



Partenariat 2012 – Thème Hydrologie des étiages - Action 5a



Reconstitution ponctuelle de chroniques hydrologiques incomplètes ou manquantes par modèle pluie-débit

Rapport final

Pierre BRIGODE, Vazken ANDRÉASSIAN, François BOURGIN et Florent LOBLIGEOIS.

Irstea - UR HBAN - Antony

Janvier 2014

- **AUTEURS**

Pierre BRIGODE, ingénieur de recherche contractuel (Irstea), pierre.brigode@irstea.fr

Vazken ANDRÉASSIAN, ICPEF (Irstea), vazken.andreassian@irstea.fr

François BOURGIN, doctorant (Irstea), francois.bourgin@irstea.fr

Florent LOBLIGEOIS, doctorant (Irstea), florent.lobligeois@irstea.fr

- **CORRESPONDANTS**

Onema : Bénédicte AUGEARD, chargée de mission Hydrologie quantitative,
benedicte.augeard@onema.fr

Irstea : Vazken ANDRÉASSIAN, ICPEF (Irstea), vazken.andreassian@irstea.fr

Droits d'usage : Accès libre
Niveau géographique : National
Couverture géographique : France métropolitaine
Niveau de lecture : Professionnels, experts

• RESUME

Les données de débit sont essentielles pour le suivi écologique des milieux aquatiques. Malheureusement, des mesures de débit ne sont disponibles que pour un nombre limité de points sur les cours d'eau français. Loin des stations de jaugeage, on a affaire à des cours d'eau que les hydrologues qualifient de « non jaugés ». Pour faciliter le suivi hydrobiologique des cours d'eau non jaugés, l'ONEMA a lancé, en collaboration avec Irstea, un travail de recherche visant à la reconstitution de chroniques de débit ponctuelles, au droit de sites d'échantillonnage écologique et/ou de suivi de la qualité de l'eau. Sans prétendre se substituer à une mesure directe des débits qui serait bien entendu préférable, ces reconstitutions visent à fournir aux gestionnaires des ordres de grandeur concernant les débits journaliers et leur variabilité au cours des jours (et des années) qui ont précédé la mesure de qualité de l'eau ou du milieu.

Une méthode de simulation de débits en site non jaugé a été développée et mise en œuvre sur 11 225 stations du réseau de référence pérenne (RRP) et de la base de données sur les milieux aquatiques et piscicoles (BDMAP) de l'ONEMA. Cette méthode se base sur la proximité géographique entre sites jaugés et non jaugés et sur la simulation hydrologique à l'aide d'un modèle pluie-débit conceptuel. Un transfert des paramètres du modèle hydrologique est ainsi réalisé entre des bassins versants jaugés (bassins « donneurs », c'est-à-dire possédant des données pour caler un modèle hydrologique) et un bassin non jaugé (bassin « receveur », c'est-à-dire ne possédant pas de données pour caler un modèle hydrologique, mais seulement des données d'entrée climatiques nécessaires à une simulation hydrologique). Les processus nivaux ont été pris en compte grâce à un module neige ajouté au modèle hydrologique. De plus, une méthode d'estimation de l'incertitude liée aux simulations en site non jaugé a été développée et appliquée à chacune des stations considérées, afin de fournir des estimations accompagnées d'un intervalle de confiance à 90%.

Ainsi, en chacune des 11 225 stations non jaugées du réseau RRP et de la BDMAP de l'ONEMA, trois séries journalières de débits ont été produites sur la période 01/08/1958 – 31/07/2012 :

- la série journalière des bornes basses de l'intervalle de confiance à 90%, correspondant au quantile 5%,
- la série journalière de débits issus de la simulation déterministe,
- la série journalière des bornes hautes de l'intervalle de confiance à 90%, correspondant au quantile 95%.

• MOTS CLES (THEMATIQUE ET GEOGRAPHIQUE)

Régionalisation, modèle hydrologique GR4J, module neige CemaNeige, proximité géographique, non jaugé, incertitude, France métropolitaine.

- **RECONSTRUCTION OF DAILY STREAMFLOW SERIES ON FRENCH UNGAUGED CATCHMENTS USING A REGIONALIZED RAINFALL-RUNOFF MODEL**

- **ABSTRACT**

Streamflow series are crucial for the ecological monitoring of aquatic environments. Unfortunately, gauging stations are available on a limited number of points on the French river network. River reaches located far from gauging stations are called “ ungauged ” by hydrologists. In order to ease the hydrobiological monitoring of ungauged river reaches, this research project launched by ONEMA in collaboration with Irstea, aims at reconstructing daily streamflow series in specific river points where sampling and/or water quality monitoring is conducted. Without replacing direct and regular streamflow measurements which would be obviously more reliable, these reconstructions aim at giving quantitative information about the variability of streamflow observed the days (and the year) before water quality measurements.

A method for streamflow simulation on ungauged stations has been developed and applied on 11,225 ONEMA stations. This method is based on spatial proximity between gauged and ungauged stations and on hydrological simulation using a conceptual rainfall-runoff model. Thus, a transfer of rainfall-runoff model parameters is done between gauged catchments (“donor” catchments, i.e. catchments where streamflow data are available for the rainfall-runoff model calibration) and an ungauged catchment (“receiver” catchment, i.e. catchment where streamflow data are not available for the rainfall-runoff model calibration but climatic input series are available for a hydrological simulation). Snow processes have been taken into account thanks to a snow module added on top of the rainfall-runoff model. Moreover, a method for the estimation of uncertainty due to simulation on ungauged stations has been developed and applied on each considered station, in order to produce streamflow series accompanied by a 90% confidence interval.

Thus, on each of the 11,225 considered ungauged stations, three series of daily streamflow have been produced for the 01/08/1958 – 31/07/2012 period:

- the daily series of lower limits of the 90% confidence interval, i.e. the 5% streamflow percentile;
- the daily streamflow series generated by the deterministic simulation;
- the daily series of upper limits of the 90% confidence interval, i.e. the 95% streamflow percentile.

- **KEY WORDS (THEMATIC AND GEOGRAPHICAL AREA)**

Regionalization, GR4J rainfall-runoff model, CemaNeige snow module, spatial proximity, ungauged catchment, uncertainty, continental France.

• SYNTHÈSE POUR L'ACTION OPERATIONNELLE

Les données de débit sont essentielles pour le suivi écologique des milieux aquatiques. Malheureusement, des mesures de débit ne sont disponibles que pour un nombre limité de points sur les cours d'eau français. Loin des stations de jaugeage, on a affaire à des cours d'eau que les hydrologues qualifient de « non jaugés ». Pour faciliter le suivi hydrobiologique des cours d'eau non jaugés, l'ONEMA a lancé, en collaboration avec Irstea, un travail de recherche visant à la reconstitution de chroniques de débit ponctuelles, au droit de sites d'échantillonnage écologique et/ou de suivi de la qualité de l'eau. Sans prétendre se substituer à une mesure directe des débits qui serait bien entendu préférable, ces reconstitutions visent à fournir aux gestionnaires des ordres de grandeur concernant les débits journaliers et leur variabilité au cours des jours (et des années) qui ont précédé la mesure de qualité de l'eau ou du milieu.

La méthode de simulation de débits en site non jaugé proposée par Oudin *et al.* (2008) a été adaptée puis appliquée sur 11 225 stations du réseau de référence pérenne (RRP) et de la base de données sur les milieux aquatiques et piscicoles (BDMAP) de l'ONEMA (cf. Figure A, carte de gauche). Cette méthode se base sur la proximité géographique entre sites jaugés et non jaugés et sur la simulation hydrologique à l'aide d'un modèle pluie-débit conceptuel, le modèle GR4J (Perrin *et al.*, 2003). Un transfert des paramètres du modèle hydrologique GR4J est ainsi réalisé entre des bassins versants jaugés (bassins « donneurs », c'est-à-dire possédant des données pour caler un modèle hydrologique, cf. Figure A, carte de droite) et un bassin non jaugé (bassin « receveur », c'est-à-dire ne possédant pas de données pour caler un modèle hydrologique, mais seulement des données d'entrée climatiques nécessaires à une simulation hydrologique). Les processus nivaux ont été pris en compte grâce au module neige CemaNeige (Valéry, 2010) ajouté au modèle hydrologique GR4J.

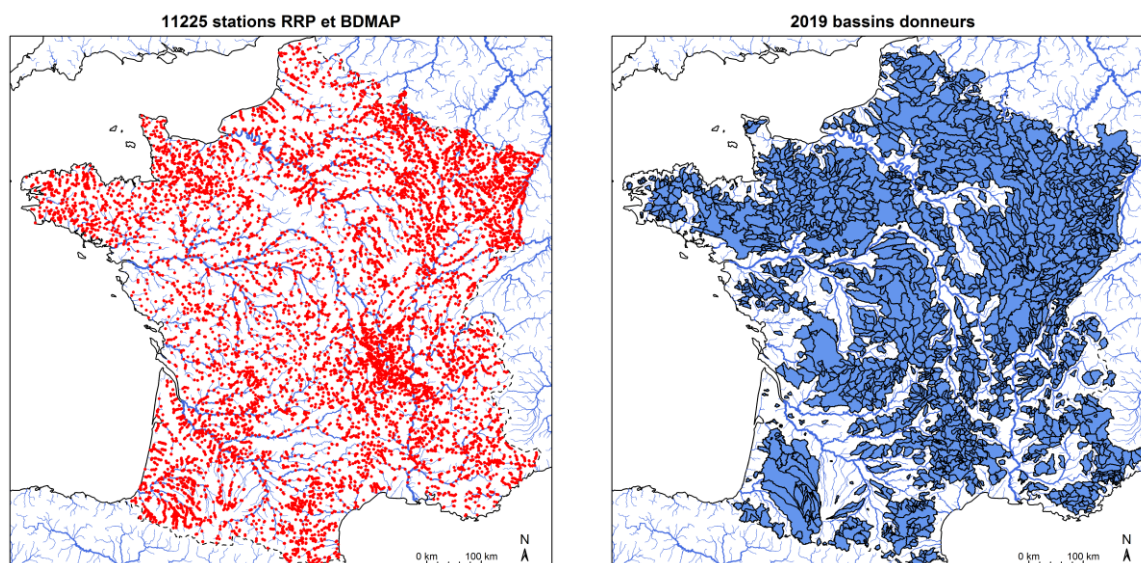


Figure A : Localisation des 11125 sites RRP et BDMAP (à gauche) et base de données des 2019 bassins versants jaugés utilisés comme donneurs dans le transfert de paramètres (à droite).

De plus, une méthode d'estimation de l'incertitude liée aux simulations en site non jaugé a été développée et appliquée à chacune des stations considérées, afin de fournir des estimations accompagnées d'un intervalle de confiance à 90% (Bourgin 2014).

Ainsi, en chacune des 11 225 stations non jaugées des réseaux RRP et BDMAP de l'ONEMA, trois séries de débits journalières ont été produites sur la période 01/08/1958 – 31/07/2012 :

- la série journalière des bornes basses de l'intervalle de confiance à 90%, correspondant au quantile 5%,
- la série journalière de débits issus de la simulation déterministe,
- la série journalière des bornes hautes de l'intervalle de confiance à 90%, correspondant au quantile 95%.

Ces trois séries sont illustrées sur la Figure B pour la station non jaugée numérotée 5240042 (réseau BDMAP, située sur un affluent de l'Isle, elle-même affluent de la Dordogne), sur la période 01/08/1958 – 31/07/1961.

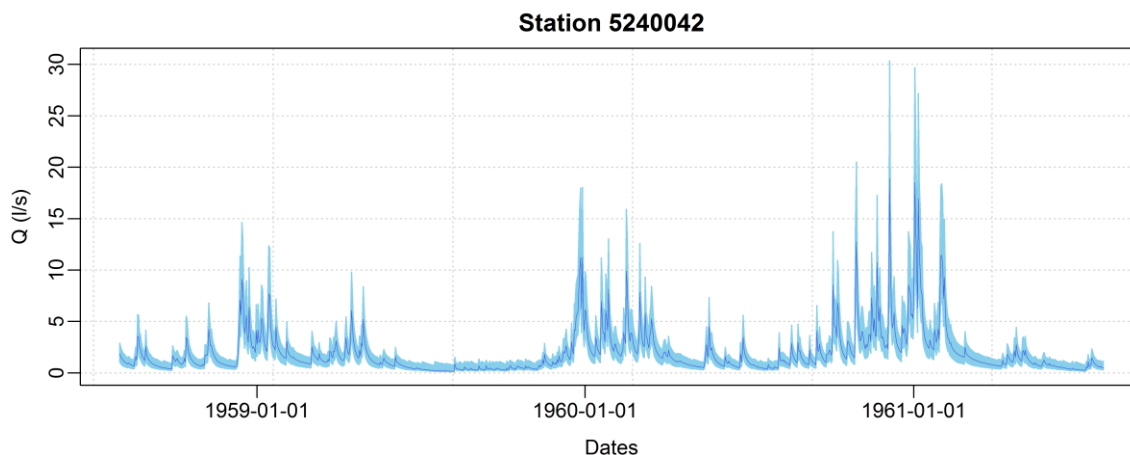


Figure B : Exemple de séries de débits simulées en un point non jaugé (station 5240042, réseau BDMAP) : le débit simulé est en bleu foncé (simulation déterministe) et l'intervalle de confiance à 90% est en bleu clair (simulation probabiliste).

L'ensemble des outils développés permet donc, en tout point du réseau hydrographique français, de simuler une série de débits journaliers sur plusieurs dizaines d'années, à partir (i) de forçages climatiques (données d'entrées du modèle hydrologique, c'est-à-dire séries journalières de pluies et d'évapotranspiration potentielles) et (ii) d'une bibliothèque de paramètres du modèle hydrologique (établie au préalable à partir de calages sur de nombreuses stations jaugées).

Pour en savoir plus :

- Bourgin, F. (2014, Soutenance prévue en avril) Comment quantifier l'incertitude prédictive en modélisation hydrologique ? Travail exploratoire sur un grand échantillon de bassins versants. AgroParisTech (Paris), Cemagref (Antony).
- Oudin, L., Andréassian, V., Perrin, C., Michel, C. & Moine, N. Le. (2008) Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments. *Water Resources Research* 44(3). doi:[10.1029/2007WR006240](https://doi.org/10.1029/2007WR006240).
- Perrin, C., Michel, C. & Andréassian, V. (2003) Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology* 279(1-4), 275–289. doi:[10.1016/S0022-1694\(03\)00225-7](https://doi.org/10.1016/S0022-1694(03)00225-7).
- Valéry, A. (2010) Modélisation précipitations-débit sous influence nivale. Élaboration d'un module neige et évaluation sur 380 bassins versants. AgroParisTech (Paris), Cemagref (Antony).
- **Contact:** Vazken ANDRÉASSIAN, Directeur Adjoint Scientifique, UR HBAN, Irstea (Antony) ; vazken.andreassian@irstea.fr; <http://webgr.irstea.fr/>.

• **SOMMAIRE**

1. Introduction	8
2. Simulation déterministe des débits en site non jaugé (travaux 2012 et 2013)	9
2.1. Méthode utilisée.....	9
2.2. Base de données des bassins versants jaugés.....	9
2.3. Transferts des jeux de paramètres de GR4J entre « donneurs » et « receveurs »	10
2.3.1. Distances entre bassins donneurs et bassins receveurs	10
2.3.2. Transformation du paramètre X4 de GR4J	11
2.3.3. Régionalisation des paramètres du module neige de GR4J (CemaNeige)	11
2.4. Simulation à partir des jeux de paramètre transférés	13
2.5. Performances de GR4J et de la méthode de reconstitution	14
3. Quantification de l'incertitude de reconstitution des débits	16
3.1. Méthode utilisée.....	16
3.2. Evaluation de la fiabilité de la méthode de calcul de l'intervalle prédictif	18
4. Présentation des simulations réalisées pour un ensemble de stations des réseaux RRP et BDMAP.....	20
5. Synthèse	21
6. Bibliographie	22
7. Table des illustrations	23
8. Annexe 1 : Distribution de paramètres de GR4J et de CemaNeige, obtenues après calage sur 2019 et 432 bassins versants jaugés, respectivement.	24

1. Introduction

Les données de débit sont essentielles pour le suivi écologique des milieux aquatiques. Malheureusement, des mesures de débit ne sont disponibles que pour un nombre très limité de points sur les cours d'eau français. Loin des stations de jaugeage, on se trouve sur ce que les hydrologues qualifient de « cours d'eau non jaugés ».

Pour faciliter le suivi des cours d'eau non jaugés, l'ONEMA a lancé, en collaboration avec Irstea, un travail de recherche visant à la reconstitution de chroniques de débit ponctuelles, au droit de sites d'échantillonnage écologique et/ou de suivi de la qualité de l'eau. Sans prétendre se substituer à une mesure directe des débits qui serait bien entendu préférable, nos reconstitutions visent à fournir aux gestionnaires des ordres de grandeur concernant les débits journaliers et leur variabilité au cours des jours (et des années) qui ont précédé la mesure de qualité de l'eau ou du milieu.

L'année 2012 a été consacrée à la préparation d'un prototype et aux travaux méthodologiques. L'année 2013 est venue compléter ces travaux en fournissant un calcul d'incertitude sur les débits reconstitués.

Ce rapport final présente :

- la méthode retenue pour la prévision déterministe des débits ;
- la méthode retenue pour l'évaluation de l'incertitude de reconstitution ;
- les simulations concernant l'ensemble des 11225 stations des réseaux RRP et BDMAP de l'ONEMA.

2. Simulation déterministe des débits en site non jaugé (travaux 2012 et 2013)

2.1. Méthode utilisée

L'année 2012 a été consacrée à la production de débits en site non jaugé. Nous avons utilisé le modèle pluie-débit GR4J (Perrin *et al.*, 2003) et avons appliqué la méthode développée par Oudin *et al.* (2008), qui s'appuie sur un transfert de jeux de paramètres entre des bassins jaugés (« donneurs ») et un bassin non jaugé (« receveur »). Pour le receveur, le transfert des paramètres du modèle GR4J est rendu nécessaire du fait de l'absence de données de débit nécessaires au calage. Cependant, les données d'entrée climatiques (pluie et évapotranspiration potentielle (ETP) au pas de temps journalier) propres au receveur sont disponibles.

Une partie des travaux de l'année 2013 a été consacrée à la prise en compte de la neige pour la simulation en site non jaugé. Les paramètres de CemaNeige (module neige à deux paramètres développé par Valéry (2010), utilisable avec le modèle GR4J) ont ainsi été obtenus par calage sur un sous-ensemble de bassins versants jaugés considérés comme influencés par la neige (part de précipitation solide moyenne annuelle supérieure à 10%). Ces jeux de paramètres ont ensuite été régionalisés par massif montagneux et utilisés pour réaliser des simulations sur les stations du réseau de référence pérenne (RRP) et de la base de données sur les milieux aquatiques et piscicoles (BDMAP) de l'ONEMA, considérées comme influencées par la neige.

2.2. Base de données des bassins versants jaugés

La Figure 1 présente à la fois les 11225 stations des réseaux RRP et BDMAP de l'ONEMA (bassins receveurs, à gauche) et les bassins versants sélectionnés comme « donneurs » de paramètres (à droite). Il s'agit de 2019 bassins versants considérés comme non influencés ou faiblement influencés, et pour lesquels des données au pas de temps journalier sont disponibles sur une durée d'au moins 6 années hydrologiques avec peu de lacunes (moins de 10%) sur la Banque Hydro (www.hydro.eaufrance.fr/).

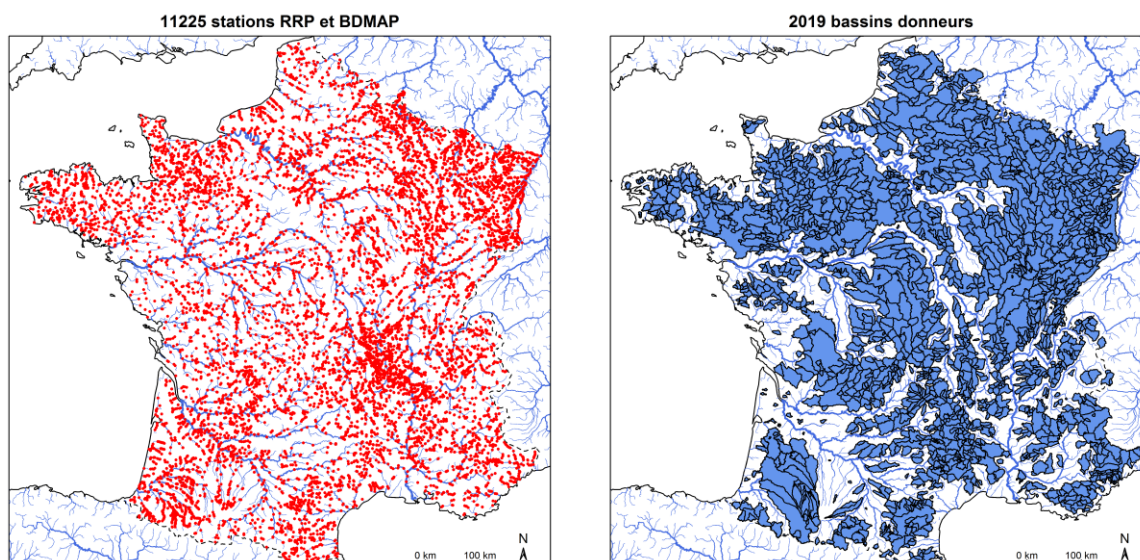


Figure 1 : Localisation des 11225 sites RRP et BDMAP (à gauche) et base de données des 2019 bassins versants jaugés utilisés comme donneurs dans le transfert de paramètres (à droite).

Le modèle pluie-débit GR4J a été calé sur chacun de ces bassins versants, et le vecteur de paramètres correspondant est donc disponible pour un transfert vers les bassins versants non jaugés. La fonction objectif utilisée pour le calage de GR4J est le critère de Nash et Sutcliffe (1970) calculé sur les racines carrées des débits. Les distributions des paramètres obtenues après calage sont présentées dans l'Annexe 1.

2.3. Transferts des jeux de paramètres de GR4J entre « donneurs » et « receveurs »

2.3.1. Distances entre bassins donneurs et bassins receveurs

La Figure 2 illustre la méthode retenue pour transférer les paramètres de GR4J entre bassins jaugés et non jaugés. Pour le bassin cible (non jaugé) figuré en gris, on commence tout d'abord à rechercher les cinq bassins versants jaugés disponibles dans notre base les plus proches géographiquement.

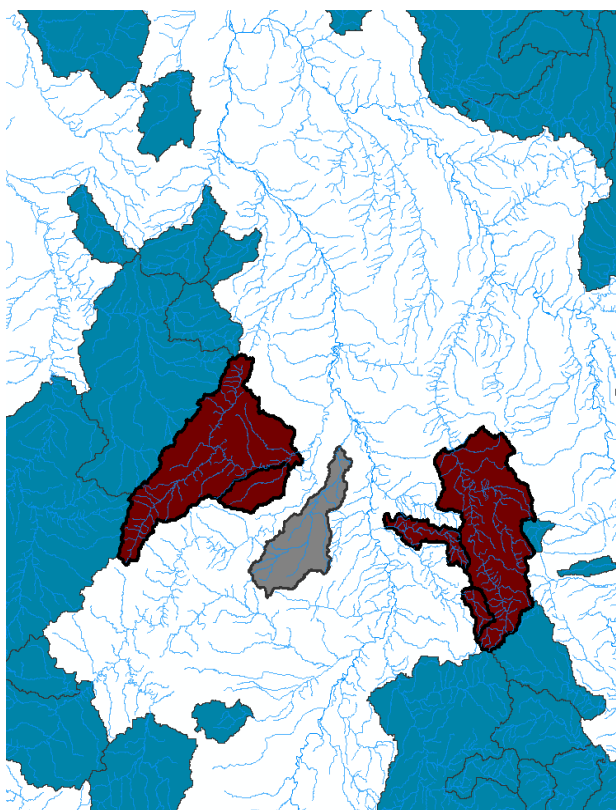


Figure 2 : « Donneurs » et « receveurs » : le bassin versant receveur (considéré comme non jaugé) est figuré en gris, les cinq bassins versants voisins jaugés donneurs sont en rouge, en bleu on a représenté les autres bassins versants jaugés

Pour éviter le transfert entre bassins versants de taille très différentes, on utilise une distance pondérée entre les centres d'inertie des bassins versants et les exutoires de ces derniers (ce qui évite le transfert de paramètres entre deux bassins concentriques de tailles très différentes) :

$$distance_{bassins} = 0.2.distance_{exutoires} + 0.8.distance_{centroïdes} \quad \text{Eq. 1}$$

La Figure 3 donne une indication des distances moyennes entre bassins donneurs et bassins receveurs. Les distances les plus importantes sont calculées pour les bassins receveurs ayant les plus grandes surfaces, et se trouvant donc sur les grands fleuves français (Seine, Loire et Rhône), ainsi que dans les zones où la densité de bassins donneurs est faible.

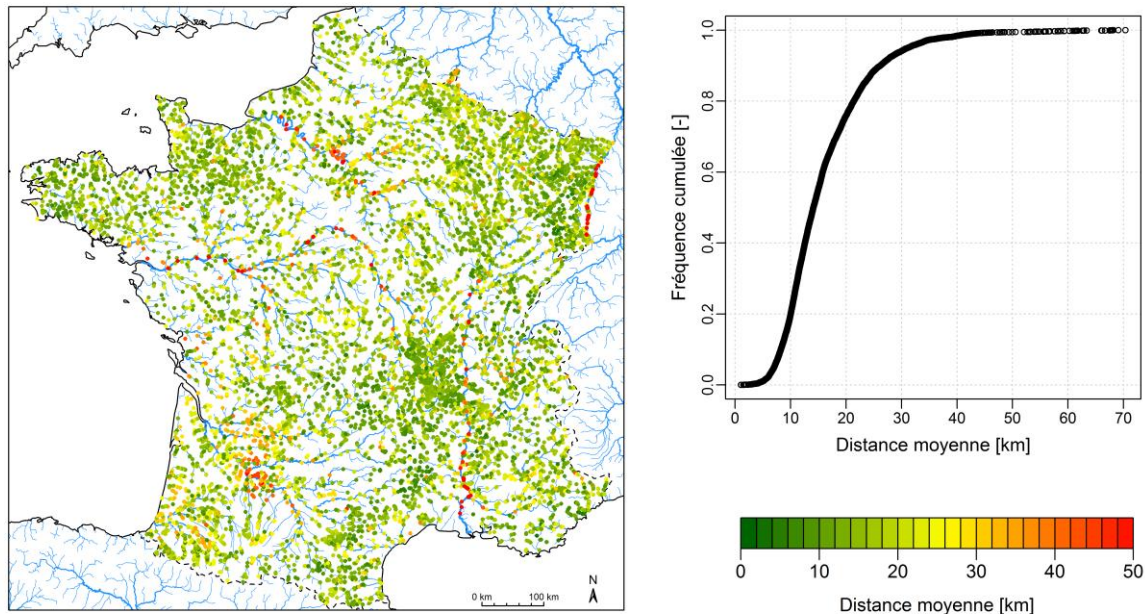


Figure 3 : Distances moyennes estimées pour chaque bassin receveur entre le bassin receveur et les cinq bassins donneurs les plus proches.

2.3.2. Transformation du paramètre X4 de GR4J

Nous utilisons ici une transformation d'un des paramètres de GR4J, le paramètre X4 (temps de base de l'hydrogramme unitaire, exprimé en jours), du fait de la dépendance significative entre la valeur de ce paramètre et la surface du bassin versant considéré. La valeur du paramètre X4 est donc modifiée en fonction du rapport entre la surface du bassin receveur (S_{receveur}) et la surface du bassin donneur (S_{donneur}), selon la formule suivante :

$$X4_{\text{receveur}} = X4_{\text{donneur}} * \frac{S_{\text{receveur}}^{0.3}}{S_{\text{donneur}}} \quad \text{Eq. 2}$$

Cette relation a été déterminée sur plus de 700 bassins versants français, au pas de temps horaire (Lobligeois, 2014).

2.3.3. Régionalisation des paramètres du module neige de GR4J (CemaNeige)

Les processus nivaux ont été pris en compte pour la simulation en site non jaugé grâce à l'utilisation du module neige appelé CemaNeige (Valéry 2010), utilisable avec le modèle GR4J. Pour cela, les bassins versants donneurs ont tout d'abord été séparés en deux groupes, en fonction de l'influence de la neige sur leur régime hydrologique. Cette influence a été estimée en calculant, pour chaque bassin versant, la part de précipitation solide sur l'ensemble des précipitations reçues en moyenne par le bassin :

- les bassins versants ayant une part de précipitations solides inférieure à 10% sont jugés pas ou peu influencés par la neige.
- les bassins versants ayant une part de précipitations solides supérieure à 10% sont jugés influencés par la neige.

Pour les bassins jaugés jugés pas ou peu influencés par la neige, CemaNeige a été utilisé pour la simulation de débits avec les valeurs de paramètres par défaut, établies par Valéry (2010) sur 380 bassins versants ($C_{Tg} = 0.25$ [-] et $K_f = 3.74$ mm/j).

Pour les bassins jaugés jugés influencés par la neige, les paramètres de CemaNeige ont été calés. La Figure 4 indique la localisation et le massif montagneux des 432 bassins versants donneurs influencés par la neige (à gauche) et les distributions des valeurs de paramètres de CemaNeige par

massif montagneux (à droite). Une régionalisation des paramètres de CemaNeige par massifs montagneux a ensuite été réalisée, au vu des différences constatées des distributions de paramètres entre massifs : par exemple, les valeurs du paramètre X1 de CemaNeige sont en médiane plus faibles pour les bassins de moyennes montagnes (Alsace, Jura et Massif Central) par rapport aux bassins de plus hautes montagnes (Alpes et Pyrénées). Des valeurs de paramètres médianes ont été déterminées par massif montagneux, et sont détaillées dans le Tableau 1. Ce sont ces valeurs médianes qui ont ensuite été utilisées pour les 432 bassins jaugés jugés influencés par la neige. Les distributions obtenues après calage des deux paramètres sont également présentées dans l'Annexe 1.

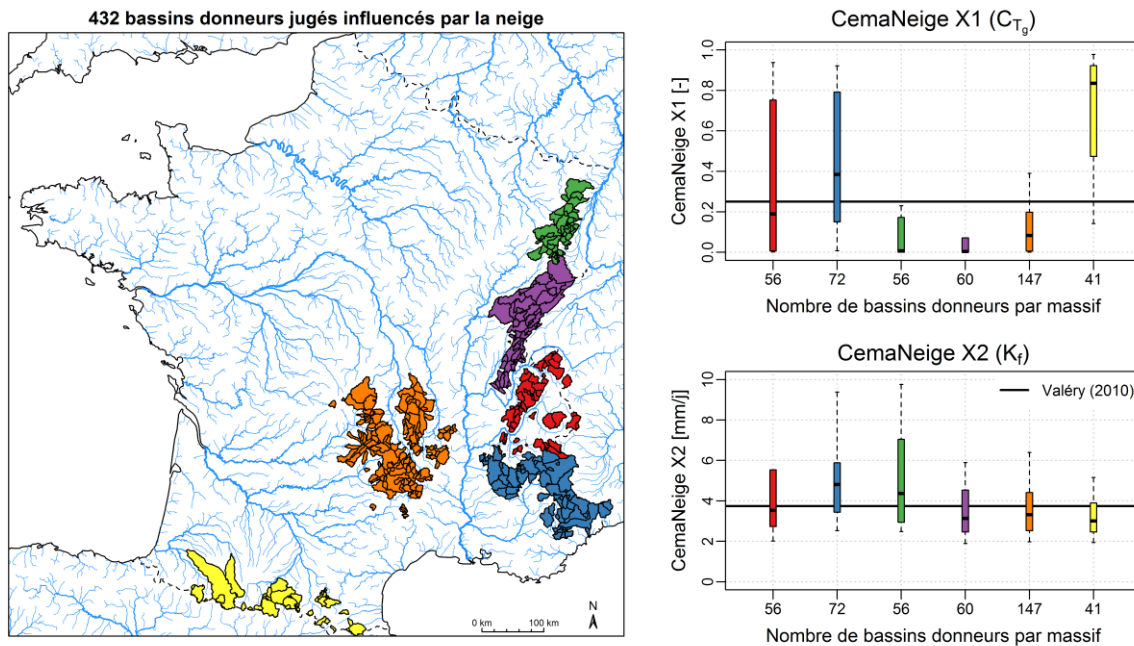


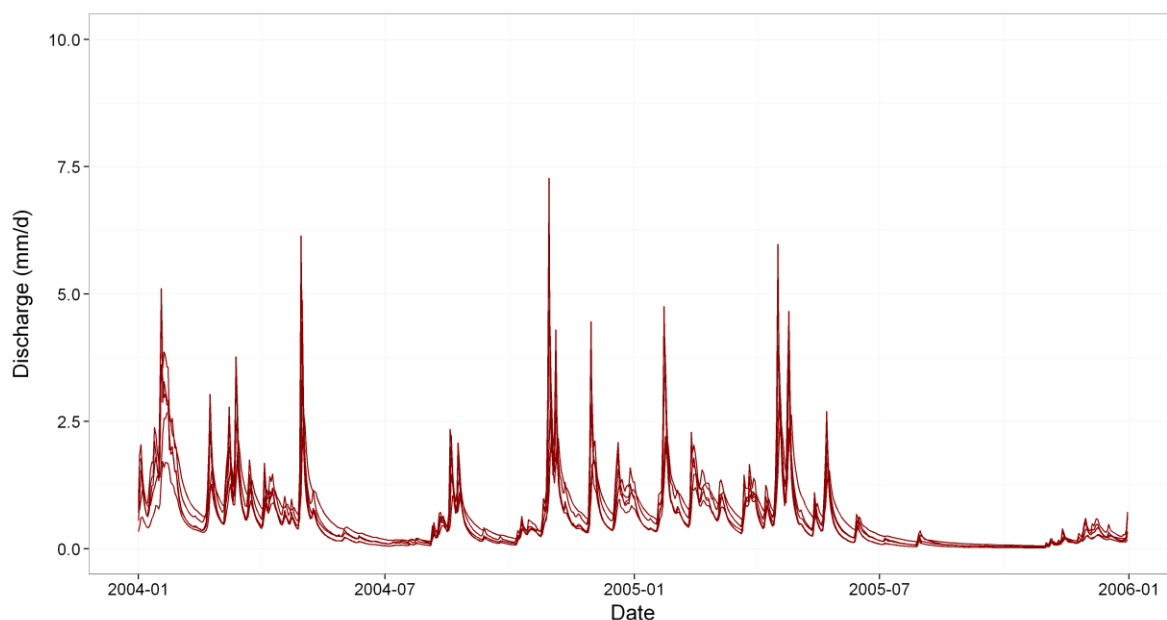
Figure 4 : Localisation et massif montagneux (Alsace, Jura, Massif Central, Alpes du Nord, Alpes du Sud ou Pyrénées) des 432 bassins donneurs considérés comme influencés par la neige (à gauche) et distributions des valeurs de paramètres de CemaNeige par massif montagneux (à droite). Les boîtes à moustaches indiquent les quantiles 0.10, 0.25, 0.50, 0.75 et 0.90 des valeurs de paramètres.

Tableau 1 – Valeurs médianes des paramètres de CemaNeige de référence (Valéry 2010) et établies pour chaque massif montagneux français sur une base de 432 bassins versants jaugés.

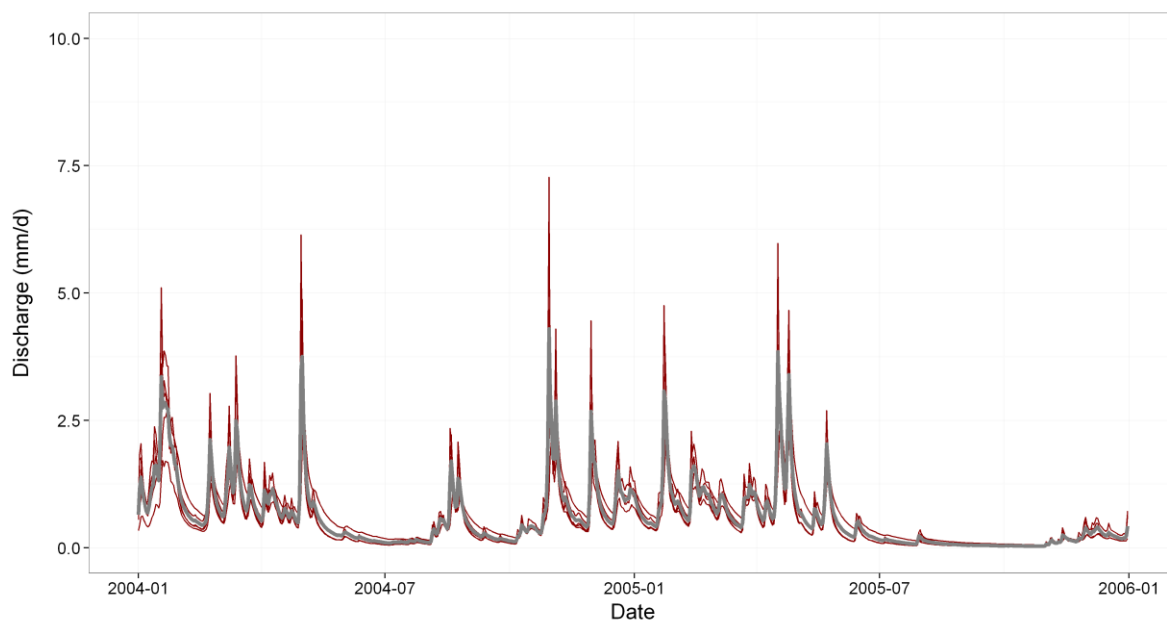
	CemaNeige X1 (C_{T_g} [-])	CemaNeige X2 (K_f [mm/j])
Par défaut (Valéry 2010)	0.250	3.74
Alsace	0.006	4.37
Jura	0.003	3.14
Massif Central	0.830	3.32
Alpes du Nord	0.189	3.53
Alpes du Sud	0.384	4.81
Pyrénées	0.835	3.00

2.4. Simulation à partir des jeux de paramètre transférés

A partir de chacun des cinq jeux de paramètres transférés, et en utilisant en entrée de GR4J les données de pluie et d'ETP propres au bassin versant jaugé, on simule cinq hydrogrammes (Figure 5-a), que l'on moyenne ensuite (Figure 5-b) : c'est cette moyenne d'ensemble qui est utilisée pour le bassin versant non jaugé.



a.



b.

Figure 5 : Simulation d'ensemble obtenue à partir du transfert des paramètres des cinq bassins les plus proches (a- cinq simulations distinctes issues des mêmes données d'entrée et de cinq jeux de paramètres différents, b- avec la moyenne des simulations en gris).

2.5. Performances de GR4J et de la méthode de reconstitution

La Figure 6 synthétise les performances obtenues par le modèle GR4J en calage sur les 2019 bassins versants donneurs. Le critère de Nash et Sutcliffe (1970) estimé sur les racines carrées des débits a été utilisé pour cette évaluation. Afin de faciliter la comparaison des performances entre bassins versants, ce critère a été borné (valeurs entre -1 et 1) grâce à la transformation C2M proposée par Mathevet *et al.* (2006). Les performances générales de GR4J sont bonnes, avec des critères C2M moyens supérieurs à 0.71 (équivalent à un critère de Nash-Sutcliffe de 0.82) pour plus de 50% des bassins versants donneurs. Certaines déficiences régionales sont cependant observées : les performances en calage de GR4J semblent plus faibles pour plusieurs bassins versants de l'Eure et de l'Eure et Loire, ainsi que dans la partie nord de l'Alsace par exemple.

Une évaluation des performances de la méthode de reconstitution a ensuite été réalisée sur chacun des 2019 bassins versants donneurs, considéré tour à tour comme non jaugé. Pour chacun de ces bassins, les cinq bassins donneurs les plus proches ont été utilisés pour simuler une série de débits journaliers, ensuite comparée avec la série de débits journaliers observés sur le bassin non jaugé cible. Les performances sont à nouveau évaluées grâce au critère de Nash-Sutcliffe sur les racines carrées des débits, transformé en C2M. La Figure 7 présente la distribution spatiale des performances de la méthode (carte à gauche), ainsi que la distribution des performances en calage des bassins versants donneurs (déjà présentée dans la Figure 6) comparée à la distribution des performances en validation. La carte des performances permet notamment d'identifier des régions où la méthode de reconstitution semble être peu performante, comme la Picardie et la Haute-Normandie. Ces régions, à l'hydrologie fortement influencée par des nappes souterraines régionales, sont connues comme étant difficilement modélisées par GR4J. On peut donc s'attendre à des reconstitutions de débit avec des incertitudes plus fortes sur ces zones. A l'opposé, des performances moyennes élevées sont obtenues pour d'autres régions, telles que la Bretagne, l'Alsace, la Lorraine, la Franche-Comté ou encore le Centre.

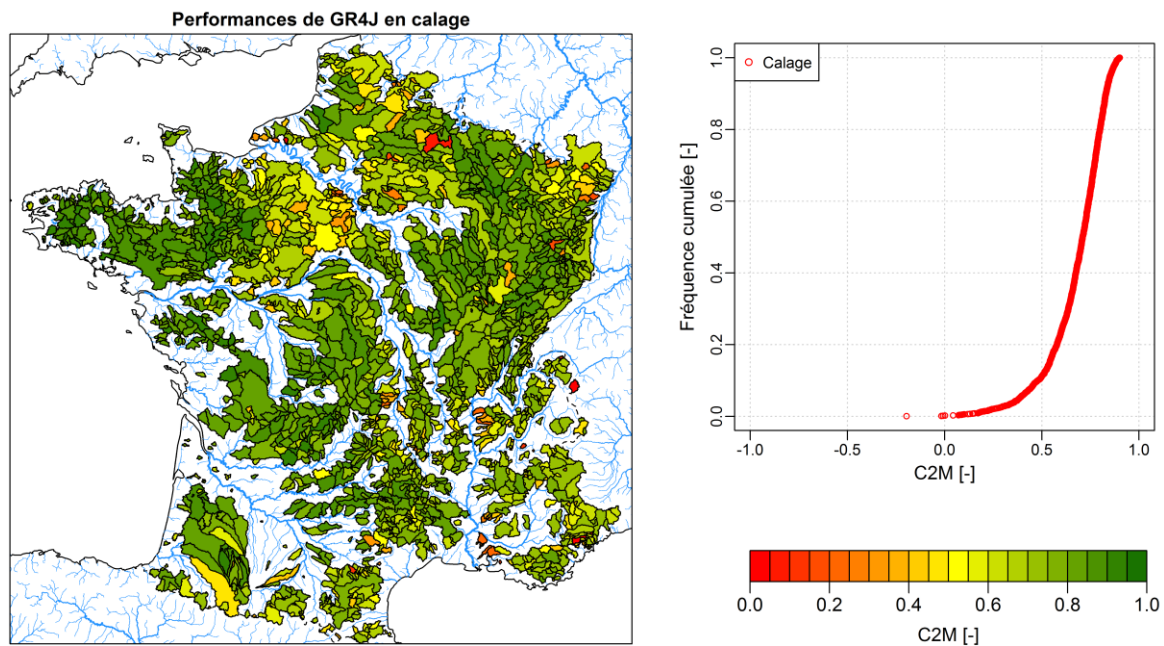


Figure 6 : Performances de GR4J en calage sur les 2019 bassins versants jaugés. Le critère de calage et le critère d'évaluation des performances est le critère de Nash-Sutcliffe, estimé sur les racines carrées des débits, puis borné entre -1 et 1 selon la transformation C2M proposée par Mathevet *et al.* (2006).

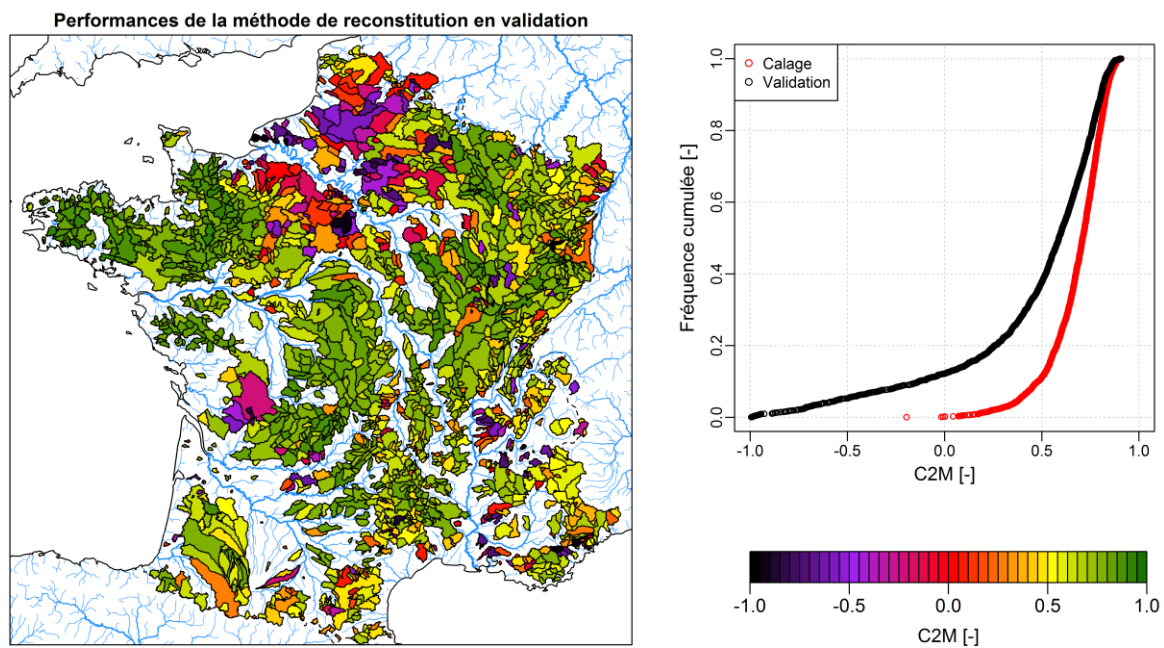


Figure 7 : Performances de la méthode de reconstitution évaluée sur les 2019 bassins versants jaugés. Le critère de calage et le critère d'évaluation des performances est le critère de Nash-Sutcliffe, estimé sur les racines carrées des débits, puis borné entre -1 et 1 selon la transformation C2M proposée par Mathevet *et al.* (2006).

3. Quantification de l'incertitude de reconstitution des débits

3.1. Méthode utilisée

Les séries de débit simulées grâce à la méthode présentée dans la section 2 sont pour l'instant déterministes. En 2013, nous avons aussi travaillé au développement d'une méthode spécifique pour quantifier l'incertitude de modélisation sur les bassins non jaugés. En effet, si de nombreuses approches ont été proposées pour les bassins jaugés, le cas des bassins non jaugés est resté peu étudié. A notre connaissance, seules des méthodes quantifiant l'incertitude à l'aide d'ensembles de jeux de paramètres ont été testées, et elles sous-estiment assez fortement les incertitudes. Nous avons donc développé une nouvelle approche qui s'appuie sur les efforts portés depuis plusieurs années sur la simulation déterministe des bassins non jaugés (Bourgin, 2014).

L'idée principale est la suivante : si nous disposons d'une méthode permettant d'obtenir une simulation déterministe sur un bassin non jaugé « receveur » à partir d'un transfert de jeux de paramètres issus de bassins « donneurs » (Figure 2), il est également possible d'appliquer cette procédure à ces bassins « donneurs ». Comme les débits observés sont en fait disponibles sur ces bassins, nous avons accès aux erreurs empiriques obtenues par la méthode de régionalisation sur chacun de ces bassins « donneurs ». Il s'agit alors de trouver une manière de regrouper et transférer ces erreurs des bassins « donneurs » au bassin « receveur ».

Nous illustrons la démarche pour un bassin versant, où la méthode de transfert de jeux de paramètres par proximité géographique ci-dessus a été appliquée.

- 1) Nous identifions dans un premier temps les 5 bassins versants « donneurs » (en rouge sur la Figure 2) d'un bassin versant « receveur » (en gris). A partir des 5 jeux de paramètres calés sur les bassins « donneurs », nous obtenons une simulation sur le bassin « receveur » en moyennant les 5 chroniques de débits obtenus.
- 2) Nous appliquons cette méthode de régionalisation à chacun des bassins « donneurs » pour obtenir des débits simulés sur chacun de ces bassins comme s'ils étaient non jaugés (Figure 8).

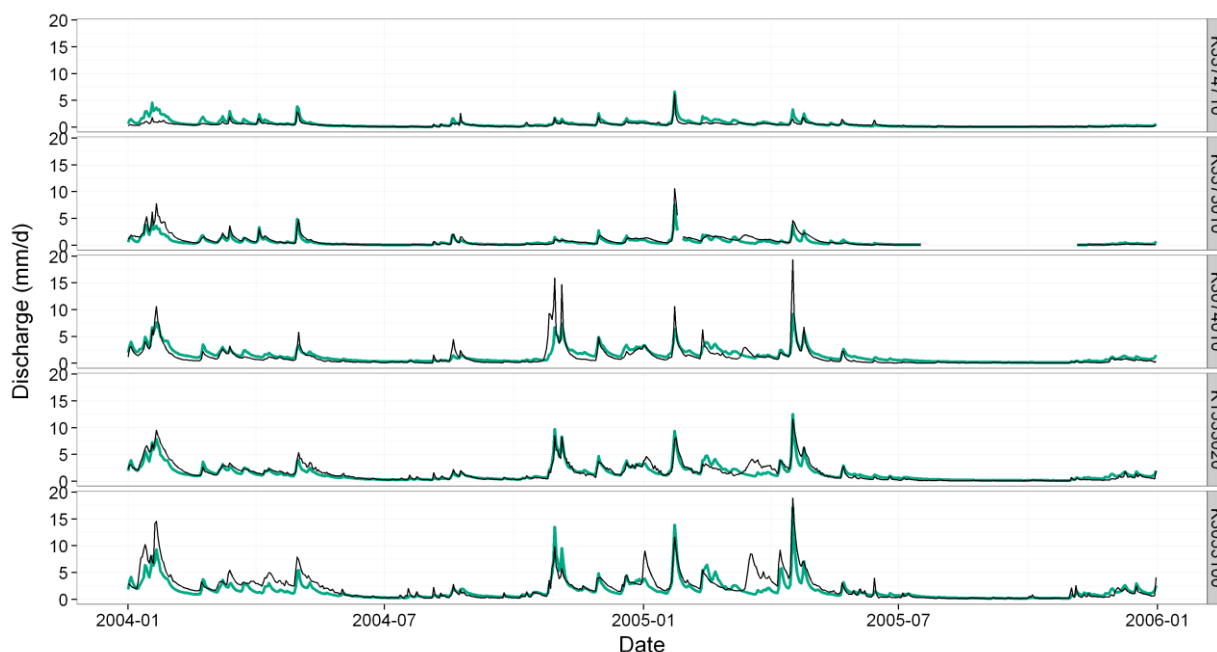


Figure 8 : Débits observés et simulations en mode non jaugé obtenus pour les bassins versants donneurs.

- 3) Les erreurs relatives calculées sur chacun des bassins « donneurs » sont séparées en dix groupes de débits, définis pour chacun des bassins en fonction des débits simulés (Figure 9).

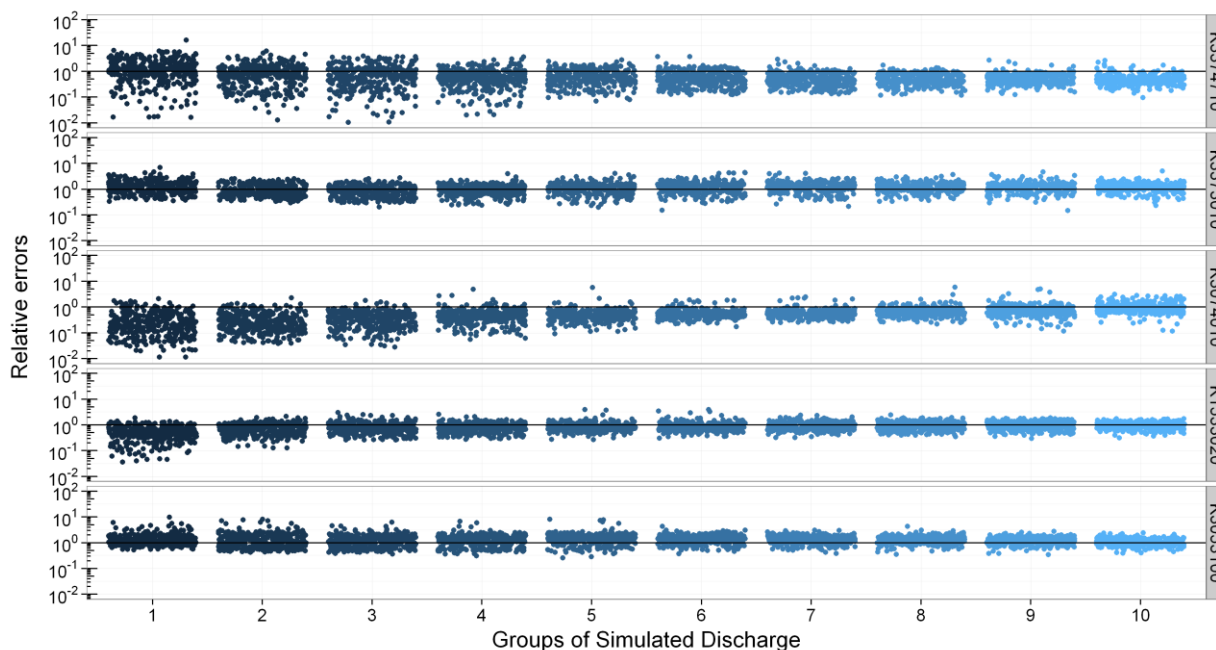


Figure 9 : Erreurs observées sur la reconstitution des débits en mode non jaugé, réparties par classes de débits.

- 4) On regroupe pour chaque groupe de débits, les erreurs relatives des 5 bassins « donneurs », de manière à calculer pour chaque groupe de débits des coefficients multiplicatifs décrivant les marges d'erreurs (Figure 10).

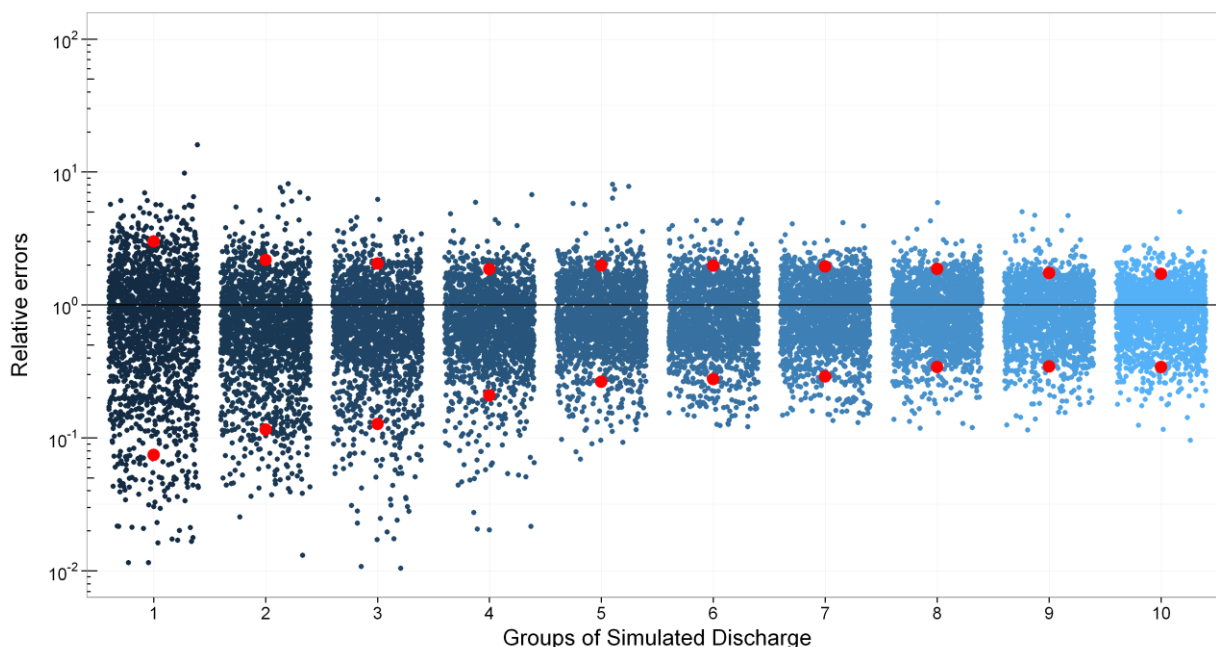


Figure 10 : Marges d'erreur regroupées par classe de débits. Les points rouges figurent les quantiles 0.05 et 0.95 des erreurs empiriques, qui sont utilisés pour définir l'intervalle de confiance à 90%.

- 5) Enfin, pour le bassin « receveur », on applique les coefficients multiplicatifs selon le groupe de débits auquel appartient le débit simulé sur ce bassin (obtenu lors de l'étape 1), pour obtenir un intervalle prédictif, défini par une borne haute et une borne basse (Figure 11).

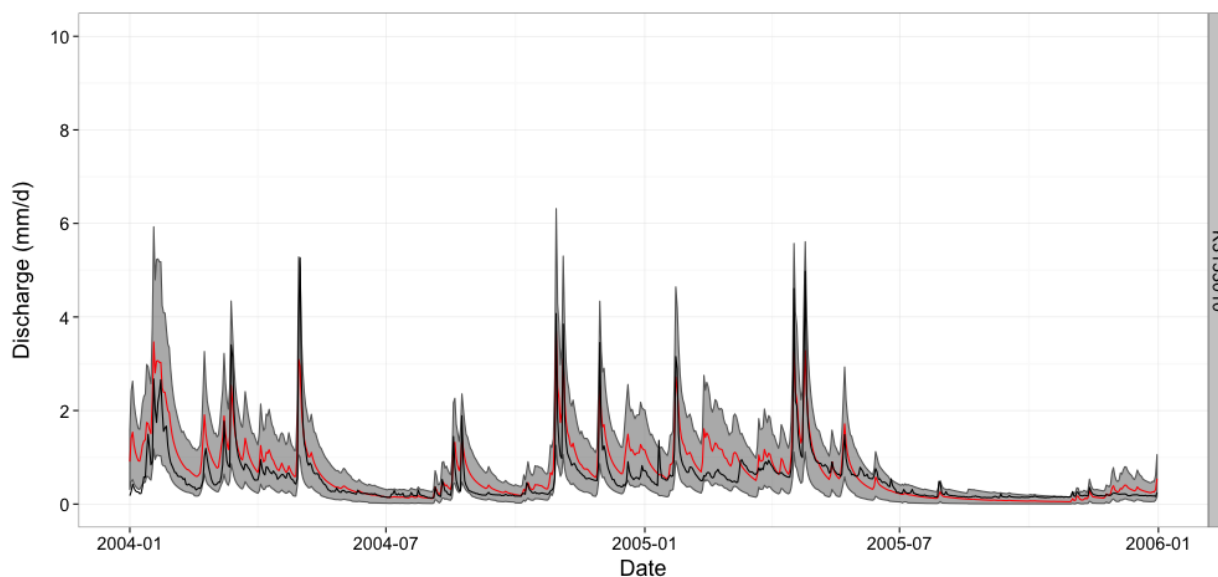


Figure 11 : Représentation graphique de l'incertitude : le débit observé est en noir, le débit (déterministe) calculé en rouge, l'intervalle de confiance à 90% est en grisé.

3.2. Evaluation de la fiabilité de la méthode de calcul de l'intervalle prédictif

Nous avons tenu à évaluer la qualité de l'intervalle de confiance à 90% qui sera proposé à l'issue de cette action. Deux figures permettent de se faire une idée de la vraisemblance des intervalles proposés :

- La Figure 12 illustre la fiabilité de la méthode de calcul des incertitudes proposée : dans plus des 75% des cas, l'intervalle de confiance proposé par notre méthode est fiable, dans le sens où une comparaison avec les débits observés nous donne entre 80 et 100% des observations comprises dans l'intervalle (qui doit en théorie comprendre 90% de ces observations).

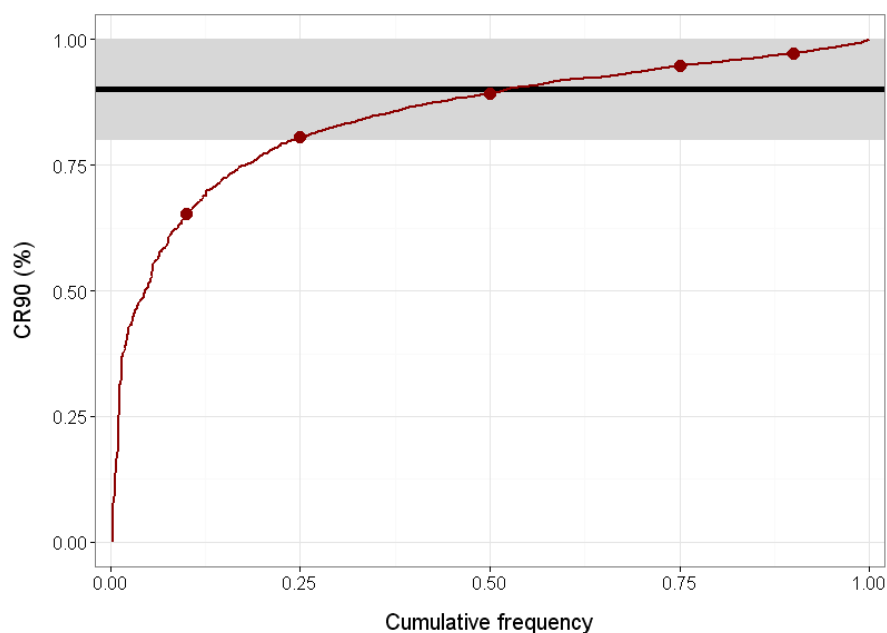


Figure 12 : Pourcentage des observations comprises dans l'intervalle de confiance à 90% (la ligne noire présente l'objectif théorique de 90%, la bande grisée un intervalle acceptable : 80 à 100%). La courbe rouge correspond à la méthode de régionalisation utilisée (proximité géographique).

- La Figure 13 illustre la finesse relative de la méthode de calcul des incertitudes proposée. Cette mesure est intéressante, dans le sens où un intervalle trop large ne serait d'aucune utilité aux utilisateurs. Dans 50 % des cas, l'intervalle moyen est inférieur à 1,1 fois le module, et dans 80% des cas, il est inférieur à 1,6 fois le module.

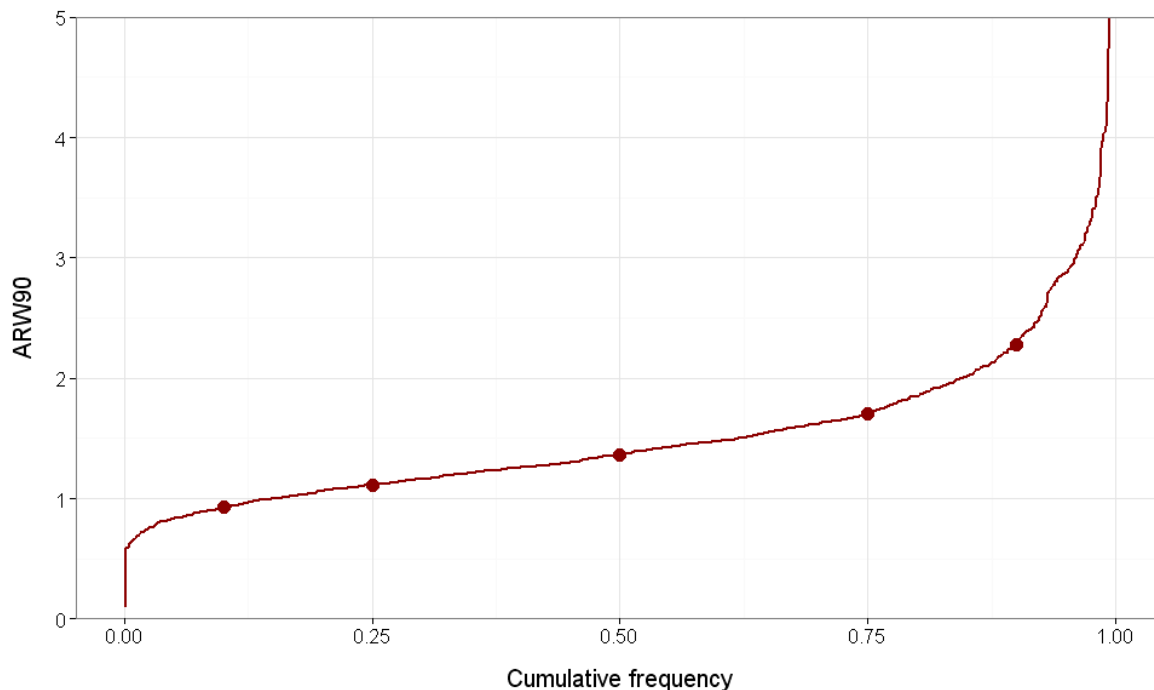


Figure 13 : Largeur relative des intervalles de confiance proposés (largeur moyenne de l'intervalle divisée par le débit moyen). La courbe rouge correspond à la méthode de régionalisation utilisée (proximité géographique).

4. Présentation des simulations réalisées pour un ensemble de stations des réseaux RRP et BDMAP

Nous avons automatisé les procédures présentées en section 2 et 0, et les avons appliquées à 11 225 stations du réseau RRP et de la BDMAP de l'ONEMA. Les fichiers texte générés présentent des simulations sur une période de 55 ans, du 01/08/1958 au 31/07/2012. La Figure 14 présente un aperçu d'un fichier texte généré, composé :

- d'un en-tête donnant des informations sur les bases de données utilisées, le bassin versant receveur, les bassins versants donneurs et les données produites ;
- du corps du fichier, comprenant les séries journalières de débits simulés.

En-tête

#	Date de création du fichier	; 15/01/2014															Description bases de données
#	Créateur	; IRSTEA - UR HBAN - Antony															
#	Données climatiques	; Livraison SAFRAN Octobre 2012															
#	Fournisseur	; Météo-France															
#	Données hydrologiques	; Simulations GR4J															
#	Fournisseur	; IRSTEA - UR HBAN - Antony															
#	Code station BDMAP RRP	; 6260342															Information bassin receveur
#	Surface [km²]	; 54808.15															
#	Coordonnées X-Y station [m, LIIE]	; 797512 ; 2004749															
#	Coordonnées X-Y centroïde [m, LIIE]	; 867862.9 ; 2181399															
#	Quantiles altitude [m]	; Zmin ; Z10 ; Z20 ; Z30 ; Z40 ; Z50 ; Z60 ; Z70 ; Z80 ; Z90 ; Zmax															
#		; 111 ; 206 ; 241 ; 285 ; 346 ; 410 ; 499 ; 635 ; 881 ; 1447 ; 4664															
#	Code des 5 stations jaugées les plus proches	; V2305610 ; V2206010 ; V2035010 ; V2039110 ; V2414010															Information bassins donneurs
#	Champs																
#	Date	; Date (AAAA-MM-JJ) [-]															Description et unité des champs de données
#	Qsim	; Débit simulé par régionalisation [m3/s]															
#	Qsim_05	; Débit simulé par régionalisation, borne basse (Q05%) [m3/s]															
#	Qsim_95	; Débit simulé par régionalisation, borne haute (Q95%) [m3/s]															
#																	
	Date ; Qsim ; Qsim_05 ; Qsim_95																
	19580801 ;	740.7171 ;	108.94166 ;	1691.83902													
	19580802 ;	743.4237 ;	109.33973 ;	1698.02105													
	19580803 ;	758.5228 ;	111.56045 ;	1732.50823													
	19580804 ;	789.0399 ;	116.04878 ;	1802.21098													
	19580805 ;	835.8958 ;	136.36164 ;	1908.01085													
	19580806 ;	903.9040 ;	147.45597 ;	2063.24597													
	19580807 ;	1003.8039 ;	169.80086 ;	2173.95272													
	19580808 ;	1143.1664 ;	193.37506 ;	2475.77212													
	19580809 ;	1316.2492 ;	223.52283 ;	2521.92215													

Données

Figure 14 : Exemple de fichier texte généré pour une station (code 6260342).

5. Synthèse

En 2012, nous avons pu automatiser la procédure de simulation des débits en site non jaugé proposée par Oudin *et al.* (2008), afin de l'appliquer à un ensemble de 11 225 stations du réseau RRP et de la BDMAP de l'ONEMA. Pour chacune de ces stations, 55 ans de débits ont été simulés (du 01/08/1958 au 31/07/2012) au moyen du modèle pluie-débit GR4J (Perrin *et al.*, 2003), et du module neige CemaNeige (Valéry, 2010) pour les bassins versants considérés comme influencés par la neige. Nous avons également développé une méthode spécifique d'estimation de l'incertitude des simulations en site non jaugé (Bourgin, 2014), afin de fournir des estimations accompagnées d'un intervalle de confiance à 90%. Ainsi, en chacun des 11 225 points non jaugés, trois séries de débits journalières sont produites sur la période 1958-2012 :

- la série journalière des bornes basses de l'intervalle de confiance à 90%, correspondant au quantile 5%,
- la série journalière de débits issus de la simulation déterministe,
- la série journalière des bornes hautes de l'intervalle de confiance à 90%, correspondant au quantile 95%.

Plusieurs perspectives d'amélioration de la méthode de simulation des débits en site non jaugé existent. Une sélection des bassins versants donneurs pourraient être réalisée en amont du transfert de paramètres, par exemple en fonction des performances du modèle GR4J lors du calage. La transformation du paramètre X4 de GR4J (en fonction des surfaces des bassins donneurs et du bassin receveur) est basée sur une relation établie au pas de temps horaire, et pourrait donc être établie à nouveau au pas de temps journalier. Enfin, l'utilisation d'un modèle hydrologique distribué permettrait d'assurer une cohérence amont-aval sur les points non-jaugés situés sur le même tronçon de rivières.

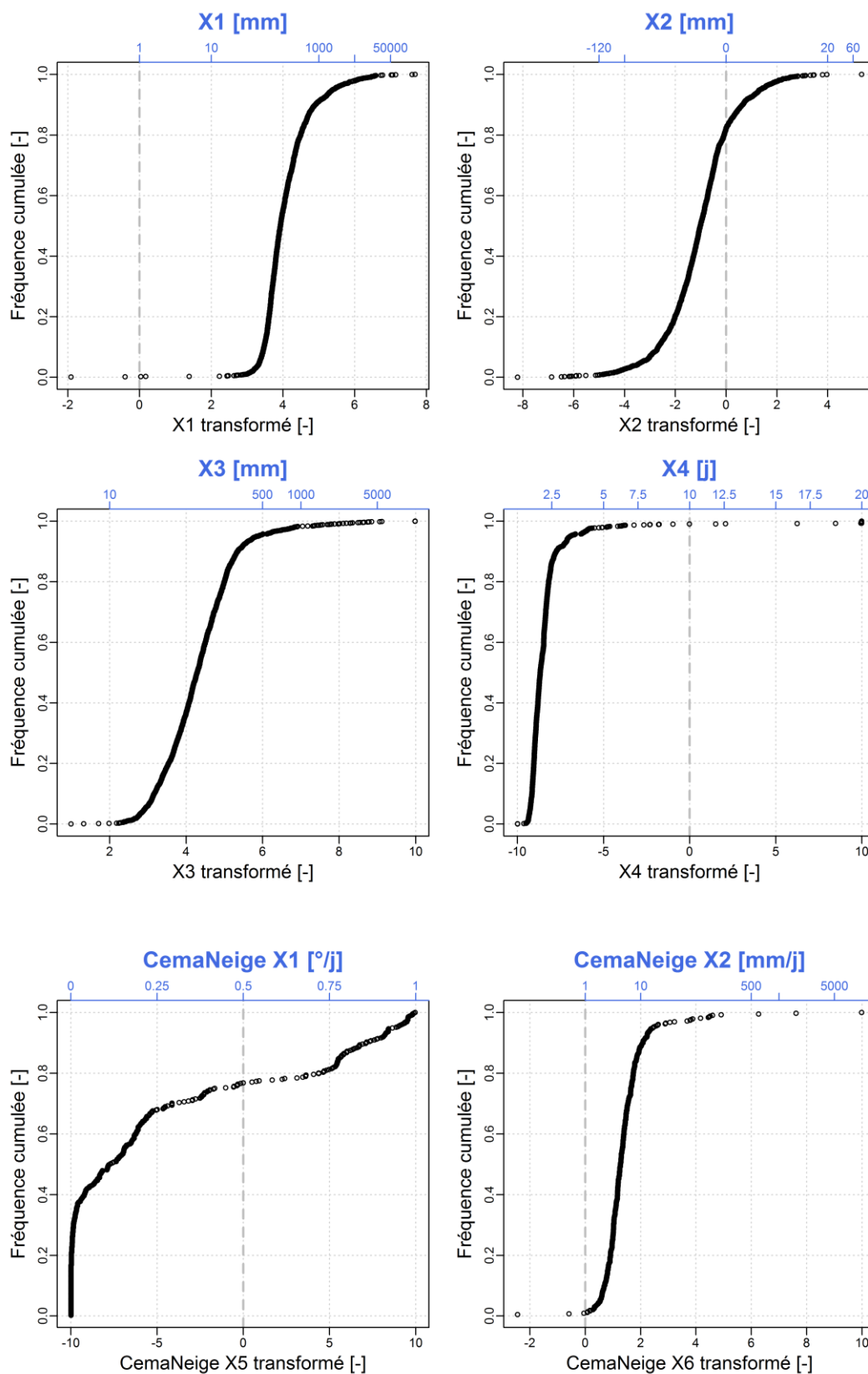
6. Bibliographie

- Bourgin, F. (2014) Comment quantifier l'incertitude prédictive en modélisation hydrologique ? Travail exploratoire sur un grand échantillon de bassins versants. AgroParisTech (Paris), Cemagref (Antony). Soutenance prévue en avril 2014.
- Lobligeois, F. (2014) Mieux connaître la distribution spatiale des précipitations améliore-t-il la modélisation des crues ? Diagnostic sur. AgroParisTech (Paris), Irstea (Antony). Soutenance prévue en mars 2014.
- Mathevet, T., Michel, C., Andréassian, V. & Perrin, C. (2006) A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins. IAHS-AISH publication 211–219.
- Nash, J. E. & Sutcliffe, J. V. (1970) River flow forecasting through conceptual models part I – A discussion of principles. *Journal of Hydrology* **10**(3), 282–290. doi:10.1016/0022-1694(70)90255-6
- Oudin, L., Andréassian, V., Perrin, C., Michel, C. & Moine, N. Le. (2008) Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments. *Water Resources Research* **44**(3). doi:10.1029/2007WR006240
- Perrin, C., Michel, C. & Andréassian, V. (2003) Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology* **279**(1-4), 275–289. doi:10.1016/S0022-1694(03)00225-7
- Valéry, A. (2010) Modélisation précipitations-débit sous influence nivale. Élaboration d'un module neige et évaluation sur 380 bassins versants. AgroParisTech (Paris), Cemagref (Antony).

7. Table des illustrations

Figure 1 : Localisation des 11125 sites RRP et BDMAP (à gauche) et base de données des 2019 bassins versants jaugés utilisés comme donneurs dans le transfert de paramètres (à droite).	9
Figure 2 : « Donneurs » et « receveurs » : le bassin versant receveur (considéré comme non jaugé) est figuré en gris, les cinq bassins versants voisins jaugés donneurs sont en rouge, en bleu on a représenté les autres bassins versants jaugés	10
Figure 3 : Distances moyennes estimées pour chaque bassin receveur entre le bassin receveur et les cinq bassins donneurs les plus proches.....	11
Figure 4 : Localisation et massif montagneux (Alsace, Jura, Massif Central, Alpes du Nord, Alpes du Sud ou Pyrénées) des 432 bassins donneurs considérés comme influencés par la neige (à gauche) et distributions des valeurs de paramètres de CemaNeige par massif montagneux (à droite). Les boîtes à moustaches indiquent les quantiles 0.10, 0.25, 0.50, 0.75 et 0.90 des valeurs de paramètres.	12
Figure 5 : Simulation d'ensemble obtenue à partir du transfert des paramètres des cinq bassins les plus proches (a- cinq simulations distinctes issues des mêmes données d'entrée et de cinq jeux de paramètres différents, b- avec la moyenne des simulations en gris).	13
Figure 6 : Performances de GR4J en calage sur les 2019 bassins versants jaugés. Le critère de calage et le critère d'évaluation des performances est le critère de Nash et Sutcliffe, estimé sur les racines carrées des débits, puis borné entre -1 et 1 selon la transformation C2M proposée par Mathevet <i>et al.</i> (2006).....	15
Figure 7 : Performances de la méthode de reconstitution évaluée sur les 2019 bassins versants jaugés. Le critère de calage et le critère d'évaluation des performances est le critère de Nash et Sutcliffe, estimé sur les racines carrées des débits, puis borné entre -1 et 1 selon la transformation C2M proposée par Mathevet <i>et al.</i> (2006).....	15
Figure 8 : Débits observés et simulations en mode non jaugé obtenus pour les bassins versants donneurs.....	16
Figure 9 : Erreurs observées sur la reconstitution des débits en mode non jaugé, réparties par classes de débits.	17
Figure 10 : Marges d'erreur regroupées par classe de débits. Les points rouges figurent les quantiles 0.05 et 0.95 des erreurs empiriques, qui sont utilisés pour définir l'intervalle de confiance à 90%.	17
Figure 11 : Représentation graphique de l'incertitude : le débit observé est en noir, le débit (déterministe) calculé en rouge, l'intervalle de confiance à 90% est en grisé.	18
Figure 12 : Pourcentage des observations comprises dans l'intervalle de confiance à 90% (la ligne noire présente l'objectif théorique de 90%, la bande grisée un intervalle acceptable : 80 à 100%). La courbe rouge correspond à la méthode de régionalisation utilisée (proximité géographique).	18
Figure 13 : Largeur relative des intervalles de confiance proposés (largeur moyenne de l'intervalle divisée par le débit moyen). La courbe rouge correspond à la méthode de régionalisation utilisée (proximité géographique).....	19
Figure 14 : Exemple de fichier texte généré pour une station (code 6260342).	20

8. Annexe 1 : Distribution de paramètres de GR4J et de CemaNeige, obtenues après calage sur 2019 et 432 bassins versants jaugeés, respectivement.



Irstea

1, rue Pierre Gilles de Gennes
CS 10030
92761 Antony
01 40 96 61 21
www.irstea.fr

Onema

Hall C – Le Nadar
5, square Félix Nadar
94300 Vincennes
01 45 14 36 00
www.onema.fr