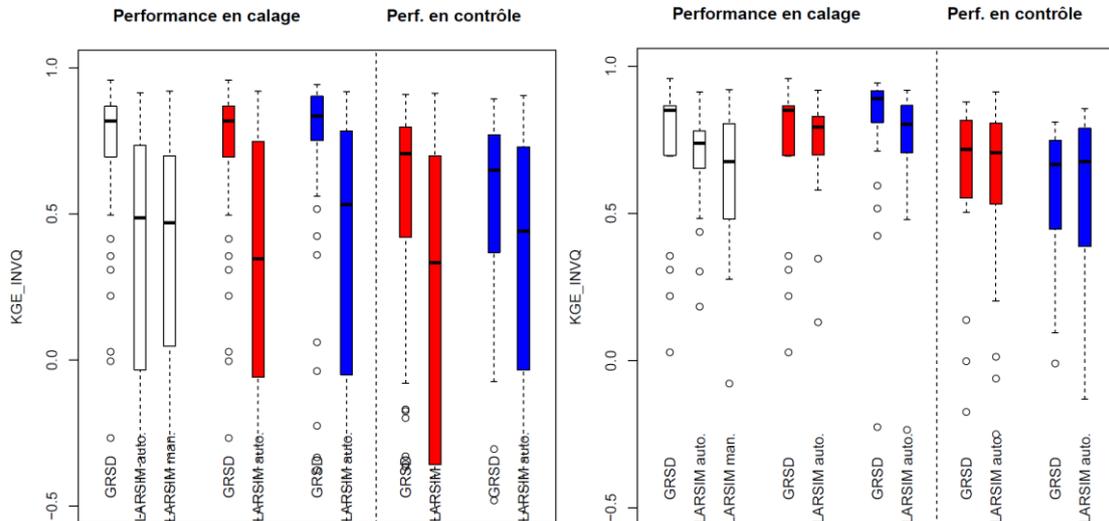


Figure 28 : Idem que Figure 26 pour le KGE_INVQ (bas débits). Gauche : les 70 stations. Droite : uniquement pour les stations calées dans LARSIM.

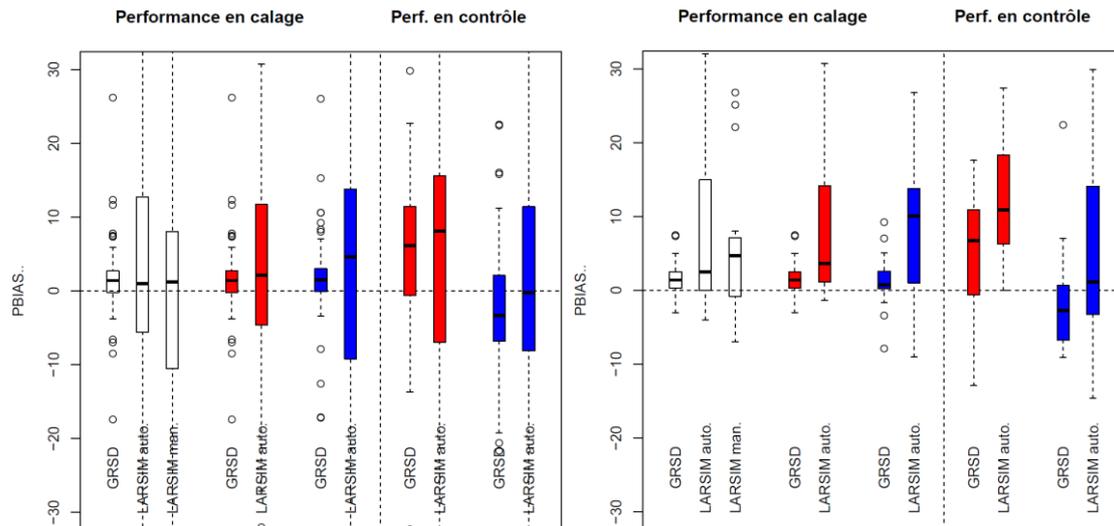


Figure 29 : Idem que Figure 26 pour le biais en pourcentage (des valeurs positives indiquent une surestimation des débits, alors que des valeurs négatives indiquent une sous-estimation des débits). Gauche : les 70 stations. Droite : uniquement pour les stations calées dans LARSIM.

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.

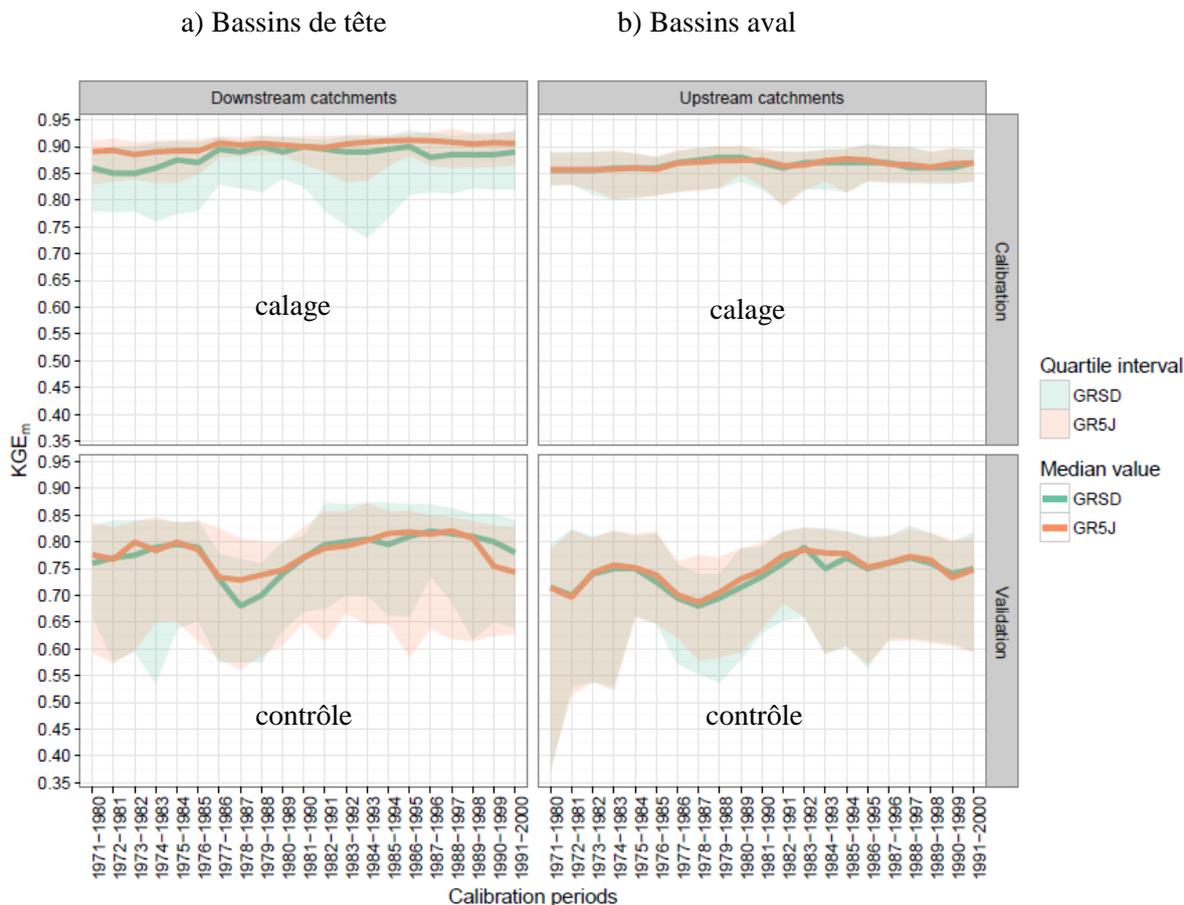


Figure 30 : Evolution de la performance (KGE) des calages de GRSD sur les périodes glissantes. Comparaison avec le modèle global GR5J sur lequel se base GRSD. a) les bassins aval ; b) les bassins de tête. Les bassins de la plaine d'Alsace ont été retirés de cette analyse. D'après de Lavenne et al. (2016).

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.

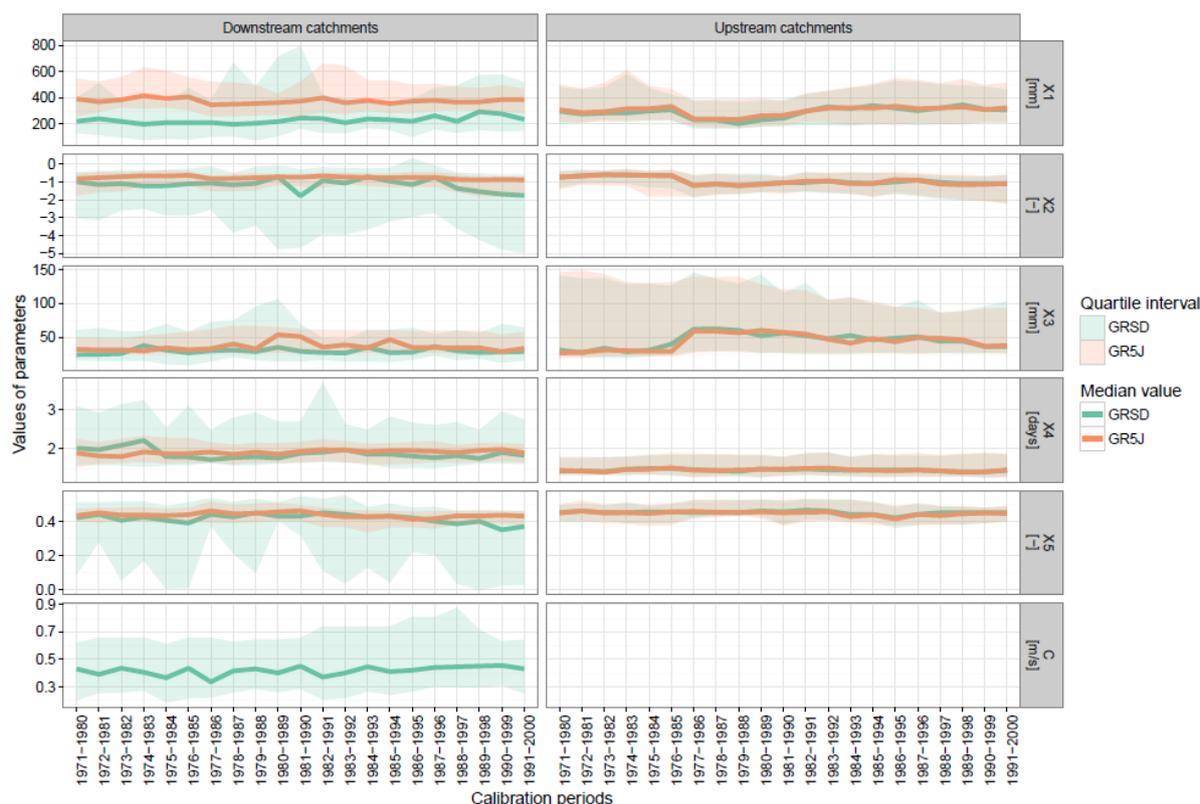


Figure 31 : Idem que Figure 30 (sans période de contrôle) pour les valeurs de paramètres de GRSD. D'après de Lavenne et al. (2016).

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.