

Amélioration des prévisions hydrologiques d'ensemble dans un contexte de modélisation multi-modèle

par

Patrice DION

MÉMOIRE PAR ARTICLE PRÉSENTÉ À L'ÉCOLE DE TECHNOLOGIE
SUPÉRIEURE COMME EXIGENCE PARTIELLE À L'OBTENTION DE
LA MAÎTRISE AVEC MÉMOIRE EN GÉNIE DE LA CONSTRUCTION
M. Sc. A.

MONTREAL, LE 11 DÉCEMBRE 2020

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC



Patrice Dion, 2020



Cette licence [Creative Commons](https://creativecommons.org/licenses/by-nc-nd/4.0/) signifie qu'il est permis de diffuser, d'imprimer ou de sauvegarder sur un autre support une partie ou la totalité de cette œuvre à condition de mentionner l'auteur, que ces utilisations soient faites à des fins non commerciales et que le contenu de l'œuvre n'ait pas été modifié.

PRÉSENTATION DU JURY
CE MÉMOIRE A ÉTÉ ÉVALUÉ
PAR UN JURY COMPOSÉ DE :

M. Richard Arsenault, directeur de mémoire
Département de génie de la construction à l'École de technologie supérieure

M. Jean-Luc Martel, codirecteur de mémoire
Département de génie de la construction à l'École de technologie supérieure

Mme Annie Levasseur, présidente du jury
Département de génie de la construction à l'École de technologie supérieure

Mme Annie Poulin, membre du jury
Département de génie de la construction à l'École de technologie supérieure

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 9 DÉCEMBRE 2020

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Je dois remercier dans un premier temps mon directeur de maîtrise, Richard Arsenault, sans qui je n'aurais probablement pas décidé de continuer aux études supérieures. Richard a été mon enseignant à trois occasions durant mon baccalauréat et cela a grandement contribué à développer ma passion pour l'hydrologie. Il a toujours su me captiver par ses connaissances lui permettant de répondre à mes nombreuses questions et me motiver à développer mon potentiel de chercheur. Merci Richard de m'avoir fait découvrir tant de choses et d'avoir été une source d'inspiration durant toute ma maîtrise.

Je dois ensuite remercier Jean-Luc, qui a agi comme codirecteur et pour qui j'ai été auxiliaire d'enseignement durant trois sessions. Tout comme Richard, ton niveau de connaissance en hydrologie m'a toujours impressionné et j'aurais aimé travailler plus souvent avec toi au laboratoire HC3 de l'université. Merci d'avoir passé autant de temps à m'aider dans plusieurs sphères de ma jeune vie professionnelle. Tu as fait de moi un meilleur chercheur et je suis clairement plus habile pour écrire un article scientifique grâce à toi.

J'aimerais aussi remercier les autres professeurs du laboratoire HC3 de l'ÉTS. Merci dans un premier temps à François Brissette, pour l'organisation de l'école du climat en 2019 et pour la distraction que m'a procuré ton utilisation de ma balle anti-stress (elle est maintenant à toi). Finalement, merci à Annie Poulin et Michel Baraër pour leur qualité d'enseignement durant mon cours de ressources hydriques et leur bonne humeur quotidienne!

Mon parcours aux cycles supérieurs n'aurait pas été aussi facile sans les sources de financement auxquels j'ai eu la chance de profiter. Merci à Richard et à l'ÉTS pour votre financement tout au long de ma maîtrise. J'aimerais aussi remercier spécialement la Fondation Molson et le Fonds de développement de l'ÉTS pour leur généreuse contribution à mon cheminement scolaire. Vous contribuez tous à soutenir les étudiants dans leur ambition de carrière tout en permettant à l'université de se démarquer au niveau de la qualité et des connaissances de leurs finissants.

Je me dois de terminer mes remerciements par les personnes qui ont été les plus importantes dans ma vie durant ces deux dernières années, c'est-à-dire mes amis et ma famille. Merci à mon groupe de soutien psychologique à l'université, Jonathan, Audrey, Andréa, Samuel, Guillaume et Khalil pour nos pauses journalières, nos fous rires et nos débats. Je dois spécialement remercier Roxanne et Magali d'avoir été mes deux piliers à l'université et à l'extérieur. Sans vous, je n'aurais pas autant apprécié mon passage aux études supérieures et ma motivation n'aurait pas toujours été aussi élevée! J'aimerais ensuite souligner le soutien de mes amis, maintenant sur le marché du travail, et faisant partie de ma deuxième famille : Mathieu, Émilie, Pierre-Karl, Clotilde, Dominik, Marie-Ève, Rachel, Philippe-Olivier et Francis. Merci d'avoir toujours cru en moi et de m'avoir tous soutenu sans exception. Finalement, merci à ma famille, sans qui je ne serais pas qui je suis aujourd'hui : Camille, ma partenaire de vie qui a vécu mes hauts et mes bas et qui a toujours été d'un réconfort absolu, mes grands-parents, ma tante Sylvie, mon oncle Robert, mon cousin Philippe, qui êtes toujours fiers de moi et pour terminer, ma sœur, ma mère et mon père, qui me soutiennent depuis toujours et qui ne pourraient être plus parfaits.

Amélioration des prévisions hydrologiques d'ensemble dans un contexte de modélisation multi-modèle

Patrice DION

RÉSUMÉ

Que ce soit en contexte de gestion de réservoirs hydroélectrique, gestion des inondations ou toutes autres opérations visant à contrôler des afflux d'eau, les hydrologues se servent de prévisions hydrologiques afin de prendre leurs décisions opérationnelles. Pour ce faire, ils doivent considérer les risques financiers, environnementaux et avant tout, ceux liés à la sécurité du public. La prévision d'ensemble est un type de prévision hydrologique qui permet d'évaluer le risque en tenant compte des incertitudes de ces prévisions. Toutefois, les prévisions hydrologiques d'ensemble sont souvent biaisées et leur distribution est souvent sous-dispersée lorsqu'elles sont comparées aux observations. Cette problématique provient principalement de l'incertitude des prévisions météorologiques d'ensemble servant à alimenter des modèles hydrologiques pour en obtenir des prévisions de débits en rivières, et aussi du ou des modèles utilisés. Ces prévisions erronées peuvent ainsi contribuer à des prises de décisions opérationnelles pouvant mettre en danger le public et les infrastructures.

Ce projet de recherche présente une nouvelle méthodologie qui vise à améliorer la précision des prévisions hydrologiques d'ensemble utilisées en gestion des ressources hydriques. Cette méthodologie a été évaluée sur cinq bassins versants du partenaire industriel *Rio Tinto* dans la région du Lac-Saint-Jean, au Québec. Les prévisions météorologiques d'ensemble (50 membres) de l'ECMWF de 2015 à 2019 ont été utilisées afin de renseigner huit modèles hydrologiques globaux sur un horizon de neuf jours. Suite à la calibration des modèles hydrologiques, une méthode d'assimilation de données basée sur les filtres d'ensemble de Kalman (EnKF) est employée afin de modifier les conditions initiales de chaque modèle hydrologique dans le but de représenter les différentes sources d'incertitudes au niveau des observations. Ensuite, un post-traitement des ensembles hydrologiques de chaque modèle est effectué en utilisant la méthode de correction de biais par quantiles (QM). Ces mêmes ensembles hydrologiques sont finalement regroupés en un grand ensemble multi-modèle dans l'objectif d'avoir un meilleur échantillonnage de l'incertitude totale, et ainsi uniformiser sa distribution.

Afin d'évaluer les résultats obtenus par la méthodologie et de ses différentes étapes, plusieurs critères de performance sont employés. Les diagrammes de Talagrand, deux métriques de quantification de la fiabilité et de la précision, ainsi que le test statistique de Kolmogorov-Smirnov, servent à évaluer la robustesse de la méthodologie. Cette évaluation est effectuée sur quatre périodes distinctes, représentant différents régimes hydrologiques sur les bassins versants. Malgré des difficultés durant la période de la crue printanière, les résultats obtenus indiquent que chaque étape de la méthodologie permet d'améliorer la précision des prévisions hydrologiques d'ensemble sur les cinq bassins versants à l'étude. Finalement, l'utilisation du multi-modèle améliore significativement la précision et l'étendue des ensembles

VIII

hydrologiques sur un horizon de neuf jours. Ce projet démontre qu'il est donc possible de corriger les biais et d'améliorer la fiabilité des prévisions hydrologiques d'ensemble.

Mots-clés : prévisions hydrologiques d'ensemble, modélisation hydrologique, assimilation de données, post-traitement des débits, multi-modèle, sous-distribution

Improvement of hydrological ensemble predictions in a multi-model modeling context

Patrice DION

ABSTRACT

Whether it is in the context of hydroelectric reservoir management, flood management or any other operation aimed at controlling water inflows, hydrologists use hydrological forecasts to make their operational decisions. To do so, they must consider financial, environmental and, above all, public safety risks. Ensemble forecasting is a type of hydrological forecasting that allows for risk assessment by taking into account the uncertainties in these forecasts. However, hydrological ensemble forecasts are often biased, and their distribution is often undispersed when compared to observations. This problem stems mainly from the uncertainty in the meteorological ensemble forecasts used to feed hydrological models to obtain forecasts of river streamflow, and also from the model(s) used. These erroneous predictions can thus contribute to operational decisions that can endanger the public and infrastructure.

This research project presents a new methodology that aims to improve the accuracy of ensemble hydrological forecasts used in water resource management. This methodology has been evaluated on five catchments of the industrial partner *Rio Tinto* in the Lac-Saint-Jean region of Quebec. ECMWF weather ensemble forecasts (50 members) from 2015 to 2019 were used to inform eight global hydrological models over a nine-day horizon. Following the calibration of the hydrological models, a data assimilation method based on Kalman ensemble filters (EnKF) is used to modify the initial conditions of each hydrological model to represent the different sources of uncertainty in the observations. Then, a post-processing of the hydrological ensembles of each model is performed using the quantile bias correction (QM) method. These same hydrological ensembles are finally grouped into a large multi-model ensemble with the objective of having a better sampling of the total uncertainty, and thus to standardize its distribution.

In order to evaluate the results obtained by the methodology and its different steps, several performance criteria are used. Talagrand diagrams, two metrics for quantifying reliability and precision, as well as the Kolmogorov-Smirnov statistical test, are used to evaluate the performance of the methodology. This evaluation is performed over four distinct periods, representing different hydrological regimes over the catchments. Despite difficulties during the spring freshet period, the results obtained indicate that each step of the methodology improves the accuracy of hydrological ensemble forecasts over the five catchments under study. Finally, the use of the multi-model significantly improves the accuracy and extent of hydrological ensembles over a nine-day horizon. This project demonstrates that it is therefore possible to correct biases and improve the reliability of hydrological ensemble forecasts.

Keywords: hydrological ensemble forecasting, hydrological modelling, data assimilation, streamflow post-processing, multi-model, under-dispersion

TABLE DES MATIÈRES

	Page
INTRODUCTION	1
CHAPITRE 1 REVUE DE LITTÉRATURE.....	7
1.1 Prévisions météorologiques	7
1.1.1 Prévision déterministe.....	8
1.1.2 Prévision probabiliste.....	8
1.1.3 Utilisation des prévisions météorologiques	9
1.2 Prévisions hydrologiques.....	10
1.3 Assimilation de données	15
1.4 Prétraitement et post-traitement.....	18
1.5 Multi-modèle hydrologique	20
1.6 Objectifs de recherche.....	21
CHAPITRE 2 HYDROLOGICAL ENSEMBLE FORECASTING USING A MULTI-MODEL FRAMEWORK	23
2.1 Abstract.....	23
2.2 Introduction.....	24
2.3 Study area.....	27
2.4 Methodology	29
2.4.1 Overview.....	29
2.4.2 Forecast data	31
2.4.3 Hydrological model calibration	31
2.4.4 Data assimilation.....	33
2.4.5 Hydrological modeling	34
2.4.6 Post-processing	34
2.4.7 Hydrological multi-model approach.....	35
2.4.8 Performance metrics	35
2.5 Results.....	38
2.5.1 Calibration of the hydrological models.....	38
2.5.2 Impact of data assimilation and bias correction on individual ESP performance	39
2.5.3 Comparison between individual ESPs and the GESP performances.....	40
2.5.4 Impact of the catchment surface area.....	43
2.5.5 Impact of data assimilation parameterization	45
2.5.6 Evaluation of the short-term forecast performance	46
2.5.7 Impact of the hydrological season	48
2.6 Discussion.....	48
2.6.1 Performance of data assimilation using EnKF.....	49
2.6.2 Post-processing performance with QM bias correction.....	50
2.6.3 Multi-model performance analysis	51
2.6.4 Impact of the hydrological season on the forecast.....	53

2.6.5	Overall methodology performance analysis	53
2.6.6	Limitations	54
2.7	Conclusion and recommendations	55
2.8	Acknowledgments.....	56
CHAPITRE 3 DISCUSSION		59
3.1	Analyse de la performance de la méthode d'assimilation de données d'EnKF.....	59
3.2	Analyse de la performance de la méthode de post-traitement	62
3.3	Analyse de la performance de la méthode multi-modèle.....	65
3.4	Impact des différentes saisons sur les prévisions d'ensemble	67
3.5	Analyse de la méthodologie globale	68
3.6	Limitations	69
3.6.1	Prévisions météorologiques d'ensemble.....	69
3.6.2	Modèles hydrologiques	70
3.6.3	Bassins versants	70
CONCLUSION.....		71
RECOMMANDATIONS		73
ANNEXE I	DIAGRAMMES DE TALAGRAND EN MATÉRIEL SUPPLÉMENTAIRE	75
ANNEXE II	BOÎTES À MOUSTACHES DES VALEURS CRPSS EN MATÉRIEL SUPPLÉMENTAIRE	79
LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES.....		83

LISTE DES TABLEAUX

	Page
Tableau 2.1	Hydrological characteristics of catchments based on observed streamflow from 1953 to 2019 provided by <i>Rio Tinto</i>28
Tableau 2.2	Number of parameters and conceptual reservoirs for each of the eight hydrological models used in this study32
Tableau 2.3	KGE scores for the eight hydrological models calibrated39

LISTE DES FIGURES

	Page
Figure 0.1	Émissions de GES d'origine anthropique par secteurs d'activité en 2010. Le secteur de l'énergie comprend la production d'électricité, de chaleur et des autres énergies. Les secteurs placés dans la section « Indirect CO ₂ emissions » représentent les utilisateurs de l'énergie électrique et de la production de chaleur. Le terme AFOLU correspond à « agriculture, forestry and other land use »2
Figure 0.2	Exemple d'un hydrogramme représentant une prévision hydrologique déterministe et le débit observé y étant associé sur une période de temps indéterminée4
Figure 0.3	Exemple d'un hydrogramme représentant une prévision hydrologique probabiliste et le débit observé y étant associé sur une période de temps indéterminée5
Figure 1.1	Exemple d'un ensemble hydrologique prévisionnel sous-dispersé. La dispersion de l'ensemble n'est pas assez grande pour inclure tous les débits observés13
Figure 1.2	Exemples de dispersion d'un ensemble de 100 membres sur des diagrammes de Talagrand. Le « Nombre de membres » en ordonnée du diagramme représente la fréquence des observations, c'est-à-dire le nombre de fois où l'observation se situe dans une des 10 classes en abscisse du diagramme de Talagrand.....14
Figure 1.3	Exemple d'un ensemble hydrologique prévisionnel systématiquement biaisé négativement. Les observations ne sont pas incluses dans l'ensemble.....14
Figure 2.1	Location and topography of the study's five catchments28
Figure 2.2	Proposed methodology implemented in this study30
Figure 2.3	Talagrand diagrams representing rank (in percentiles) occupied by observed streamflows in the hydrological forecast ensemble using the GR4J model on the LSJ catchment. "OL" refers to the open-loop simulation. "DA" refers to the validation members of the 1250 ensemble members after performing data assimilation and "QM" refers to the same members after performing data assimilation and post-processing with quantile-mapping. Blue (red) diagrams indicate that the K-S test failed to reject (rejected) the hypothesis of uniformity ($\alpha = 0.05$)40

Figure 2.4	Talagrand diagrams representing rank (in percentiles) occupied by observed streamflows in the ESPs of eight different hydrological models after performing DA and QM on the LSJ catchment. Blue (red) diagrams indicate that the K-S test failed to reject (rejected) the hypothesis of uniformity ($\alpha = 0.05$).....	42
Figure 2.5	Talagrand diagrams representing rank (in percentiles) occupied by observed streamflows in the hydrological multi-model (GESP) forecast after performing DA and QM on the five catchments. Blue (red) diagram indicates that the K-S test failed to reject (rejected) the hypothesis of uniformity ($\alpha = 0.05$).....	44
Figure 2.6	Talagrand diagrams representing rank (in percentiles) occupied by observed streamflows in the hydrological multi-model (GESP) forecasts on the LSJ catchment after performing the DA manipulation only, with different standard deviation values for observed streamflows (Q) and the precipitation forecasts (P). Blue (red) diagrams indicate that the K-S test failed to reject (rejected) the hypothesis of uniformity ($\alpha = 0.05$).....	46
Figure 2.7	Boxplot of the <i>CRPSS</i> values of the members of the GR4J hydrological model ensemble and multi-model (GESP) ensemble on the LSJ catchment. <i>CRPSS</i> greater than zero indicates an improvement in overall forecast accuracy, as compared to the GR4J open-loop (OL) simulations. Members with <i>CRPSS</i> values smaller than -1 are considered very poor and are purposely left out to avoid overloading the graph	47
Figure 3.1	Diagrammes de Talagrand représentant le rang (en percentiles) occupé par les débits observés dans l'ensemble des prévisions hydrologiques à l'aide du modèle IHACRES sur le bassin versant LSJ. "OL" fait référence à la simulation de référence. "DA" réfère aux membres de validation des 1250 membres de l'ensemble après avoir effectué l'assimilation des données et "QM" réfère aux mêmes membres après avoir effectué l'assimilation des données et le post-traitement avec la correction par quantiles. Les diagrammes en bleu (rouge) indiquent que l'hypothèse d'uniformité n'est pas rejetée (rejetée) par le test K-S ($\alpha = 0,05$)	60

Figure 3.2	Boîtes à moustaches des valeurs CRPSS des membres de l'ensemble hydrologique du modèle IHACRES et de l'ensemble multi-modèles (GESP) sur le bassin versant LSJ. Un CRPSS supérieur à zéro indique une amélioration de la précision globale des prévisions par rapport aux simulations de référence (OL) d'IHACRES. Les membres dont les valeurs du CRPSS sont inférieures à -1 sont considérés comme très mauvais et ne sont pas indiqués afin d'éviter de surcharger le graphique61
Figure 3.3	Diagrammes de Talagrand représentant le rang (en percentiles) occupé par les débits observés dans les ESPs des huit modèles hydrologiques différents après avoir effectué le DA et le QM sur le bassin versant PD. Les diagrammes en bleu (rouge) indiquent que l'hypothèse d'uniformité n'est pas rejetée (rejetée) par le test K-S ($\alpha = 0,05$)64

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

ABDU	Average Bin Distance to Uniformity
AM	Saison de la fonte de neige – Avril à Mai
CD	Bassin versant Chute-du-Diable
CEQUEAU	Modèle hydrologique développé par l' <i>Institut national de la recherche scientifique</i> (INRS) de l'Université du Québec
CRPS	Continuous rank probability score
CRPSS	Continuous rank probability skill score
CS	Bassin versant Chute-à-la-Savane
DA	Assimilation de données
DJFM	Saison hivernale – Décembre à Mars
ECCC	Environnement et Changement climatique Canada
ECMWF	European centre for medium-range weather forecasts
EnKF	Filtres d'ensemble de Kalman (Ensemble Kalman filter)
ESP	Prévision hydrologique d'ensemble (Ensemble streamflow prediction)
GES	Gaz à effet de serre
GESP	Grand ensemble multi-modèle (Grand ensemble streamflow prediction)
GR4J	Modèle hydrologique Génie rural à 4 paramètres – Journalier, développé par le <i>Centre national du machinisme agricole du génie rural, des eaux et des forêts</i> (CEMAGREF) en France, maintenant devenu l' <i>Institut national de recherche pour l'agriculture, l'alimentation et l'environnement</i> (INRAE)
HBV	Modèle hydrologique « Hydrologiska Byråns Vattenbalansavdelning model » développé par le <i>Swedish meteorological and Hydrological Institute</i> (SMHI)
HOOPLA	Hydrological Prediction Laboratory

HYMOD	Modèle hydrologique « HYdrological MODel » développé par l' <i>Institute of Hydrology</i> à Wallingford en Angleterre.
IHACRES	Modèle hydrologique développé conjointement par le <i>Centre for Ecology & Hydrology</i> en Angleterre et le <i>Centre for Resource and Environmental Studies</i> à l'Université national d'Australie.
IPCC	Groupe d'experts intergouvernemental sur l'évolution du climat (Intergovernmental Panel on Climate Change)
JJA	Saison d'été – Juin à Août
KGE	Kling–Gupta Efficiency
LM	Bassin versant Lac-Manouane
LSJ	Bassin versant Lac-Saint-Jean
MOHYSE	Modèle hydrologique simplifié à l'extrême développé par l' <i>Institut national de la recherche scientifique</i> (INRS) de l'Université du Québec
NWPE	Prévision météorologique d'ensemble (Numerical weather prediction ensemble)
PD	Bassin versant Passes-Dangereuses
PET	Évapotranspiration potentielle
SAR	Snow accounting routine
SIMHYD	Modèle hydrologique simplifié à partir du modèle HYDROLOG et développé par le <i>Cooperative Research Centre for Catchment Hydrology</i> en Australie
SON	Saison d'automne – Septembre à Novembre
TOPMODEL	Modèle hydrologique topographique développé par l' <i>Institute of Hydrology</i> à Wallingford en Angleterre.

INTRODUCTION

La modification rapide des écosystèmes par l'homme et les changements climatiques provoqués par les émissions de gaz à effet de serre (GES) d'origine anthropique représenteront certainement un des plus grands défis auxquels sera confrontée l'humanité au cours du prochain siècle. Leurs impacts se font déjà ressentir partout sur la planète (IPCC et al., 2019). Selon cette source, depuis la période préindustrielle (1850-1900), la modification des écosystèmes et les changements climatiques provoquent une hausse de la température moyenne de la surface des continents, déjà évaluée à +1,5 °C. Cette hausse de température engendre, par exemple, des épisodes de chaleur extrêmes, une accélération de la désertification des milieux naturels et une modification du cycle hydrologique. La production d'énergie est l'un des secteurs qui contribuent à l'augmentation des GES. En 2010, ce secteur a représenté 34,6 % de tous les GES émis sur la planète (IPCC et al., 2014), plus que n'importe quel autre secteur (voir figure 0.1). Plusieurs méthodes alternatives d'énergie renouvelable, telles que le solaire, l'éolien et l'hydroélectricité, sont disponibles et continuellement en développement pour en améliorer la performance. Ce projet de recherche se concentre principalement sur cette dernière, de par sa forte présence au Québec.

L'énergie hydroélectrique a plusieurs impacts sur les cours d'eau, les milieux environnants et les populations. Malgré que cette source d'énergie contribue à la lutte aux changements climatiques en produisant de l'électricité à partir d'une source renouvelable avec des émissions de GES moindre que les sources d'énergie plus traditionnelles (p. ex. : le charbon et le gaz naturel), la construction de nouvelles centrales et des ouvrages connexes (p. ex. : les digues, les barrages, l'évacuateur de crue, etc.) affectent considérablement les écosystèmes. Il est donc primordial d'exploiter le plus efficacement possible les centrales hydroélectriques existantes dans le but de maximiser leur production énergétique et de retarder la construction de nouveaux ouvrages. Une gestion efficace de l'eau permet aussi de réduire les risques pour le public et les infrastructures environnantes. Il existe deux principaux types de centrales hydroélectriques, c'est-à-dire au fil de l'eau et à réservoir. La centrale hydroélectrique à réservoir implique la construction de plusieurs digues et barrages afin de créer une réserve d'eau artificielle en

quantité suffisante pour produire de l'électricité à longueur d'année. La centrale hydroélectrique au fil de l'eau est plus limitée puisqu'elle ne crée pas de réserve d'eau et sa production d'énergie dépend directement du débit d'eau de la rivière exploitée en temps réel. Ce projet de recherche se concentre sur l'amélioration de la gestion de l'eau en contexte opérationnel sur ces deux types de centrales, mais les conclusions pourront aussi bien servir en contexte de gestion des inondations ou tout autre type de gestion hydrique sur des rivières naturelles.

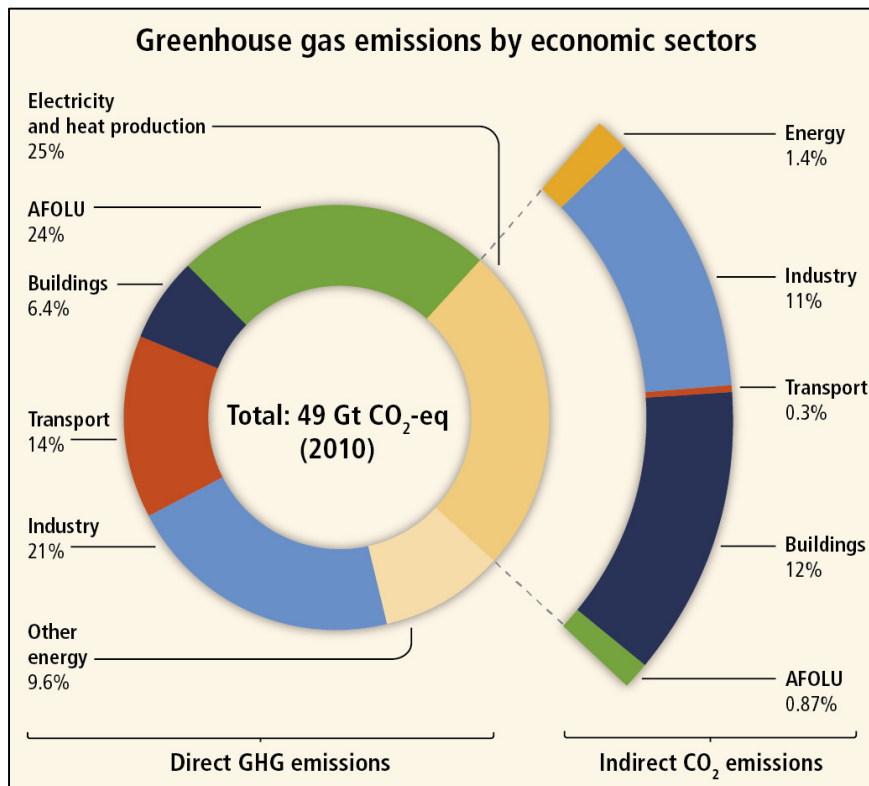


Figure 0.1 Émissions de GES d'origine anthropique par secteurs d'activité en 2010. Le secteur de l'énergie comprend la production d'électricité, de chaleur et des autres énergies. Les secteurs placés dans la section « Indirect CO₂ emissions » représentent les utilisateurs de l'énergie électrique et de la production de chaleur. Le terme AFOLU correspond à « agriculture, forestry and other land use »

Tirée du rapport synthèse AR5 (p.47) IPCC et al. (2014)

La modélisation hydrologique est l'outil utilisé par les gestionnaires de centrales hydroélectriques afin de prévoir les apports naturels en eau alimentant les réservoirs. Elle est aussi utilisée pour la prévention des inondations ou toutes autres opérations visant à contrôler des afflux d'eau. Le modèle hydrologique simule le cycle de l'eau en utilisant des données météorologiques (précipitations et températures) et permet de prévoir, lorsque des prévisions météorologiques sont utilisées, les débits naturels qui alimenteront le réservoir. Toutefois, ce processus comporte plusieurs sources d'incertitudes provenant notamment de la variabilité naturelle du climat de la Terre, de la nature chaotique du système climatique engendrant des prévisions imparfaites, de la simplification des processus hydrologiques réels et des modèles utilisés (calibration, erreurs dans les données fournies aux modèles, etc.).

La modélisation hydrologique peut servir à effectuer des prévisions hydrologiques déterministes (voir figure 0.2) ou probabilistes (voir figure 0.3). Ces deux types de prévisions reposent sur l'utilisation d'un modèle hydrologique réputé comme très performant sur le bassin versant étudié. Toutefois, à l'opposé de la méthode probabiliste, la méthode déterministe ne donne qu'une seule prévision de débits, extraite à partir de la simulation du modèle hydrologique. Les gestionnaires l'utilisent alors pour tenter d'optimiser la production d'énergie ou prévenir les inondations, sans pouvoir tenir compte des incertitudes sur cette prévision. Les résolutions spatiale et temporelle de ce type de prévision sont généralement plus précises que celles provenant de la méthode probabiliste puisqu'elles demandent moins de puissance de calcul et peuvent donc être produites à haute résolution. Toutefois, en contexte de gestion des réservoirs d'eau, elle peut parfois donner une fausse impression de certitude (Krzysztofowicz, 2001) pouvant se solder par une mauvaise optimisation de la ressource en eau si la prévision s'avère inexacte.

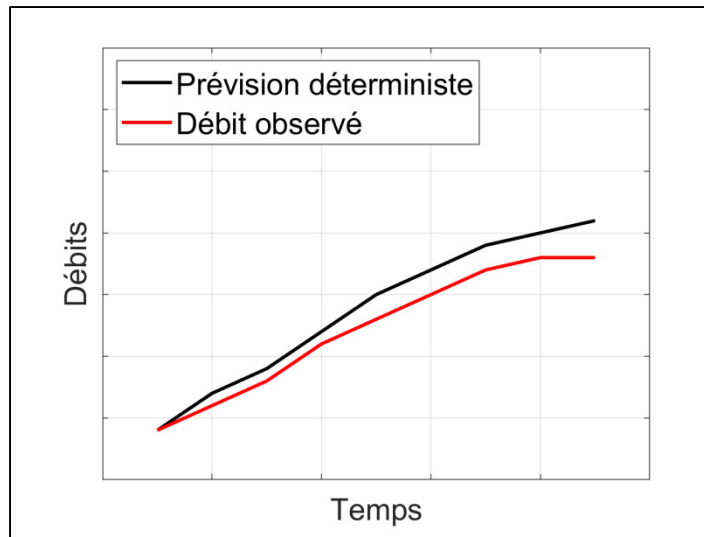


Figure 0.2 Exemple d'un hydrogramme représentant une prévision hydrologique déterministe et le débit observé y étant associé sur une période de temps indéterminée

Pour leur part, les prévisions hydrologiques probabilistes, regroupant un ensemble de plusieurs simulations hydrologiques, permettent une optimisation de la ressource en eau en quantifiant l'incertitude, ce qui n'est pas possible avec la prévision déterministe. Cela permet aux gestionnaires d'être mieux outillés en contexte de gestion des risques en ayant à leur disposition une enveloppe de probabilité du volume d'apports naturels dans leurs réservoirs. Toutefois, l'idée de les utiliser en contexte opérationnel peut être écartée lorsque les gestionnaires ne sont pas outillés pour bien les interpréter. De plus, les membres de l'ensemble de prévisions probabilistes sont souvent biaisés et sous-dispersés, ce qui indique un manque de variabilité des membres et se traduit souvent par une incapacité à englober les observations à venir en contexte prévisionnel. Les prévisions déterministes présentent aussi ces problématiques, mais il n'est alors pas possible de les quantifier.

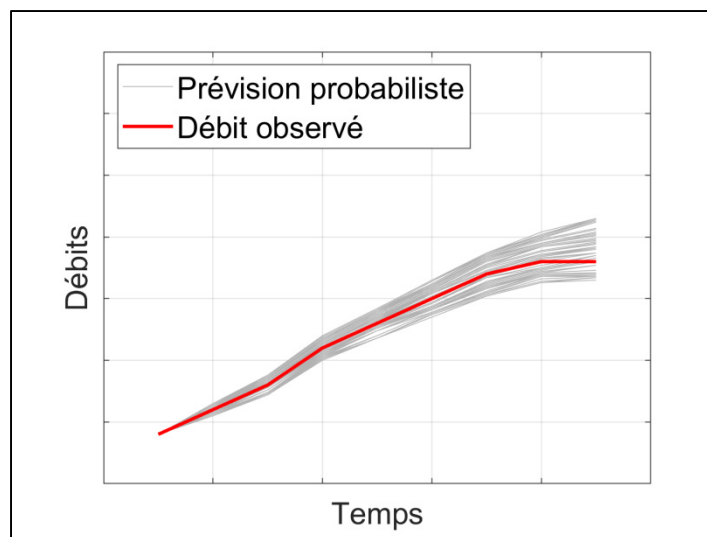


Figure 0.3 Exemple d'un hydrogramme représentant une prévision hydrologique probabiliste et le débit observé y étant associé sur une période de temps indéterminée

L'objectif de cette recherche est d'établir une méthodologie pour utiliser les prévisions hydrologiques d'ensemble en contexte opérationnel de gestion hydrique tout en abordant la problématique du manque de variabilité dans ces prévisions. Pour y parvenir, plusieurs techniques de manipulation des prévisions hydrologiques seront employées, telles que l'assimilation des débits et le post-traitement des prévisions, ainsi que l'utilisation du multi-modèle hydrologique (c.-à-d. des prévisions provenant de plusieurs modèles hydrologiques). Cette recherche a été effectuée en partenariat avec l'entreprise *Rio Tinto*¹, opérant plusieurs complexes hydroélectriques au Canada, notamment au Québec et en Colombie-Britannique. L'entreprise exploite six centrales hydroélectriques dans la région du Saguenay-Lac-Saint-Jean au Québec afin d'y produire près de 50 % (1,26 MT) de sa production mondiale annuelle d'aluminium.

¹ <https://www.riotinto.com/fr-CA/can/>

Ce mémoire est divisé en trois chapitres et renseigne son lecteur sur les différentes étapes réalisées pour parvenir à établir la méthodologie développée. Le premier chapitre présente une revue de la littérature existante sur les différentes techniques d'optimisation des prévisions existantes et leur fonctionnement. Le deuxième chapitre présente un article scientifique soumis pour publication dans le journal scientifique *Hydrology and Earth System Sciences* présentant l'ensemble des résultats de ce projet de recherche. Toutes les informations concernant les données utilisées, les bassins versants à l'étude, la méthodologie, les résultats et leur interprétation, y sont incluses. Le troisième chapitre est une discussion qui vise à approfondir les conclusions de l'article par l'ajout de résultats additionnels les supportant. Le dernier chapitre de ce mémoire est consacré à la conclusion et aux recommandations qui permettront à tous les lecteurs de retenir les points essentiels de la présente recherche.

CHAPITRE 1

REVUE DE LITTÉRATURE

Cette section traitera de l'état de l'art en lien avec ce projet de recherche. Les sujets couverts sont les prévisions météorologiques, les prévisions hydrologiques, l'assimilation de données hydrologiques, le post-traitement des simulations de débit ainsi que l'utilisation de plusieurs modèles hydrologiques en prévision probabiliste (appelée ci-après « multi-modèle »).

1.1 Prévisions météorologiques

Une prévision météorologique représente un outil indispensable à la modélisation hydrologique dans un contexte de production hydroélectrique (Schaake, Hamill, et al., 2007). Celle-ci permet d'alimenter les modèles hydrologiques en leur fournissant principalement les prévisions de précipitations et de température dans le cas de modèles hydrologiques simples (conceptuels). Dans le cas des modèles hydrologiques plus complexes (typiquement à bases physiques), ceux-ci ont généralement recours à un plus grand nombre de variables telles que le vent et l'humidité relative de l'air. La modélisation hydrologique permet ensuite d'obtenir des prévisions de débits, servant notamment à la gestion des réservoirs. La durée des prévisions météorologiques dans ce contexte précis peut s'étendre de quelques minutes à plusieurs mois d'horizon selon l'objectif recherché par le modélisateur.

La prévision météorologique peut être déterministe, soit une seule prévision provenant d'un unique modèle météorologique, ou bien probabiliste (typiquement appelée prévision météorologique d'ensemble ou de l'anglais NWPE²; numerical weather prediction ensemble), soit plusieurs prévisions provenant d'un ou plusieurs modèles météorologiques. Les principaux centres météorologiques, tels qu'Environnement et Changement climatique Canada (ECCC) et l'*European Centre for Medium-Range Weather Forecasts* (ECMWF), fournissent les deux

² Les abréviations anglaises de l'article scientifique seront employées dans tout le mémoire pour faciliter la compréhension.

types de prévisions météorologiques. Leur rôle consiste à émettre, typiquement quelques fois par jour, de nouvelles prévisions météorologiques sur un horizon de temps à court terme (une à 48 heures) ou à moyen terme (trois à 15 jours) sur un territoire et une résolution spatiale prédéfinis. Pour ce faire, ils utilisent un ou plusieurs modèles météorologiques qui permettent de simuler l'état de l'atmosphère terrestre à partir de ses conditions initiales. Ces conditions initiales (p. ex. : l'intensité du rayonnement solaire ou la direction du vent) affectent directement les prévisions météorologiques puisqu'elles modifient la simulation et peuvent ainsi entraîner un changement dans la prévision (p. ex. : la température maximale ou le déplacement d'un orage).

1.1.1 Prévision déterministe

La prévision déterministe est constituée d'une seule série de précipitations et de températures provenant d'un unique modèle météorologique. Il est alors alimenté par les paramètres initiaux représentant l'état de l'atmosphère terrestre avec le niveau de confiance le plus élevé. Cette hypothèse est surtout utilisée pour une prévision à court terme au cours de laquelle les déplacements de masse d'air et autres études de l'atmosphère sont analysés uniquement pour les prochaines heures et sont donc plus prévisibles (Pappenberger et al., 2011). Toutefois, ce type de prévision se bute à des limitations importantes en raison de l'absence d'informations sur l'incertitude, ce qui empêche ses utilisateurs à tenir compte de tous les scénarios météorologiques probables pouvant survenir (Schaake, Hamill, et al., 2007).

1.1.2 Prévision probabiliste

Les prévisions probabilistes sont représentées par un ensemble de scénarios météorologiques équiprobables. En général, lorsqu'une référence aux différentes simulations d'un même NWPE est effectuée, le terme « membre » est employé. Ces NWPEs sont constituées d'un membre de contrôle, typiquement obtenu à partir du modèle météorologique alimenté par les mêmes conditions initiales de l'atmosphère que celles utilisées pour la prévision déterministe, et d'autres membres. Les conditions initiales du membre de contrôle sont perturbées plusieurs fois en fonction de l'incertitude attribuée aux observations afin d'obtenir les autres membres

de l'ensemble. L'avantage principal d'un NWPE est de pouvoir quantifier l'incertitude en explorant la variabilité du système, et ainsi créer un intervalle d'incertitude dans lequel on tente d'inclure les observations. Il permet aussi de mieux simuler les événements extrêmes ayant une probabilité plus faible de se produire (Wetterhall et al., 2014).

1.1.3 Utilisation des prévisions météorologiques

En général, la performance des prévisions déterministes par rapport à celles probabilistes dépend grandement de la qualité des conditions initiales fournies aux modèles météorologiques (Schwanenberg et al., 2015). Ainsi, moins les conditions initiales sont précises, moins l'utilisation des prévisions déterministes est appropriée. Toutefois, dans la littérature, les NWPEs sont de plus en plus utilisées (Brochero et al., 2011; Velázquez et al., 2011). Comme indiqué précédemment, les NWPEs permettent d'ajouter de la variabilité et d'obtenir un ensemble de différentes précipitations et températures qui ont la même probabilité de se produire sur un intervalle de temps défini (p. ex. : 12 heures ou 9 jours). D'ailleurs, plusieurs études effectuées (Abaza et al., 2013; Thibault et al., 2016) utilisent les NWPEs dans un contexte de prévisions hydrologiques dans le but d'ajouter de la variabilité et ainsi augmenter les probabilités que les précipitations et températures observées fassent partie des scénarios de l'ensemble. Même dans les cas de prévisions à court terme, la moyenne des membres du NWPE s'approche souvent plus près des observations que la prévision déterministe (McCollor & Stull, 2008). Lorsque les prévisions météorologiques sont effectuées sur plus de 48 heures en contexte de prévisions hydrologiques, Roulin (2007) a observé que l'utilisation des NWPEs permet d'obtenir un critère de compétence de Brier (*Brier Skill Score*) supérieur aux prévisions déterministes afin d'évaluer la probabilité de dépassement d'un certain niveau d'eau dans une rivière. Le critère de compétence de Brier est une métrique permettant de mesurer l'amélioration de la précision d'une prévision (voir formule présentée par Roulin (2007)). Un post-traitement des NWPEs (c.-à-d. une correction des variables météorologiques des NWPEs par rapport aux observations) permet aussi d'améliorer la qualité de la prévision d'ensemble en recentrant l'intervalle de probabilité d'occurrence des prévisions sur les observations (Georgakakos et al., 2014).

En contexte de gestion hydroélectrique, des prévisions déterministes de haute résolution peuvent s'avérer plus précises que les NWPEs afin d'évaluer le potentiel de production hydroélectrique à partir de la quantité d'eau qui sera disponible pour être turbinée (Boucher et al., 2012). En plus d'être disponibles en meilleure résolution, les prévisions météorologiques déterministes demandent aussi moins de temps de calcul que leur homologue probabiliste (Demeritt et al., 2007). C'est pourquoi les prévisions déterministes demeurent plus utilisées en contexte opérationnel que les ensembles prévisionnels. Puisque la prévision météorologique parfaite n'existe pas pour bien prévoir les apports en eau et que des décisions difficiles doivent être prises pour gérer le risque, les NWPEs deviennent toutefois un excellent outil en contexte opérationnel (p. ex. : optimisation de réservoirs hydroélectriques ou prévention des inondations) afin de prévoir les scénarios extrêmes possibles et ainsi mieux reconnaître les incertitudes de la météo (Demeritt et al., 2007; Schaake, Hamill, et al., 2007).

Peu importe le type de prévisions météorologiques choisi, les données fournies par les centres météorologiques présentent inévitablement des biais par rapport aux réelles observations. Certains biais peuvent provenir de la résolution et précision des modèles numériques de surfaces (Mazrooei et al., 2015), des modèles de prévision de l'état de l'atmosphère terrestre qui permettent de déterminer les conditions initiales et des méthodes d'observation de la météo (p. ex : stations, satellites, radars) alors que d'autres proviennent directement de l'incertitude des modèles météorologiques utilisant ces valeurs (Council, 2006; Ramos et al., 2010). Il est toutefois encore ardu de déterminer d'où proviennent ces biais et même de les caractériser (Leutbecher et al., 2017).

1.2 Prévisions hydrologiques

Les prévisions hydrologiques en contexte de gestion hydroélectrique sont réalisées par des modèles hydrologiques effectuant la conversion pluie-débit à l'aide des prévisions météorologiques (typiquement les variables de précipitation et de température). Similairement

aux prévisions météorologiques, une prévision hydrologique peut être déterministe ou probabiliste.

Bien que la prévision déterministe ait encore un temps de calcul inférieur et une plus grande résolution que son homologue probabiliste, elle peut avoir un effet pervers en contexte opérationnel en donnant une impression de certitude puisqu'il y a qu'une seule valeur de prévision (Krzysztofowicz, 2001). Considérant qu'aucune prévision n'est parfaite, cette interprétation pourrait s'avérer d'autant plus dangereuse si elle se trouve à sous-estimer ou surestimer la réalité.

La prévision probabiliste (appelée prévision hydrologique d'ensemble ou de l'anglais ESP; ensemble streamflow forecast), pour sa part, permet d'inclure une partie de l'incertitude (p. ex. : la précision des mesures ou les simplifications du cycle de l'eau effectuées par le modèle hydrologique) et donc de communiquer un certain risque associé à la prévision afin d'aider à la prise de décisions rationnelles afin de réduire les risques économiques (p. ex. : pertes de revenus en contexte de gestion hydroélectrique suite à une surverse de réservoir ou coût supplémentaire lié à la gestion d'inondations) (Montanari & Grossi, 2008). Cette technique permet de donner une appréciation objective du niveau de confiance envers les prévisions et de quantifier l'incertitude pour chaque bassin versant, ce qui aide les gestionnaires de réservoirs à effectuer des investissements ciblés et adéquats à chaque milieu (p. ex. : installation de stations de mesure) dans le but de réduire cette incertitude (Demargne et al., 2014). Finalement, toujours selon Demargne et al. (2014), la prévision hydrologique probabiliste permet d'exploiter les NWPEs qui contiennent elles aussi de l'incertitude (typiquement associée à la précision des conditions initiales). Toutefois, l'incertitude du modèle hydrologique est habituellement prédominante sur l'incertitude météorologique en contexte de prévisions à court terme (Bennett et al., 2014).

En plus des avantages énumérés ci-haut, les ESPs permettent une meilleure représentativité des événements extrêmes. En effet, les événements plus rares peuvent alors faire partie des scénarios probables de la prévision, représentés par les membres supérieurs ou inférieurs de

l'enveloppe. Dans plusieurs cas, la moyenne de tous les membres d'un ESP peut s'avérer plus performante qu'une prévision déterministe (Wetterhall et al., 2013).

Même si l'ESP semble être une meilleure alternative que la prévision déterministe, il est important de considérer également les incertitudes liées à la structure des modèles hydrologiques, qui s'ajoutent à celles des modèles météorologiques présentés au sous-chapitre 1.1. L'incertitude associée au modèle hydrologique provient de la structure du modèle hydrologique lui-même, de ses paramètres, de ses conditions initiales, de sa calibration et des débits observés qui lui servent de référence (Ajami et al., 2008). Les différentes sources d'incertitudes des prévisions hydrologiques probabilistes peuvent être réduites en utilisant conjointement diverses techniques qui ont déjà démontré leur efficacité. Elles incluent notamment l'assimilation de données (Liu et al., 2012), le post-traitement des données hydrométriques (Wetterhall & Smith, 2019) et l'utilisation du multi-modèle (c.-à-d. de plusieurs modèles hydrologiques) (Leandro et al., 2019). Ces techniques seront présentées dans les prochaines sections.

Les conclusions de plusieurs études (Hemri et al., 2013; Arsenault et al., 2016; Boucher et al., 2019) suggèrent qu'il y a deux problématiques récurrentes en lien avec l'utilisation des ESPs : elles sont généralement sous-dispersées et biaisées. La sous-dispersion d'un ESP survient lorsque les membres de l'ensemble ne varient pas assez les uns par rapport aux autres. Dans une telle situation, les observations se situent fréquemment dans le haut ou le bas de l'intervalle de l'enveloppe créé par tous les membres (voir figure 1.1).

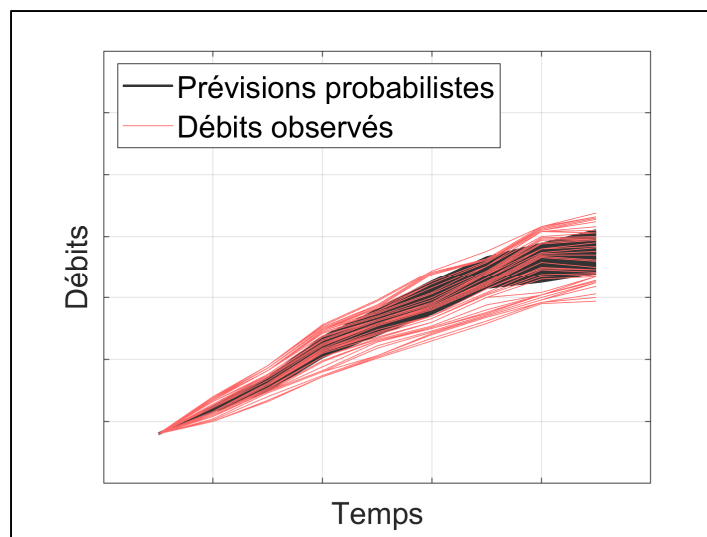


Figure 1.1 Exemple d'un ensemble hydrologique prévisionnel sous-dispersé. La dispersion de l'ensemble n'est pas assez grande pour inclure tous les débits observés

Il est aussi possible de représenter cette problématique sur un diagramme de Talagrand (voir figure 1.2). Ce diagramme permet de vérifier visuellement que tous les intervalles de distribution des ensembles sont fiables. Si ces ensembles sont bien distribués, les observations, qui correspondent aux membres en ordonnée, devraient être réparties uniformément dans les intervalles du diagramme. Cette figure montre aussi les autres types de dispersion d'un ESP possibles. Une surdispersion indique que la distribution de l'ESP est trop large et variable. Elle inclut alors des membres aberrants qui sont beaucoup trop bas ou élevés par rapport aux observations. Les biais systématiques négatifs (voir figure 1.3) et positifs indiquent que l'ESP ne représente pas du tout les observations et que celles-ci sont toujours plus élevées (biais négatif) ou plus basses (biais positif) que les membres de l'ESP. L'uniformité « réaliste » d'un ESP est l'objectif à atteindre, sachant que l'uniformité « théorique » est impossible à atteindre dans la réalité.

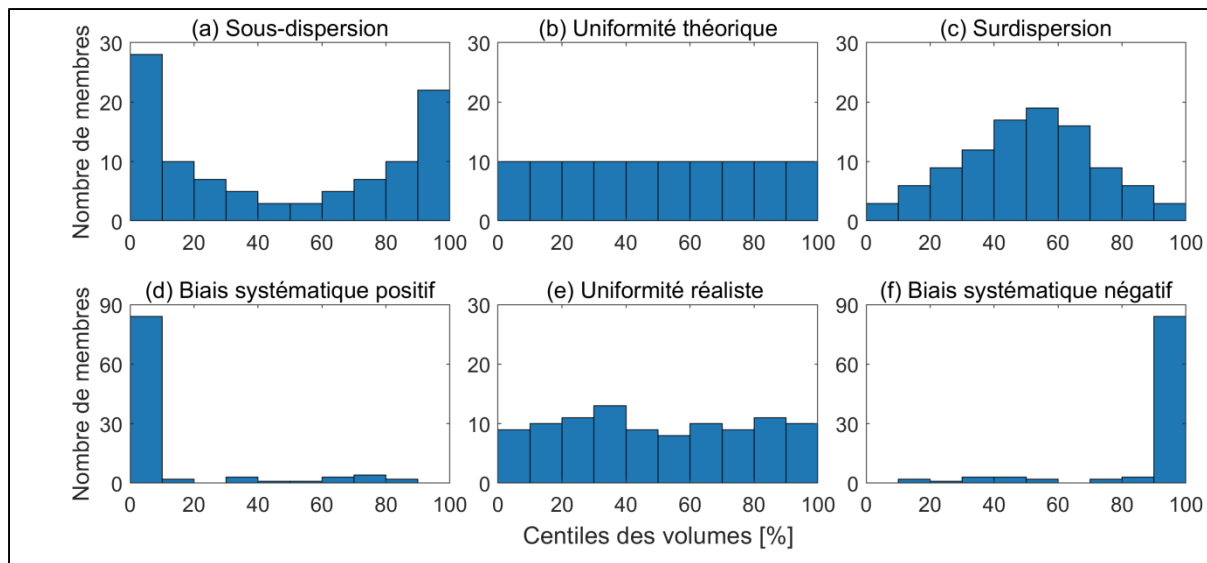


Figure 1.2 Exemples de dispersion d'un ensemble de 100 membres sur des diagrammes de Talagrand. Le « Nombre de membres » en ordonnée du diagramme représente la fréquence des observations, c'est-à-dire le nombre de fois où l'observation se situe dans une des 10 classes en abscisse du diagramme de Talagrand

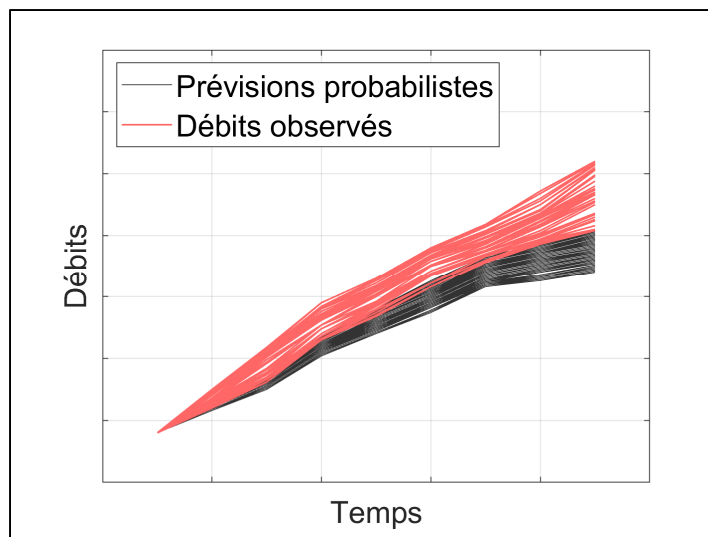


Figure 1.3 Exemple d'un ensemble hydrologique prévisionnel systématiquement biaisé négativement. Les observations ne sont pas incluses dans l'ensemble

L'une des hypothèses avancées par Boucher et al. (2019) et Hemri et al. (2013) pour expliquer ces problèmes est que les modèles hydrologiques utilisent comme variables d'entrée les NWPEs, qui sont aussi typiquement biaisées et sous-dispersées. La sous-dispersion des prévisions implique que les observations sont souvent placées dans l'intervalle supérieur ou inférieur de l'enveloppe des prévisions, ce qui ne représente pas un avantage en contexte opérationnel. Dans un tel contexte, il est important que les observations se situent à l'intérieur de l'enveloppe des prévisions afin que la gestion des réservoirs soit effectuée de manière optimale (p. ex. : volume d'eau turbiné) en se basant sur des prévisions d'apport en eau véridiques. D'ailleurs, plusieurs défis s'opposent à l'utilisation d'un système de l'ESP en contexte opérationnel (Cloke & Pappenberger, 2009; Cuo et al., 2011) :

- les faibles résolutions spatiales et temporelles des prévisions météorologiques ainsi que leurs difficultés à prédire l'avenir avec précision;
- le manque de compréhension et de communication de toutes les incertitudes du système;
- l'assimilation de données (c.-à-d. : la correction d'une simulation hydrologique par rapport aux observations à l'aide de mesures in situ) plus ou moins performante;
- la puissance informatique insuffisante;
- le manque d'expérience pour son implantation dans les opérations concrètes telles que la gestion des barrages hydroélectriques et/ou la prévention des inondations.

Arsenault et al. (2016) ont démontré qu'il était possible de corriger la sous-dispersion d'un ESP durant l'hiver en ajoutant de la variabilité aux conditions initiales du modèle hydrologique conceptuel CEQUEAU. L'ajout de variabilité s'effectuait en insérant un terme d'erreur (calculé préalablement entre les débits simulés et observés en période de référence) à la variable d'état de l'équivalent en eau de la neige du modèle hydrologique.

1.3 Assimilation de données

Une simulation provenant d'un modèle hydrologique calibré présente typiquement un écart au niveau de l'amplitude des débits et/ou un décalage temporel de certains événements (Liu et al.,

2011). Sans post-traitement, la simulation effectuée pour les prévisions à venir pourrait présenter un biais considérable sur les conditions initiales. L'assimilation de données (appelé en anglais *data assimilation*; DA) vise à corriger cette problématique en réalignant quasi parfaitement les conditions initiales d'un modèle hydrologique par rapport aux observations à chaque début de prévisions hydrologiques à l'aide de mesures in situ. Le réaligement permet de débiter une nouvelle simulation hydrologique près des conditions réellement observées sur un bassin versant à un moment précis. Cette correction est donc importante et bénéfique dans le processus visant à produire des prévisions hydrologiques représentatives de la réalité (Seo et al., 2014; Wetterhall & Smith, 2019).

En prévision opérationnelle (p. ex. : pour la gestion des réservoirs), l'assimilation de données (c.-à-d. la correction des conditions initiales des modèles hydrologiques) se fait souvent manuellement, mais ce procédé est laborieux et difficile à appliquer lorsque les ESPs sont utilisées (Seo et al., 2014). Il y a alors un nombre trop important de conditions initiales à corriger manuellement. Différentes techniques d'assimilation de données automatiques existent afin de rendre le processus plus rapide et performant. Celles-ci peuvent avoir comme objectif de réduire le biais au niveau de l'humidité du sol (Reichle et al., 2002), de la quantité de neige (Andreadis & Lettenmaier, 2006; DeChant & Moradkhani, 2011) ou encore du débit simulé (Moradkhani et al., 2005; R. Lanthier, 2018). Toutes ces études ont démontré que l'utilisation d'une technique d'assimilation de données automatique permettait de réduire les biais et produire des prévisions hydrologiques plus représentatives des observations. Différentes méthodes d'assimilation de données existent et sont utilisées dans la littérature, telles que les filtres particuliers (*Particular Filters*; PF) (DeChant & Moradkhani, 2011) et l'approche bayésienne (*Bayesian forecasting system*; BFS) (Reggiani & Weerts, 2008), mais la technique d'ensemble des filtres de Kalman (*Ensemble Kalman Filters*; EnKF) est la méthode la plus utilisée en hydrologie en raison de sa robustesse et de sa facilité d'utilisation (Liu et al., 2012).

Sachant que les observations in situ fournies à un modèle hydrologique (p. ex. : débits de référence et variables météorologiques servant d'intrants aux modèles hydrologiques)

contiennent inévitablement des incertitudes liées aux appareils de mesure, la technique de l'EnKF consiste à échantillonner l'incertitude de ces intrants et à leur ajouter la variabilité correspondante. L'algorithme de l'EnKF crée alors un nouvel ensemble de conditions initiales (c.-à-d. de nouveaux membres permettant de représenter la variabilité ajoutée) qu'il tentera d'aligner sur le débit observé. Elle permet dans un même temps de corriger les distributions asymétriques entre la simulation initiale et les observations et par conséquent, la problématique de non-linéarité entre les conditions initiales du modèle hydrologique et les débits observés (c.-à-d. : le biais entre la simulation et les observations de référence à chaque pas de temps) (Noh et al., 2019). Abaza et al. (2014) ont démontré que l'EnKF permettait aussi d'augmenter la fiabilité de la simulation hydrologique sur toutes les périodes de l'année. De plus, leurs résultats indiquent que même si l'ajout par l'algorithme de l'EnKF d'un ensemble de 1000 membres demeure le plus performant, un choix de 20 à 50 membres aléatoires est généralement suffisant lorsqu'un post-traitement des débits est effectué.

Bien que l'utilisation de l'EnKF soit répandue, il est important de garder à l'esprit que cette technique a certaines lacunes. Dans un premier lieu, l'ajout de membres additionnels requiert évidemment une puissance de calcul plus importante. Ensuite, les conditions initiales générées peuvent être incohérentes et l'amélioration d'un critère de performance suite à l'utilisation de l'EnKF peut se faire au détriment d'un autre (Thibault & Anctil, 2015). De plus, en sous-estimant l'incertitude observationnelle des variables d'entrée et des débits observés (c.-à-d. : de référence), la méthode pourrait ne pas considérer l'incertitude de ces débits, puisqu'elle jugerait alors que ces débits de référence sont « parfaits », alors que ce n'est pas le cas (Maxwell et al., 2018). L'EnKF peut aussi compromettre la fiabilité des prévisions en étant trop confiant au niveau l'exactitude des conditions initiales du modèle hydrologique et ajouter une trop grande variabilité à l'ensemble par la création des nouveaux membres (DeChant & Moradkhani, 2012).

Bien que la méthode de l'EnKF présente certains inconvénients, elle demeure très efficace dans la majorité des études pour réduire le biais entre la simulation et les observations (Bergeron et al., 2016; Maxwell et al., 2018; Jiménez et al., 2019). En effet, elle semble très

performante pour améliorer les prévisions à court terme, notamment sur les bassins versants au Canada (Bergeron et al., 2016) et même sur de longues périodes de prévisions sur des bassins versants en Amérique du Sud (Jiménez et al., 2019). Afin de remédier à ses inconvénients, Thibault et al. (2016) suggère d'employer les NWPEs en plus d'une technique de multi-modèle hydrologique. De plus, il est recommandé d'utiliser une technique de post-traitement des données hydrométriques en supplément d'une technique d'assimilation de données étant donné que les deux sont complémentaires (Bourgin et al., 2014). L'assimilation de données permet ainsi d'obtenir des débits simulés pratiquement identiques aux observations dans les premières journées d'une prévision, alors que le post-traitement permet des prévisions plus précises sur les jours qui suivent.

1.4 Prétraitement et post-traitement

En modélisation hydrologique, le prétraitement des données correspond à la correction des variables météorologiques par rapport aux observations avant de les fournir au modèle hydrologique. Le post-traitement des données, lui, représente plutôt la correction des débits simulés par rapport aux débits observés. L'objectif de ces deux types de traitements reste toutefois similaire : générer des prévisions météorologiques ou hydrologiques les plus fidèles possible aux observations (Hopson et al., 2019).

Le prétraitement des prévisions météorologiques permet de réduire les biais des simulations produites par les modèles météorologiques afin de mieux représenter les précipitations et températures observées (Georgakakos et al., 2014), retirant ainsi les erreurs systématiques. Toutefois, pour être utilisable en modélisation hydrologique, ce prétraitement doit s'effectuer en conservant la cohérence et la corrélation entre les variables qui alimenteront le modèle hydrologique (Wetterhall & Smith, 2019). Ainsi, les variables météorologiques doivent être corrigées en même temps puisqu'un modèle hydrologique nécessite des variables cohérentes et uniformes entre les températures, précipitations et l'évapotranspiration. Cette nécessité est encore plus importante lors de la période de fonte du couvert de neige puisque la température joue un rôle prédominant dans la modélisation des processus liés à la fonte. Il est aussi possible

que le biais systématique des prévisions météorologiques ait été corrigé par prétraitement, mais que cette correction ne soit pas visible à la suite de la simulation hydrologique. Dépendamment des conditions initiales du modèle hydrologique, le prétraitement pourrait donner de bons résultats de simulation hydrologique sur une certaine période de temps, mais moins bons sur une autre (Hopson et al., 2019).

Plusieurs études ont d'ailleurs démontré que le post-traitement est plus performant pour corriger les prévisions hydrologiques que le prétraitement dans la situation où une seule technique doit être utilisée (Arheimer et al., 2011; Wetterhall & Smith, 2019). L'utilisation conjointe des deux techniques pourrait se trouver à être l'option optimale pour certaines situations, comme en hiver lorsque les erreurs dans les conditions initiales du modèle sont plus dommageables que les incertitudes engendrées par la structure des différents modèles hydrologiques (Roulin & Vannitsem, 2015).

Peu importe si le prétraitement, le post-traitement ou les deux sont réalisés, il est difficile de choisir la meilleure méthode de traitement puisque le choix final dépendra assurément de chaque situation et de ses limitations (Muhammad et al., 2018; Boucher et al., 2019). Une revue de la littérature faite par Li et al. (2017) décrit et compare plusieurs méthodes de traitement des données. En complément, Zhao et al. (2017) évoque trois critères pour évaluer une méthode de prétraitement ou de post-traitement :

- 1) corriger les biais;
- 2) conserver la fiabilité de la prévision dans son ensemble;
- 3) les prévisions corrigées doivent être cohérentes (c.-à-d. : elles doivent respecter les processus qui régissent la météorologie ou l'hydrologie).

Parmi les méthodes disponibles, telles que les statistiques de sortie du modèle d'ensemble (*Ensemble model output statistics*; EMOS) (Hemri & Klein, 2017) ou la régression par quantile (*Quantile regression*; QR) (Muthusamy et al., 2016), la correction de biais par quantile (*Quantile mapping*; QM) (Wood & Schaake, 2008) en est une qui a fait ses preuves et qui est

répandue. Cette méthode ne semble toutefois pas appropriée pour le prétraitement puisque son effet est nul sur les prévisions hydrologiques finales selon Verkade et al. (2013). La méthode de correction par quantile QM permet d'ajuster la distribution des débits simulés de l'ESP par rapport aux observations. La méthode est simple, performante pour réduire les biais et efficace sur les faibles débits, comparativement à d'autres méthodes comme celle de la régression par quantile (QR) (Hashino et al., 2007). La principale problématique de cette dernière est le fait qu'elle corrige toutes les données d'une période définie en même temps selon les quantiles de sa distribution, sans conserver la connexion entre la simulation et les observations (Li et al., 2017). Toutefois, lorsqu'elle est utilisée en post-traitement sur les simulations de débits, la méthode de correction par quantile améliore la fiabilité, réduit l'incertitude des résultats et permet d'améliorer la performance des ESPs dans la plupart des cas (Hashino et al., 2007; Wood & Schaake, 2008; Muhammad et al., 2018).

1.5 Multi-modèle hydrologique

Il est possible de combiner les résultats de différents modèles hydrologiques afin d'exploiter les forces de chacun pour compenser certaines faiblesses (Arsenault et al., 2015). Cette technique, connue sous le nom de multi-modèle hydrologique, est de plus en plus utilisée en prévision hydrologique (Devineni et al., 2008; Velázquez et al., 2011; Leandro et al., 2019). La plupart des modèles hydrologiques ont des structures différentes pour un ou plusieurs processus hydrologiques, engendrant par le fait même des forces et des faiblesses. Le terme « multi-modèle » signifie simplement que plusieurs modèles hydrologiques sont utilisés, mais il peut être interprété de différentes façons et la communauté scientifique ne s'entend pas toujours sur son sens. Une des interprétations du terme implique l'utilisation de méthodes multi-modèles connues afin de combiner plusieurs modèles hydrologiques avec des techniques de pondération de ces modèles. Toutefois, il s'agit souvent d'effectuer une moyenne plus ou moins complexe des modèles (p. ex. : moyenne simple (Georgakakos et al., 2004), moyennes de Granger-Ramanathan (Granger & Ramanathan, 1984) ou moyenne du modèle Bayésien (*Bayesian Model Average*; BMA) (Neuman, 2003)).

Dans ce projet de recherche, l'interprétation du terme est différente puisque le but n'est pas d'obtenir une moyenne pondérée des ESPs, mais plutôt de les combiner afin de créer un grand ensemble multi-modèle avec beaucoup plus de variabilité (appelé en anglais Grand ensemble streamflow prediction; GESP). Velázquez et al. (2011) et Brochero et al. (2011) ont tenté l'expérience en utilisant les ESPs sur 9 jours du ECMWF sur une période totale de 507 jours et ont combiné un total de 16 modèles hydrologiques sur 29 bassins versants en France. Leur GESP ainsi créé a obtenu des résultats prometteurs en démontrant que la fiabilité des prévisions augmentait même si leur précision diminuait durant la progression temporelle de la prévision. Hopson et Webster (2010) ont aussi démontré que l'utilisation du GESP regroupant deux modèles hydrologiques sur deux bassins versants au Bangladesh avait permis d'améliorer la fiabilité des ESPs en contexte de prévention des inondations.

Plusieurs études ont ainsi déjà tenté de combiner ces différentes manipulations d'ESP en GESP, mais leur méthode d'assimilation des données était parfois une simple correction (p. ex. : ajouter l'erreur moyenne de la simulation hydrologique de référence par rapport aux observations aux prévisions hydrologiques à chaque pas de temps (Brochero et al., 2011; Velázquez et al., 2011) ou optimiser les états initiaux des modèles hydrologiques à l'aide du critère Nash-Sutcliffe sur les trois jours précédents la prévision (R. Lanthier, 2018)). Par ailleurs, certaines études utilisaient un nombre limité de modèles hydrologiques pour former leur GESP (p. ex. : deux modèles hydrologiques (Hopson & Webster, 2010)), ce qui ne permettait possiblement pas de générer assez de variabilité afin d'uniformiser la distribution des prévisions hydrologiques d'ensemble.

1.6 Objectifs de recherche

Concrètement, l'objectif principal de ce projet de recherche est de développer une méthodologie pour combiner toutes les manipulations (p. ex. : l'assimilation de données, le post-traitement des données et le multi-modèle) qui se sont avérées performantes dans la littérature afin de corriger les biais et la problématique de sous-dispersion des ESPs. L'utilisation des NWPEs comme intrants de huit modèles hydrologiques, lesquels seront tous

soumis à une assimilation de données, permettra d'ajouter la variabilité nécessaire aux ESPs afin de bien représenter les observations. Une technique de post-traitement de ces prévisions permettra à son tour de réduire les biais des prévisions. Enfin, les ESPs des huit modèles hydrologiques seront réunis afin de créer un GESP avec davantage de variabilité. De plus, l'impact des différentes saisons (périodes de régimes hydrologiques) sur les résultats sera analysé. Pour ce faire, différents critères de performance seront utilisés afin d'analyser les résultats.

CHAPITRE 2

HYDROLOGICAL ENSEMBLE FORECASTING USING A MULTI-MODEL FRAMEWORK

Patrice Dion ^a, Richard Arsenault ^b, Jean-Luc Martel ^c

^{a, b, c} Department of Construction Engineering, École de technologie supérieure,
1100 Notre-Dame West, Montreal, Quebec, Canada H3C 1K3

This paper has been submitted for publication
in *Hydrology and Earth System Sciences*,
October 2020

2.1 Abstract

Ensemble streamflow predictions (ESP) from a single hydrological model tend to under-sample the variability needed to provide a good representation of streamflow observations. ESPs typically end up being biased and under-dispersed, making them imprecise and less valuable in an operational context. In this study, a proposed methodology based on a multi-hydrological model approach is applied to address these issues. The forecast methodology includes eight hydrological models whose states were updated through data assimilation using an Ensemble Kalman Filter (EnKF) to provide initial states to the models, allowing them to better sample uncertainty in the hydrometeorological observations and reduce initial condition errors. ESP forecasts are then driven by numerical weather prediction ensemble forecasts of the European Centre for Medium-Range Weather Forecasts (ECMWF) for a nine-day lead-time on five snowmelt-dominated catchments in Quebec, Canada. These eight assimilated individual ESPs are then post-processed using a quantile mapping bias correction method, after which they are combined in a larger ensemble (known as a Grand Ensemble Streamflow Prediction (GESPP)) using a multi-model approach. Results indicate that the proposed methodology allowed reducing both forecast biases and under-dispersion problems for all catchments and four periods corresponding to the seasonality of the streamflow. On their own, the data assimilation (DA) and post-processing manipulations allowed shifting the hydrological ensembles to better encompass the observed streamflow. The multi-model

approach further improved the forecasts by increasing their reliability. Talagrand diagrams, supported by the Kolmogorov-Smirnov test for uniformity, as well as *ABDU* and *CRPSS* scores, were chosen as performance metrics to evaluate the robustness of the methodology. These findings suggest that it is possible to improve hydrological short-term forecast reliability through a multi-model approach in an operational forecasting context.

Keywords: hydrological ensemble forecasting, hydrological modelling, data assimilation, flow post-processing, multi-model, under-dispersion

2.2 Introduction

Streamflow forecasts play a crucial role in the operational contexts of hydropower generation, flood forecasting and water resource management. For hydropower and water resource management, the goal is to exploit the water resource as efficiently as possible. For flood forecasting, the objective is to protect and possibly evacuate affected zones as quickly and as safely as possible. Streamflow forecasts are available in the form of deterministic forecasts or ensemble streamflow forecasts (better known as ensemble streamflow predictions (ESP)). A deterministic forecast provides a single estimation of incoming streamflow, while an ensemble forecast provides a set of several equiprobable streamflow scenarios. The term ESP is used in this study in the context of numerical weather prediction ensemble (NWPE) forecasting and should not be confused with extended streamflow prediction performed using past climate observations.

Although a deterministic streamflow forecast requires less computing time and typically has a higher resolution than an ESP, it can have the perverse effect of giving the impression of certainty by providing a single forecasted value (Krzysztofowicz, 2001). This interpretation can be especially dangerous in an operational context, resulting in an overestimation, or worse, an underestimation of reality. On the other hand, ESP makes it possible to include part of the uncertainty associated with the overall hydrological forecasting processes to assist in rational decision-making, reducing economic risk (Krzysztofowicz, 2001). For instance, including

uncertainty allows for a better representation of extreme events, which can result in decisions that ultimately lead to less economic loss. Thus, ESPs are an important tool for water resource system managers.

The NWPE (Demargne et al., 2014) and the hydrological model (Bennett et al., 2014) are two of the different sources of uncertainty included in ESPs. The latter has been found to be usually predominant in the context of short-term forecasts (Krzysztofowicz & Maranzano, 2004; Bennett et al., 2014). However, several studies (Hemri et al., 2013; Arsenault et al., 2016; Boucher et al., 2019) have reported two recurring problems related to the use of ESPs: the forecasts are generally biased and under-dispersed. When forecasts are biased and under-dispersed, the realization (i.e., observations when working in hindcast) is more likely to fall outside or near the extremities of the forecast envelope. One of the elements that contributes to these scenarios is the predominant uncertainty associated with hydrological models (Hemri et al., 2013). It may be due to the hydrological model structure, the parameters used, the initial states, the system calibration, and even the observed streamflow used during the model calibration (Ajami et al., 2008).

Several data manipulation techniques are available to address these issues specifically. First and foremost, automatic data assimilation (Cloke & Pappenberger, 2009; Cuo et al., 2011) can be used to realign the hydrological model's initial states to reproduce the best estimate of the true observed streamflow. This manipulation is typically done at the beginning of each hydrological forecast using in situ information (i.e., streamflow and snow water equivalent measurements, as well as observed meteorological and evapotranspiration data). Data assimilation of deterministic forecasts is often done manually, but the process is laborious and unsuitable for ESPs (Seo et al., 2014). Studies have shown that automatic data assimilation methods such as the Particle Filters (PF) (DeChant & Moradkhani, 2011) and Ensemble Kalman Filters (EnKF) (Liu et al., 2012) methods, as well as the Bayesian forecasting system (BFS) (Reggiani & Weerts, 2008), are able to improve the streamflow forecast distribution (i.e., a forecast envelope that better encompasses reality).

Performing bias correction (i.e., pre-processing on NWPEs or post-processing on ESPs) is another way to improve the distribution and accuracy of ESPs, since it allows producing forecasts that are statistically inseparable from past observations (Hopson et al., 2019). NWPEs pre-processing reduces the bias in weather simulations by removing systematic errors (Georgakakos et al., 2014). However, in order to be usable in a hydrological modeling framework, this pre-processing must be carried out while maintaining consistency and correlation between the meteorological variables (e.g., precipitation and temperature) that will be fed to the hydrological model (Wetterhall & Smith, 2019). Alternatively, a simpler post-processing of the ESPs can be used to correct the hydrological forecast distribution with the one obtained from the observed streamflow (Muhammad et al., 2018). When only one technique is used, some studies have shown that post-processing is more efficient than pre-processing (Arheimer et al., 2011; Wetterhall & Smith, 2019). Several streamflow post-processing methods have proven to be effective, and are used in different studies, such as those involving ensemble model output statistics (EMOS) (Hemri & Klein, 2017), quantile regression (QR) (Muthusamy et al., 2016) and quantile mapping (QM) (Wood & Schaake, 2008).

The variability of ESPs can also be improved by addressing the uncertainty of hydrological models, which can be different from one model to another. In order to exploit the strengths of several different hydrological models and compensate for certain weaknesses, the multi-hydrological model approach may represent a promising solution (Arsenault et al., 2015). Several algorithms are available to average multiple ESPs. These include the maximum likelihood Bayesian model averaging (MLBMA) (Neuman, 2003) and the ensemble model output statistics (EMOS) (Hemri, 2019). ESPs from various hydrological models can also be combined to create a larger ensemble, called the grand ensemble streamflow prediction (GESP), resulting in a better representation of uncertainty (Brochero et al., 2011; Velázquez et al., 2011). The studies referenced above obtained promising results by showing that the overall reliability of the ensemble forecasts increased.

The main objective of this study is to develop a methodology that allows combining methods that have proven to be successful in the literature in a bid to correct the biases and the under-dispersion issues in ESP forecasts. To this end, ESPs from eight different hydrological models were combined by a multi-model framework, which was then evaluated on four hydrological period regimes over five catchments in Canada. The study area and methodology are presented in Sect. 2.3 and 2.4, respectively, followed by the results in Sect. 2.5, and a discussion in Sect. 2.6. The conclusion and recommendations of this paper are presented in Sect. 2.7.

2.3 Study area

The study area is comprised of five sub-catchments of the 74,200 km² Lac-Saint-Jean catchment located in the province of Quebec, Canada. The climate of this region is characterized by hard winters, with significant snow accumulation from November to April, leading to a snowmelt-driven hydrological regime, with the freshet occurring during the spring. Several reservoir hydropower plants on the catchments are managed by the aluminum division of *Rio Tinto*, which also operates the hydrometric stations used in this study. The snowmelt-dominated hydrological regime makes reservoir management more complicated by the fact that large volumes of water enter the reservoirs during the spring freshet, forcing managers to estimate the optimal amount of water to be discharged depending on their forecasts (Arsenault et al., 2013). Short-term forecasts are nonetheless necessary to estimate the spillway opening timing during the freshet as well as for the rest of the year. This is because the reservoir head plays a major role in the hydropower generating station efficiency, and the reservoir level must be kept as high as possible, while preventing any spills, or worse, overtopping. Therefore, the proposed methodology developed in this study could be used by managers to improve their decision-making process through the use of a GESP on a forecast horizon typically spanning 1-10 days.

The Lac-Saint-Jean catchment is divided into five sub-catchments in the present study (referred to as catchments hereafter) in order to allow considering the position of the retaining structures at their outlet, where known observed streamflow data are available. The five-catchment

sample also allows assessing the impact of the surface area on the techniques analyzed in this study. Table 2.1 presents the hydrological characteristics of each catchment, while Figure 2.1 gives their locations and ground elevations.

Tableau 2.1 Hydrological characteristics of catchments based on observed streamflow from 1953 to 2019 provided by *Rio Tinto*

ID	Catchments	Area [km ²]	Average daily streamflow [m ³ /s]	Average annual peak streamflow [m ³ /s]
LM	Lac-Manouane	4898	119	525
PD	Passes-Dangereuses	11 531	262	1207
CD	Chute-du-Diable	9731	188	960
CS	Chute-à-la-Savane	1297	34	166
LSJ	Lac-Saint-Jean	46 747	873	4183

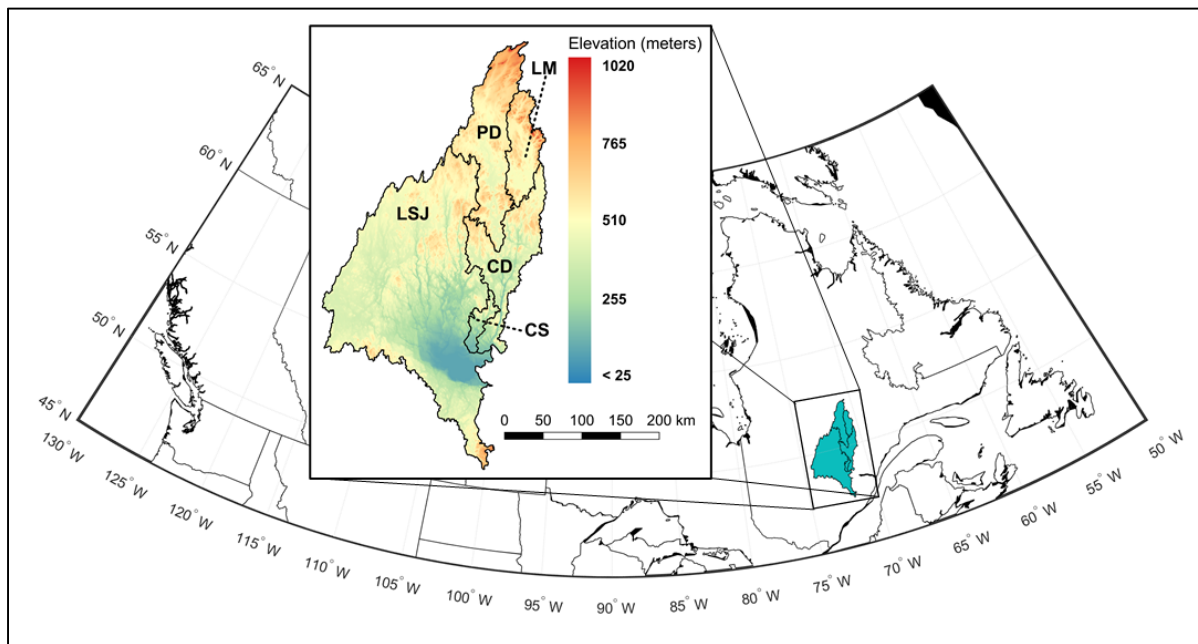


Figure 2.1 Location and topography of the study's five catchments

2.4 Methodology

This section details the different steps of the proposed methodology. The source of meteorological and hydrometric data employed, as well as the techniques used to perform calibration, data assimilation, hydrological modeling, data post-processing and multi-model clustering are discussed. Finally, several performance metrics are introduced to evaluate the performance of the methodology at each step.

2.4.1 Overview

Ensemble forecasting systems (EFSs) include all prediction systems that can be designed in different ways. One of them, the GESP, is formed by feeding an NWPE into more than one hydrological model (Velázquez et al., 2011). A GESP is also formed when two or more NWPEs are fed into at least one hydrological model (Bourdin & Stull, 2013). However, the methodology described in this paper is based on the former GESP type.

To integrate a GESP into the decision-making process, many steps must be performed to transform weather forecasts into representative streamflow forecasts. Figure 2.2 shows the proposed methodology used in this paper. The steps are summarized next and are further described in the following sub-sections.

First and foremost, forecast data from an NWPE must be obtained in order to feed the hydrological models to generate the ESPs. In this study, a total of eight different hydrological models are calibrated using historical climate data. Data assimilation is then performed with these models, using the EnKF method, to generate 25 perturbed initial states. The NWPE in this study provides 50 members, leading to the hydrological modeling of a 1250-member ensemble (50 NWPE members $\times$ 25 EnKF perturbed initial states) for each day of forecast. A post-processing bias correction technique is then performed, followed by a multi-model technique, to combine the resulting ESPs within a GESP. To tackle hydrological period regime variability (referred to as “seasons” hereafter), the last steps are repeated to separate the forecasts into four distinct seasons. In order to compare each step of the methodology to the

same reference, the 50-member NWPE used as input in the hydrological models before performing data assimilation, post-processing and the multi-model GESP is considered as the open-loop (OL) simulation, and serves as a baseline for the analyses.

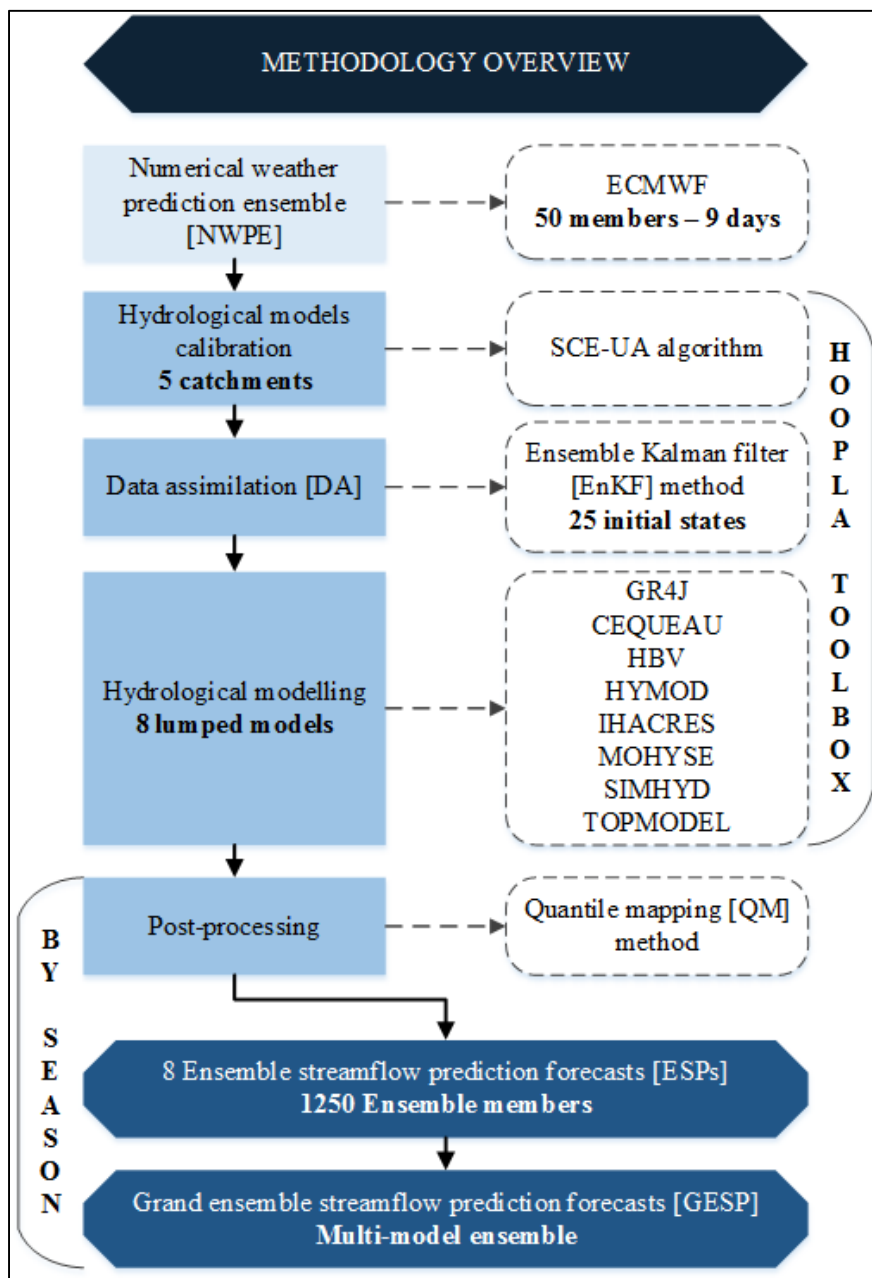


Figure 2.2 Proposed methodology implemented in this study

The HydrOlOgical Prediction LABoratory (HOOPLA) toolbox (Thiboult et al., 2019) is a software application developed at the Department of Civil and Water Engineering of Université Laval (Quebec, Canada). It was created to perform multiple integrated manipulations, such as potential evapotranspiration (PET) and snow accounting routine (SAR) computations, and data assimilation and hydrological modeling on a large set of hydrological models. In this study, HOOPLA is used to perform the hydrological model calibration, data assimilation and simulation steps as shown in Figure 2.2.

2.4.2 Forecast data

The NWPE forecasting data used in this study was extracted from the European Centre for Medium-Range Weather Forecasts (ECMWF) database for the five study catchments. A total of 50 simulations (also known as members) were available at an 18 km resolution and covered the January 2015 to June 2019 time period. The 50-member NWPE was generated for each day, with a forecast lead-time of 9 days. This means that the January 1, 2015 forecast provides information until January 9, 2015. The precipitation and minimum and maximum temperature variables were extracted from the forecast at a 6-hour time step. However, considering that the hydrological models used in this study only perform at the daily time step, the 6-hour time step was converted to a daily one by aggregating the precipitations and taking the maximum and minimum air temperature values over the period.

2.4.3 Hydrological model calibration

The HOOPLA toolbox (Thiboult et al., 2019) was used to perform the calibration and simulations using eight different hydrological conceptual lumped models. Table 2.2 lists all eight hydrological models along with details of their structure, namely, the number of parameters and conceptual reservoirs. The PET was calculated with the Oudin formula (Oudin et al., 2005) and the CemaNeige (Valéry et al., 2014) SAR computation, having previously shown good performance over similar catchments in the same region (Poissant et al., 2017), was selected for all hydrological models.

Tableau 2.2 Number of parameters and conceptual reservoirs for each of the eight hydrological models used in this study
Adapted from Thiboult et al. (2019)

Hydrological model	Number of parameters	Number of reservoirs	Hydrological model derived from
CEQUEAU	9	2	Girard et al. (1972)
GR4J	4	2	Perrin et al. (2003)
HBV	9	3	Bergström et Forsman (1973)
HYMOD	6	5	Wagener et al. (2001)
IHACRES	7	3	Jakeman et al. (1990)
MOHYSE	7	2	Fortin et Turcotte (2007)
SIMHYD	8	3	Chiew et al. (2002)
TOPMODEL	7	3	Beven et al. (1984)

The Shuffled Complex Evolution – University of Arizona (SCE-UA; Duan et al. (1992)) optimization algorithm was chosen for the calibration, as it had previously performed very highly with simple hydrological models similar to the ones used in this study (Arsenault et al., 2014). The Kling–Gupta efficiency (KGE; Gupta et al. (2009)) was selected as the performance metric for the calibration process, and is defined following Eq. (1):

$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2}, \quad (2.1)$$

where r is the correlation coefficient, α is a measure of variability and β is the bias between observed and simulated streamflows. A KGE score of 1 indicates a perfect fit between observations and simulations, and a score greater than -0.41 means that the simulation is better than the average observed streamflow, as shown by Knoben et al. (2019).

The calibration of the eight hydrological models was performed over the 1953-2019 period using hydrometeorological observations provided by *Rio Tinto* for the five catchments. Since the observed streamflows were reconstituted by water mass balance, some pre-processing was required. Negative and zero streamflow rates were not considered in the calibration.

2.4.4 Data assimilation

Data assimilation (DA) aims mainly to modify the hydrological model initial states to better represent the prevailing hydrological conditions on the catchment, allowing simulations to better approximate the true observed streamflow (accounting for uncertainties), and negating the effects of model drift (Seo et al., 2014). In a streamflow forecasting context, these optimal hydrological model initial states, which provide the closest possible simulation of the true observed streamflow at the beginning of the forecast, are used as starting model states.

In the present case, the objective of DA was to reduce the bias between each member of the eight ESP forecasts and observations. To this end, the Ensemble Kalman Filter (EnKF) (Evensen, 1994) was used as the DA method, and was implemented using the HOOPLA toolbox. The EnKF method and its interpretation are explained in detail in Evensen (2003).

One of the challenges of the EnKF method is the need to quantify the uncertainty related to hydrometric and meteorological measurements. Considering that the study catchments' natural inflows are derived from mass-balance estimations, they are expected to be further affected by some level of bias, as compared to real natural streamflow, in addition to other sources of uncertainty, such as precision errors in measurements. With respect to meteorological variables such as precipitation, it is also difficult to obtain perfect representations over the catchment due to the spatial distribution of weather stations, which inevitably leads to errors in real observations.

Given the uncertainties in the reconstitution of inflows (considered to be within 10-15 % in this study) and in meteorological measurements (considered to be between 10-50 % for precipitation; 1-2 °C for temperature in this study), the EnKF makes it possible to sample the uncertainty of these variables and to then create an optimal set of initial states closer to the true observed streamflow (Noh et al., 2019).

In many studies, the EnKF method has proven very effective in reducing the bias between simulations and observations (Bergeron et al., 2016; Maxwell et al., 2018; Jiménez et al., 2019). Abaza et al. (2014) showed that using the EnKF also increased the reliability of hydrological simulations for any period of the year, except for the freshet season, which was not assessed in their study. Moreover, they found that while using an ensemble of 1000 members with the EnKF algorithm provided optimal results, a number between 20 and 50 members was sufficient when forecasts are post-processed. In this study, an ensemble of 25 initial states derived from the EnKF method, combined with the 50-member NWPE, allowed creating an ensemble of 1250 members for each forecast.

2.4.5 Hydrological modeling

Hydrological modeling is performed using the eight hydrological models and their previously calibrated parameters. The 25 daily initial states of each hydrological model, obtained at the DA stage, are used to initialize the forecast. The 50 NWPE forecast members are then provided to the hydrological models to obtain eight individual ESPs. Thus, each ESP, containing 1250 members (25 initial states, combined with the 50 NWPE members), was simulated for each of the nine-day forecast lead-times during the January 2015 to June 2019 period. These ESPs are then post-processed and grouped into a GESP to complete the methodology.

2.4.6 Post-processing

While DA allows reducing the error of the initial states, residual biases and other sources of errors from the various processes are present in final streamflow forecasts. Performing some post-processing to bias-correct simulated streamflow is one of the preferred ways to improve the forecast reliability, bias and accuracy (Wetterhall & Smith, 2019). A quantile mapping (QM) post-processing was thus performed on streamflow forecasts over four individual seasons corresponding to the seasonality of the streamflow on the study catchments: December to March (DJFM; winter low flows), April and May (AM; spring freshet), June to August (JJA; summer low flows and convective storms) and September to November (SON; fall synoptic storms). This post-processing method corrects the streamflow percentiles of the simulated

distributions with those of the observed distributions. Hashino et al. (2007) concluded that the QM method is simple, efficient in reducing bias, and more effective for low flows than other methods such as QR. Other studies (Wood & Schaake, 2008; Muhammad et al., 2018) demonstrated that the QM improves the reliability and the performance of ESP forecasts in most cases when used for post-processing streamflow simulations.

In order to perform bias correction on ESPs, the QM method must be performed after the DA is completed. Otherwise, the QM method would attempt unsuccessfully to correct the spread of wider distributions obtained from the biased initial states of the hydrological models. These inefficiently corrected ESPs would not be able to achieve the objectives of the proposed methodology. For this reason, the QM method is not independently compared to the DA method.

To ensure the robustness of the bias correction, a split-sample validation using 80 % of the data selected at regular intervals to calculate the correction coefficients to be applied to all the individual ESP distributions of a given season was selected. The remaining 20 % of data was used to validate the robustness of the correction. All post-processed results shown hereafter pertain uniquely to the validation dataset.

2.4.7 Hydrological multi-model approach

A simple multi-model approach was preferred in order to evaluate the compromise between simplicity and performance needed in an operational context. The selected method consisted in grouping the eight ESPs together, effectively generating a GESP, to better sample the overall uncertainty. The ESP grouping was tested before and after the post-processing step to better assess the best option.

2.4.8 Performance metrics

Cloke et Pappenberger (2008) suggest using several performance metrics to ensure the reliability of forecasts obtained. Thus, different performance metrics are used to compare the

results of each additional step (DA, QM Post-Processing (henceforth referred to simply as “QM”) and GESP) with those of the OL simulations to confirm their added value. Hereafter, the term “manipulation” will be used to refer to the DA, QM and GESP steps. It is important to note that the QM manipulation includes DA and the GESP manipulation includes both DA and QM.

Anctil et Ramos (2019) reviewed several verification metrics, including the Talagrand diagram, the continuous ranked probability score (*CRPS*) and the continuous ranked probability skill score (*CRPSS*). Talagrand diagrams are commonly used to illustrate the ensemble dispersion versus observed data in the context of hydrological forecasting (Olsson & Lindström, 2008; Voisin et al., 2011; Fan et al., 2014). Although well-suited to assess the reliability of ESPs (Zappa et al., 2013), their visual aspect may be misinterpreted, and it is preferable to combine them with a complementary statistical test for validation (Hamill, 2001). The shape of the Talagrand diagram is essential in order to assess the dispersion of an ESP. A U-shape diagram indicates that the ESP is under-dispersed and is not wide enough to include observations. Observations are then found at the extreme limits of the ESP, either in the first or last percentiles, hence the U-shape. Inversely, an $\cap$ -shape diagram points to an ESP that is over-dispersed, and therefore, too wide. A $\lfloor$ or $\rfloor$ shape indicates a positive or a negative systematic bias, respectively. These shapes, which are completely outside the left or right side of the diagram, are obtained when the ESP does not represent the observations at all. Finally, a flat diagram is the desired shape, indicating a uniform distribution of observations in the ESP, and providing the right amount of variability without any biases.

The Kolmogorov-Smirnov test (K-S test) was chosen to reject (or not) the hypothesis of uniformity ($\alpha = 5\%$). If the hypothesis of uniformity is not rejected, it means that the ESP forecast is likely to provide the right amount of variability (represented by a flat Talagrand diagram), which allows a better evaluation of the overall uncertainty. The Average Bin Distance to Uniformity (*ABDU*) score was also added to quantitatively measure the deviation from uniformity, previously used by Arsenault et al. (2016) on the LSJ catchment and is defined following Eq. (2):

$$ABDU = \frac{1}{N} \sum_{k=1}^N \left| s_k - \frac{M}{N} \right|, \quad (2.2)$$

where s_k is the number of occurrences in bin k , M is the number of y values, and N is the number of histogram bins. An *ABDU* score is equal to zero when there is no deviation from uniformity. In this study, this metric cannot compare seasons with each other since it is directly affected by the total number of values (N) forming the Talagrand diagram and the lengths of the seasons used are not equal. Instead, it will serve as a quantitative metric to compare the distribution between two ensembles and determine which is closer to uniformity over a given season.

Finally, the *CRPS* and *CRPSS* are calculated to evaluate the overall accuracy of the forecasts. These two metrics are commonly used in hydrological forecasting (Schaake, Demargne, et al., 2007; Pagano et al., 2013; Muhammad et al., 2018). The *CRPS* is defined following Eq. (3):

$$CRPS(k) = \int_{-\infty}^{\infty} [F'_k(x) - F_k^o]^2 dx, \quad (2.3)$$

where $F'_k(x)$ is the predictive cumulative distribution function between a given ESP forecast and the observed streamflow (k), x is the predicted variable, and F_k^o is the corresponding observed cumulative distribution function (Anctil & Ramos, 2019). The *CRPS* score will not be analyzed directly, and is only used to calculate the *CRPSS* score, which allows a direct comparison of the OL simulation with other simulations obtained after DA, QM and GESP manipulations. The *CRPSS* is defined following Eq. (4):

$$CRPSS = 1 - \frac{CRPS_b(k)}{CRPS_{ref}(k)}, \quad (2.4)$$

where $CRPS_b$ is the score calculated from the modified forecast and $CRPS_{ref}$ is the score of the OL simulation. A *CRPSS* score greater than zero indicates an improvement in the forecast ensemble accuracy, as compared to the OL ensemble. The *CRPSS* values range varies from $-\infty$ to 1, with 1 being the best possible score and a score below zero indicates a decrease in

performance. Thus, results under -1 will be ignored and simply considered as pointing to very poor performance.

A time delay must be kept between the forecasts used to calculate these verification metrics to ensure their temporal independence (Hamill, 2001). A three-day delay was kept between each data point of the metrics, as suggested by Pappenberger et al. (2011).

2.5 Results

This section details the results of the different manipulations of the proposed methodology. Different topics are investigated, such as the hydrological model calibrations, the respective performances of the DA and QM methods, a comparison between ESP and GESP performance, the impacts of the catchments area, the parameterization of the DA and an evaluation of the short-term forecast performance.

2.5.1 Calibration of the hydrological models

The KGE scores obtained following the calibration of the hydrological models on the five catchments are presented in Table 2.3. The highest KGE scores are obtained on the largest catchment (LSJ) and the lowest KGE scores (by a relatively small margin) are seen on the smallest catchments (CS). Overall, all the hydrological models evaluated show good skills over the study catchments. However, some models, such as MOHYSE, SIMHYD and TOPMODEL, exhibited poorer performances on smaller catchments (CS and LM), while other models, such as GR4J and HBV, performed very well on them. When hydrological models are used on the larger catchments (CD, PD and LSJ), their KGE scores are similar.

Of the eight hydrological models tested, the GR4J provided the best performance or was among the best in terms of KGE scores after calibration, for the five catchments. For the sake of simplicity, the results shown hereafter are focused on the GR4J hydrological model and the

LSJ catchment. The conclusions obtained with GR4J and LSJ are consistent with those obtained for the other hydrological models and catchments.

Tableau 2.3 KGE scores for the eight hydrological models calibrated on the five study catchments

Hydrological models	Catchments				
	LM	PD	CD	CS	LSJ
CEQUEAU	0.86	0.89	0.91	0.82	0.93
GR4J	0.88	0.90	0.92	0.86	0.94
HBV	0.88	0.86	0.92	0.84	0.94
HYMOD	0.84	0.88	0.91	0.80	0.94
IHACRES	0.86	0.87	0.90	0.82	0.92
MOHYSE	0.81	0.86	0.92	0.82	0.94
SIMHYD	0.79	0.86	0.92	0.77	0.92
TOPMODEL	0.82	0.88	0.91	0.76	0.93

2.5.2 Impact of data assimilation and bias correction on individual ESP performance

The Talagrand diagrams for the OL simulations (Figure 2.3) indicate that the observed streamflow is almost always outside of the forecast ensemble envelope for all four seasons (DJFM, AM, JJA and SON). While systematic biases are obtained for DJFM (negative), JJA and SON (positive), AM exhibits a U-shape pattern, indicating under-dispersion of the ensemble. Other studies (Hemri et al., 2013; Arsenault et al., 2016; Boucher et al., 2019) have reported similar issues with the OL simulations with other hydrological models for different periods of the year. These observations prove the relevance of using the DA and post-processing methods on streamflows. Figure 2.3 also indicates that performing DA with the EnKF algorithm provides additional variability to the ensemble for each season, allowing the forecast to better encompass the observed streamflow, and leading to a decrease in the *ABDU* score. Finally, when data assimilation (EnKF) is combined with bias correction (QM), the additional variability obtained further increases the dispersion of the ESPs and reduces the *ABDU* score over all seasons except for AM. The K-S test rejected the hypothesis of uniformity ($\alpha = 0.05$) for all ensembles, except for the (f), (h), (k) and (l) ensembles in Figure 2.3.

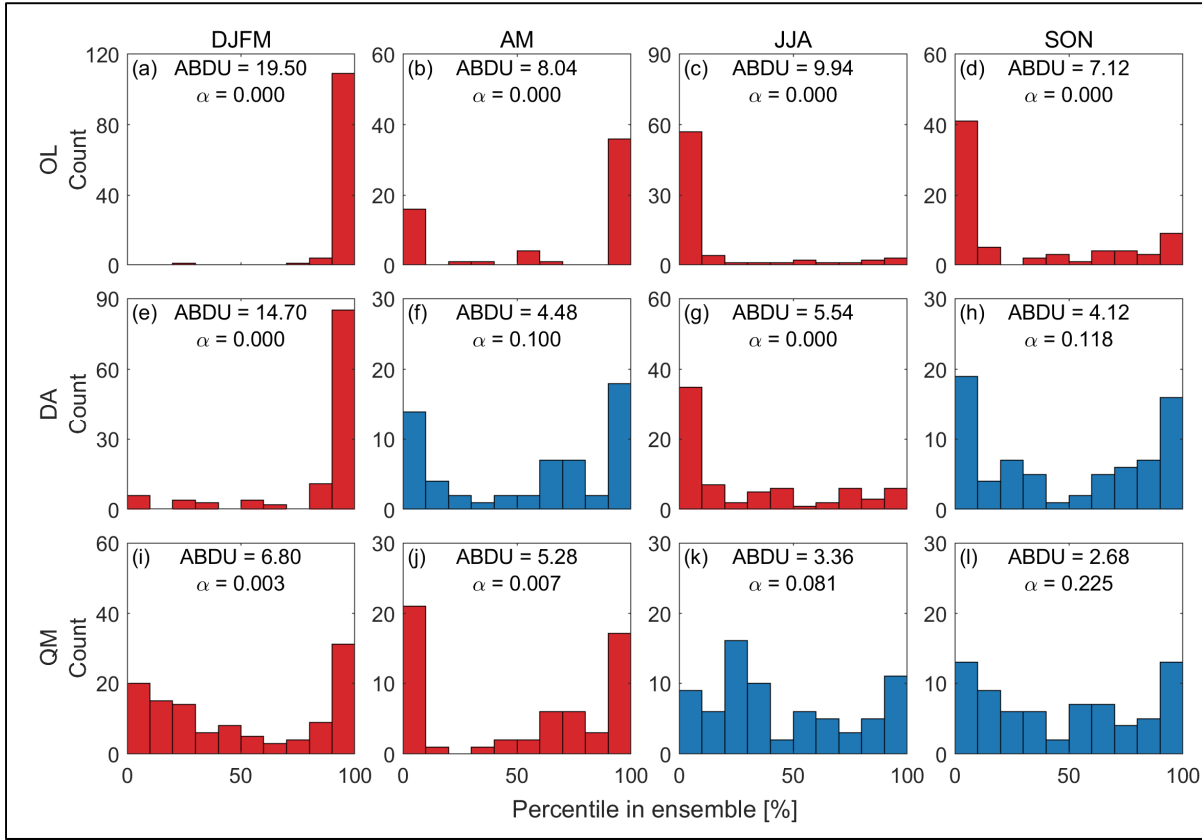


Figure 2.3 Talagrand diagrams representing rank (in percentiles) occupied by observed streamflows in the hydrological forecast ensemble using the GR4J model on the LSJ catchment. “OL” refers to the open-loop simulation. “DA” refers to the validation members of the 1250 ensemble members after performing data assimilation and “QM” refers to the same members after performing data assimilation and post-processing with quantile-mapping. Blue (red) diagrams indicate that the K-S test failed to reject (rejected) the hypothesis of uniformity ($\alpha = 0.05$)

2.5.3 Comparison between individual ESPs and the GESP performances

The objective of the multi-model (GESP) manipulation was to add enough variability to the individual ESPs (after DA and QM) by grouping them together into a GESP. The same manipulations were performed with the other seven hydrological models, as displayed in Figure 2.4. The results of the GESP multi-model set are also presented in Figure 2.4 for ease of comparison. Even after performing DA and QM, all individual ESPs remain under-dispersed for at least one or two seasons. All ESPs created by the eight individual hydrological models

in the AM season (corresponding to the spring freshet) present U shapes, indicating under-dispersion relative to streamflow observations. Figure 2.4 also shows that the K-S test rejected the hypothesis of uniformity for the majority (14/16) of the ensembles during the DJMF and AM seasons, but failed to reject the same hypothesis for more than half (11/16) of the ensembles during the other two seasons (JJA and SON).

Overall, no single ESP (or hydrological model) outperformed all other ESPs, and their performance varied strongly from season to season, as well as from catchment to catchment. This also indicates that each ESP is likely to contribute to additional variability to the forecast in a GESP, allowing to better encompass variabilities in the observed streamflow. This can best be observed in the last row of Figure 2.4, which shows the results with the GESP. While the *ABDU* values from different single ESPs still outperform the GESP in some specific cases (CEQUEAU and IHACRES for DJFM; HYMOD and TOPMODEL for JJA), the GESP provided the lowest *ABDU* values for the AM and SON seasons (with the largest improvements seen over the AM season). Furthermore, the K-S test failed to reject the hypothesis of uniformity over all seasons, most notably over the AM season, for which the K-S test rejected the hypothesis for every single ESP.

It should be noted that very similar results were obtained over the other catchments. The main difference is that the K-S test rejected the hypothesis of uniformity for the GESP over three seasons over the CD and CS catchments (as shown in Figure 2.5). However, the GESP still outperforms almost all individual ESPs in terms of the *ABDU* score.

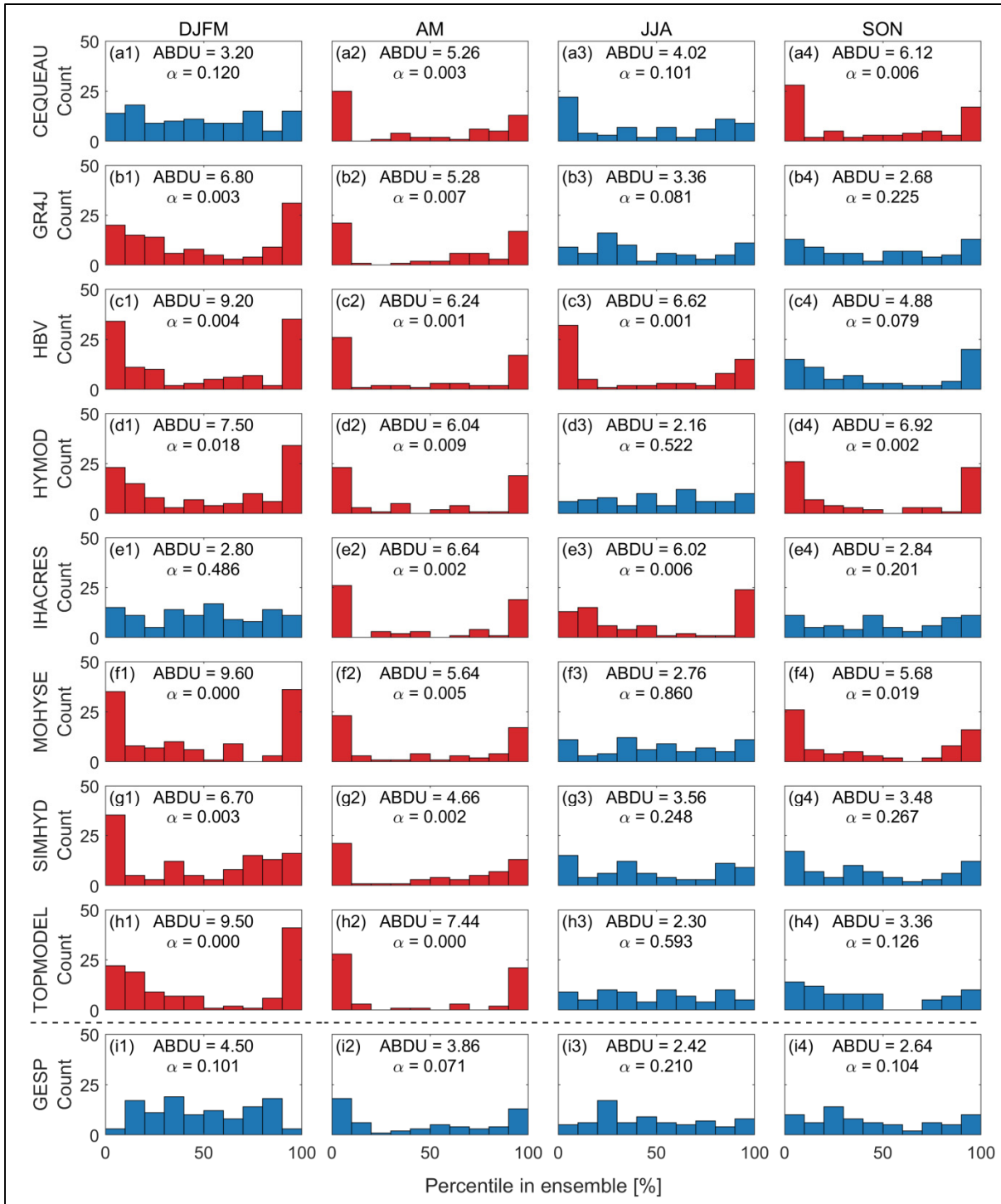


Figure 2.4 Talagrand diagrams representing rank (in percentiles) occupied by observed streamflows in the ESPs of eight different hydrological models after performing DA and QM on the LSJ catchment. Blue (red) diagrams indicate that the K-S test failed to reject (rejected) the hypothesis of uniformity ($\alpha = 0.05$)

2.5.4 Impact of the catchment surface area

As discussed in the previous section, when comparing individual ESPs with each other, no single ESP (or hydrological model) outperformed all other models. However, results suggest that this behavior is also dependent on the catchment surface area, since the models that performed better than the others over a given season vary from catchment to catchment.

Figure 2.5 presents the GESP forecast distribution for all five catchments. In most cases, a more uniform distribution of the forecast members (shown with flatter Talagrand diagrams) indicates that the use of the GESP has brought the ensemble closer to uniformity, compared to individual ESPs (Figure 2.4 for LSJ catchment, Figure 3.3 for PD catchment and Figures-A I-1 to I-3 for the other three catchments). Indeed, the K-S test failed to reject the hypothesis of uniformity for all ensembles, except for three of them: (j), (l) and (n). It can be seen that these three cases occurred on two of the smallest catchments (CD and CS), suggesting that the catchment size could have an impact on potential forecast improvements. Furthermore, some ensembles remain under-dispersed (e.g., (g), (j), (n) and (r)) or even over-dispersed (e.g., (q)).

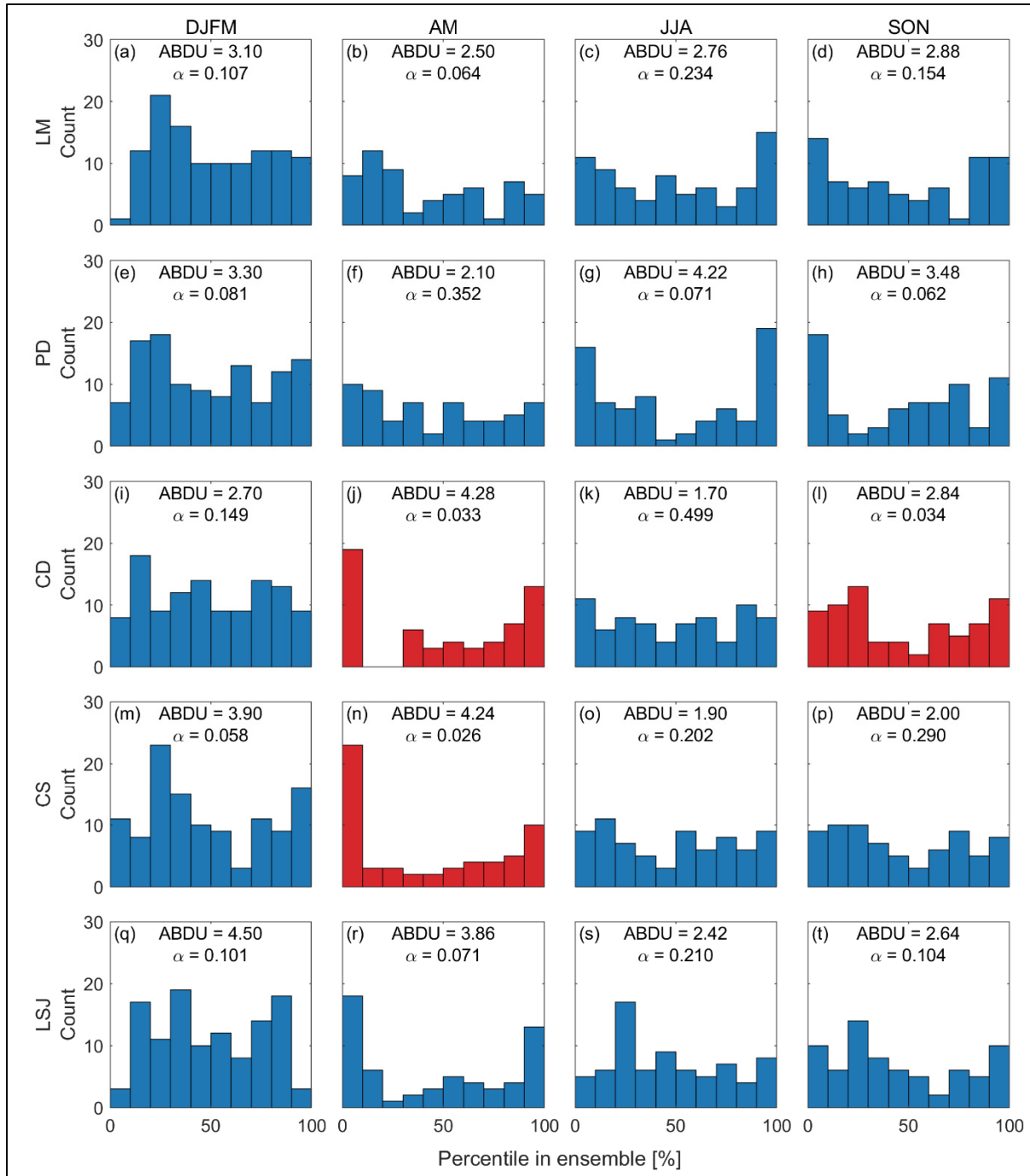


Figure 2.5 Talagrand diagrams representing rank (in percentiles) occupied by observed streamflows in the hydrological multi-model (GESP) forecast after performing DA and QM on the five catchments. Blue (red) diagram indicates that the K-S test failed to reject (rejected) the hypothesis of uniformity ($\alpha = 0.05$)

2.5.5 Impact of data assimilation parameterization

Different data assimilation parameters were tested to add different levels of uncertainty to the hydrological model calibration data (i.e., observed precipitation, temperature and streamflow). The standard deviation of the uncertainty distribution for each observation was set between 10 % and 50 % of precipitation, 1 and 2 °C for temperature and 10 % and 15 % for streamflow. It should be noted that the previous results were obtained with a standard deviation equal to 15 %, 30 % and 2 °C of the streamflow, precipitation and temperature observations, respectively.

Figure 2.6 presents the impact of different DA parameters on the GESP using Talagrand diagrams. The results indicate that changes in standard deviation of both streamflow and precipitation have only a marginal impact on ensemble dispersion, and no clear pattern can be observed in terms of *ABDU* score. Estimated temperature uncertainty ranged from 1 to 2 °C and provided very marginal impact (not shown).

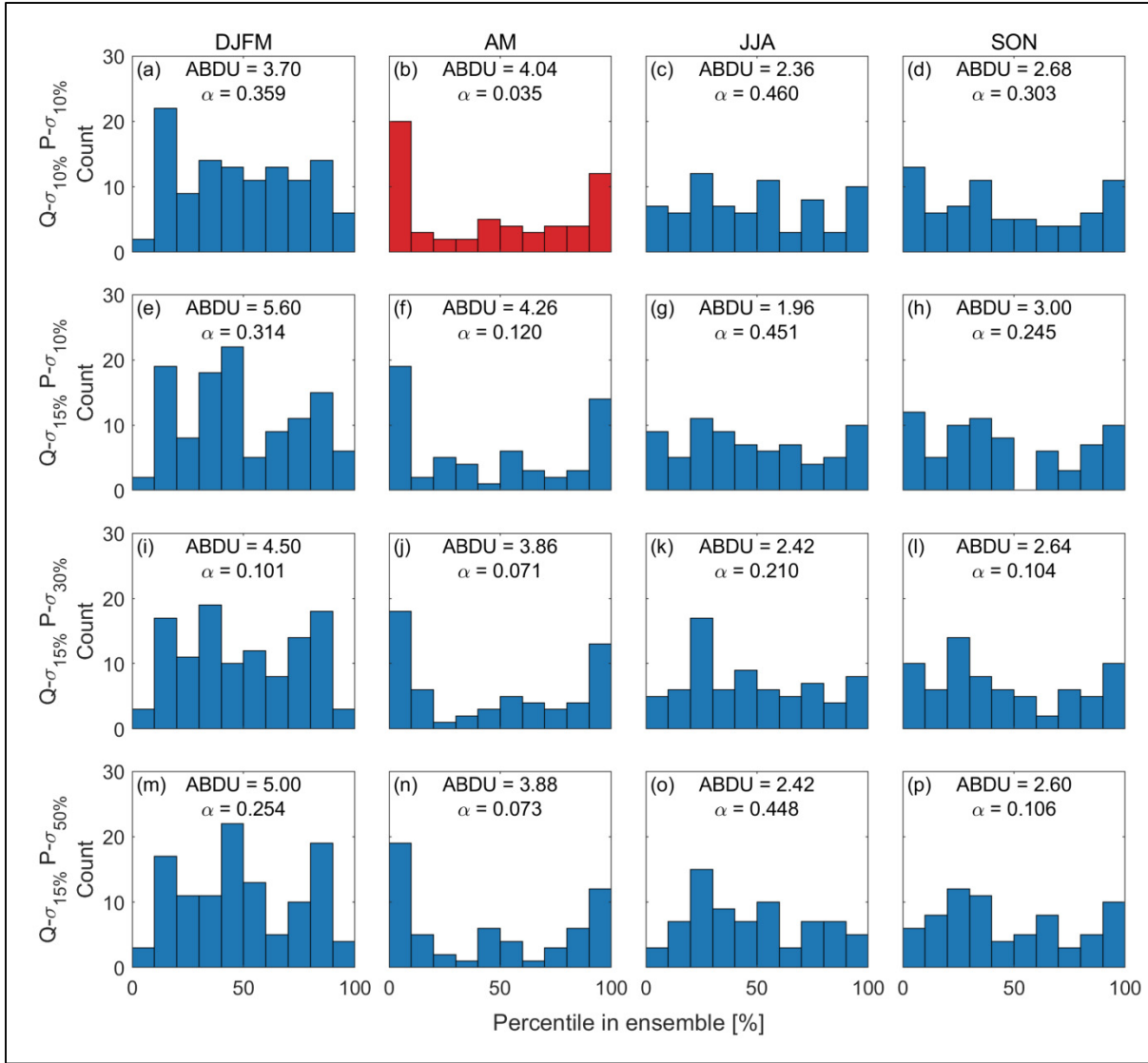


Figure 2.6 Talagrand diagrams representing rank (in percentiles) occupied by observed streamflows in the hydrological multi-model (GESP) forecasts on the LSJ catchment after performing the DA manipulation only, with different standard deviation values for observed streamflows (Q) and the precipitation forecasts (P). Blue (red) diagrams indicate that the K-S test failed to reject (rejected) the hypothesis of uniformity ($\alpha = 0.05$)

2.5.6 Evaluation of the short-term forecast performance

In order to compare the added value of each methodological manipulation (DA, post-processing with QM and combination of ESPs with the GESP) to the forecast accuracy, their *CRPSS* value was computed against OL simulations. Furthermore, the *CRPSS* values were also

computed over the nine forecasting days for each manipulation to assess their skill in the short term. These results are presented with Figure 2.7 boxplots for the LSJ catchment.

It can be seen that the median *CRPSS* value of each boxplot is above 0 for all cases, indicating that the short-term forecast accuracy is improved with each data manipulation performed. As expected, the forecast accuracy is gradually decreasing over time. In most cases, bias-correction and the use of the GESP improve the *CRPSS* score, as compared to using DA only. However, combining all ESPs into a GESP does not always increase the accuracy of the forecast when compared with the GR4J bias-corrected simulations.

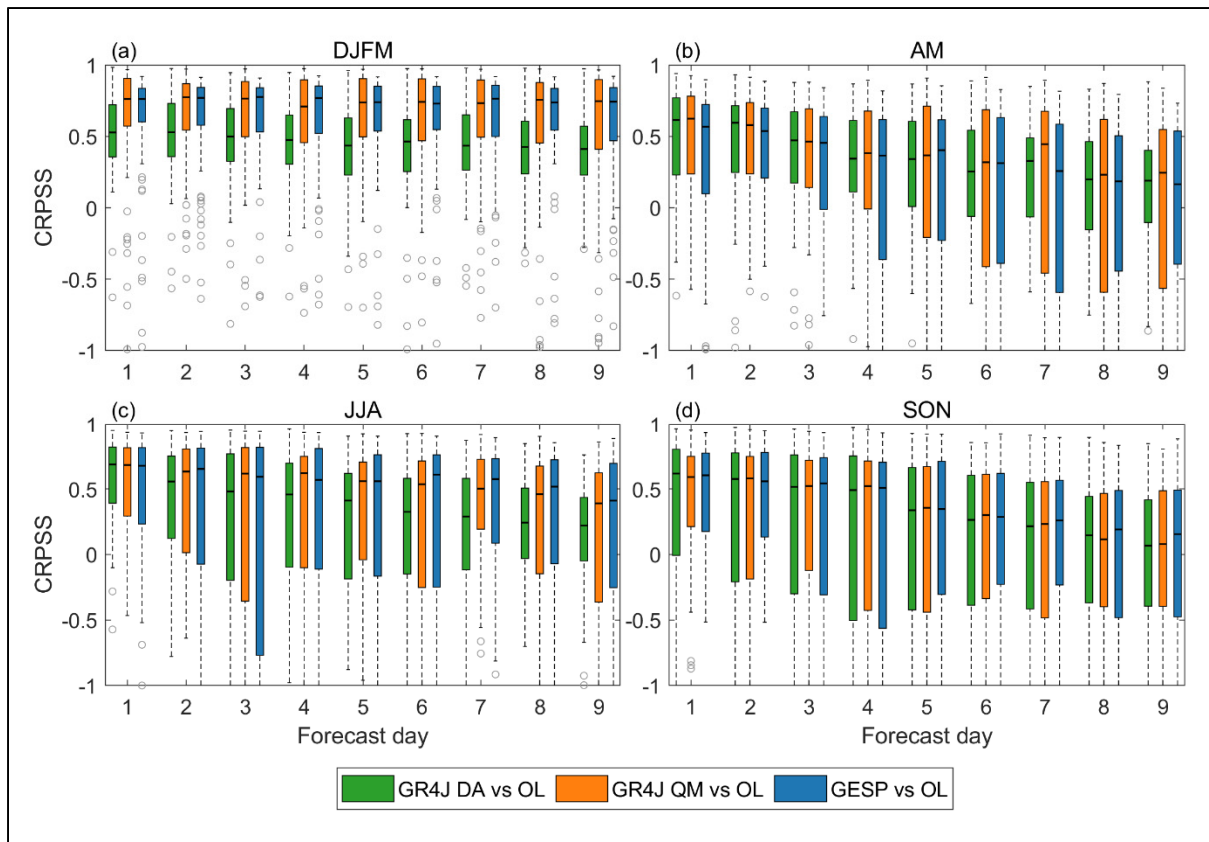


Figure 2.7 Boxplot of the *CRPSS* values of the members of the GR4J hydrological model ensemble and multi-model (GESP) ensemble on the LSJ catchment. *CRPSS* greater than zero indicates an improvement in overall forecast accuracy, as compared to the GR4J open-loop (OL) simulations. Members with *CRPSS* values smaller than -1 are considered very poor and are purposely left out to avoid overloading the graph

2.5.7 Impact of the hydrological season

By analyzing the different performance metrics after each manipulation (DA, QM and GESP), it is possible to confirm that the results depend on the forecast season. Figure 2.3 shows that the DA reduces *ABDU* scores over all seasons, while the QM does not seem to perform as well in reducing the *ABDU* score during the freshet season (AM). When the hydrological ensembles are grouped in a GESP (Figures 2.4 and 2.5), the distribution of this larger ensemble is similarly improved during the DJFM, JJA and SON seasons. The distribution of GESP ensembles during the AM season still shows under-dispersion over most of the catchments (3/5).

The *CRPSS* score results (Figure 2.7) also allow to observe the impact of the different hydrological seasons. While the three methods used (DA, QM, GESP) lead to a significant improvement of the *CRPSS* during the DJFM season, they seem to have a lesser impact on the results of the AM season. On the LM and CS catchments (e.g., Figures-A II-2 and II-4 for the CS catchment with the GR4J and IHACRES models, respectively), *CRPSS* scores even drop below 0 on the last forecast day, and are close to 0 by the fifth day. This indicates that the proposed methodology is unable to improve ensemble accuracy over the two smallest catchments during the AM season beyond a few days. For the other two seasons, JJA and SON, the *CRPSS* scores are improved to a lesser extent than those of the DJFM season, but still remain above 0 for all nine forecasted days, indicating an improvement in the accuracy over the OL simulation.

2.6 Discussion

This section assesses the performance of different manipulations performed on ESP forecasts in order to meet the main objective of this study, which was to develop a methodology to correct the problems of bias and under-dispersion in forecast ensembles. The effectiveness of each streamflow manipulation method (DA, QM, and GESP) is discussed first, followed by a discussion of the impact of the hydrological season, the performance of the overall methodology and their limitations.

2.6.1 Performance of data assimilation using EnKF

The use of EnKF at the data assimilation stage is intended to reduce the bias between the members of the ESP forecasts and the observations. This is achieved by sampling the uncertainties in hydrometeorological (i.e., streamflow, precipitation and temperature) measurements to correct the hydrological models' initial states.

As shown in Figure 2.3, the EnKF method is able to reduce the *ABDU* scores between ESPs and observations for all seasons over the LSJ catchment. Overall, there is an improvement in the spread of the ESPs, as exhibited in the Talagrand diagrams, for all seasons. While these improvements were significant for both the AM and SON seasons, resulting in a failure to reject the hypothesis of uniformity based on the K-S test, the under-dispersion (displayed as a U-shape in the Talagrand diagram) remains problematic. In the case of the other seasons (DJFM and JJA), Figure 2.3 suggests that DA has been unable to correct the systematic biases (both positive and negative) in the ESP forecasts. This could be due to the fact that the method is effective in reducing errors in the initial conditions that improve simulated streamflow, but does not allow full consideration of the uncertainty of the observations, and may compromise the reliability of the predictions by adding random uncertainty, as Maxwell et al. (2018) had warned. There is also no guarantee that the post-assimilation initial states would generate a streamflow that is more dispersed.

It can be seen in Figure 2.6 that there is no optimal value in the hydrometeorological estimation provided to the EnKF method, and that this estimate is strongly dependent on the season and the catchment analyzed (not shown). Thibault et Anctil (2015) also observed that both the catchment and the hydrological model had an impact of the EnKF performance. However, what is most notable in the results in Figure 2.6 is that no matter the value used as an estimate of the uncertainty in the hydrometeorological measurements, data assimilation always improves the forecast performance when compared to the OL simulation. These findings are

also supported by Abaza et al. (2014), who used an ensemble of 1000 members (different initial states using the EnKF method) over a forecast period of 10 days.

As shown in Figure 2.7, the gradual decrease in the forecast performance as the lead-time increases with a gradual decrease in the *CRPSS* median value indicates that the DA tends to be more efficient in the shorter term (up to 2-3 days). Bergeron et al. (2016) obtained similar conclusions, where data assimilation of streamflow data using EnKF was notably more efficient over only a few days of hydrological forecasts. However, the *CRPSS* median value remained above 0 for the nine forecast days, indicating an improvement brought about by data assimilation over the baseline OL simulations.

2.6.2 Post-processing performance with QM bias correction

Figure 2.3 shows that for the LSJ catchment and the GR4J hydrological model, the post-processing of ESP forecasts using the QM method allows solving some of the under-dispersion over all seasons, with the exception of the freshet season (AM). This is also supported by a decrease in the *ABDU* score and a failure to reject the hypothesis of uniformity with the K-S test for both the JJA and SON seasons. As noted by Li et al. (2017), the decrease in performance obtained during the freshet (versus the DA) can be explained by the fact that the correction does not preserve the temporal connection in the snowmelt process. It should be mentioned that while post-processing with the QM is expected to improve the bias (evaluated with the *CRPSS*), it is not designed to improve the forecast dispersion (evaluated with the *ABDU*). Overall, the results shown in Figure 2.4 for all eight individual ESPs indicate that the QM post-processing is not able to correct the under-dispersion problem.

Figure 2.7 shows the boxplots of the daily *CRPSS* obtained by the forecasts over the LSJ catchment, simulated by the GR4J model, over a nine-day lead-time. The median *CRPSS* of the boxplots computed from the post-processed (QM) forecasts are always above zero, showing better performance than for the OL simulations, and are almost always higher than the median values obtained for the DA. Similar results were obtained for the other hydrological

models and catchments (e.g., the IHACRES model (Figures 3.2 and A II-3) on the LSJ and LM catchments, respectively) with the exception of the smaller CS catchment, for which the median CRPSS values for DA are often higher than those for QM depending on the season (e.g., Figures-A II-2 and II-4 with the GR4J and IHACRES models, respectively). This suggests that for most cases, the post-processing of ESPs further improves the forecast accuracy versus both the OL and DA ensembles.

Overall, the results show that post-processing using the QM bias correction technique allows reducing the biases in the ESPs in most cases, but shows little skill when it comes to improving the under-dispersion problem. This finding is also supported by Hashino et al. (2007), who showed that QM was effective in reducing the biases of monthly streamflows, while the present study investigated daily streamflows.

2.6.3 Multi-model performance analysis

The final methodological manipulation brought together the different ESP forecasts of the eight hydrological models into a single GESP. The goal here was to aggregate the strengths and weaknesses from each individual hydrological model, and to then obtain a larger ensemble that best encompasses the overall uncertainty. It can be seen that this additional manipulation is indeed capable of reducing the *ABDU* score as well as the under-dispersion problem (shown with U-shaped Talagrand diagrams) for all seasons, as compared to each individual ESP (Figure 2.4). By combining the strengths and the weaknesses of the different ESPs, the GESP significantly improved the forecast distribution and failed to reject the hypothesis of uniformity for all seasons. The largest improvement was seen over the AM season, in which all individual ESPs were strongly under-dispersed (Figure 2.4). These results were found to be consistent from one catchment to the next, as shown in Figure 2.5.

When comparing the *CRPSS* obtained from the GESP to the ESPs (Figure 2.7; GR4J hydrological model over the LSJ catchment), results indicate that the forecast accuracy of GESP is strongly improved when compared to the OL simulation, but does not improve beyond

what is achieved with the QM post-processing. This was fairly consistent over the nine-day forecast lead-time. While the GESP is not able to outperform the QM in this case, analyses of the other seven individual hydrological models (e.g., Figures 3.2, A II-3 and A II-4 for the IHACRES model) indicate that the GESP forecast accuracy is usually greater than or equal to the one from the post-processing. Furthermore, on the last day of the forecast, the median *CRPSS* of the GESP is almost always greater than or equal to the ones from individual ESPs after DA and QM. Overall, these results indicate that there is an advantage in grouping together ESPs from different hydrological models into a GESP, especially as the lead-time increases. These findings are supported by Velázquez et al. (2011), who conducted a similar analysis using 16 hydrological models. Their study unsurprisingly found that the accuracy of GESP forecasts decreases as the lead-time increases, as also shown in Figure 2.7.

The present study was carried out by grouping together the forecast ensembles of eight hydrological models, but the results obtained could be altered by using a larger or smaller number of models. The dispersion level should further improve if the number of hydrological models used is increased; however, there would probably be diminishing returns at the cost of increased computing expenses. Velázquez et al. (2011) obtained promising improvements using 16 hydrological models in their GESP. Similarly, using fewer than eight models would likely still provide significant improvements as compared to individual ESPs. However, there is probably no optimal number of hydrological models that can be used, since results have been shown to vary from catchment to catchment and from season to season.

Further tests were carried out by combining the ESPs into a GESP before post-processing to see if additional improvements could be obtained. However, this method resulted in poorer overall performance in terms of both the *ABDU* and *CRPSS* scores. This could be explained by a reduction in the GESP variability (accounting for less uncertainty) that would arise from applying a unique set of corrections to all merged ESPs, rather than correcting each ESP individually, which leading to more adapted corrections.

2.6.4 Impact of the hydrological season on the forecast

Looking at the results in Figure 2.7, it can be seen that there is little to no drop in accuracy for the DJFM over the nine forecasted days. This is likely due to the winter hydrological regime of the DJFM season, which is characterized by low flows and almost no liquid precipitation, which makes it easy to estimate short-term flow. For the other seasons, the accuracy of the ensembles drops progressively as the lead-time increases. In contrast to the DJFM season, the hydrological regime of the other seasons can vary greatly, and involves more variable processes (e.g., snowmelt, rainfall-runoff, etc.) that are harder to forecast, which could in turn explain the progressive decrease in ensemble accuracy over the 9 forecast days. This finding is supported by the larger inter-quartile range of the boxplots for the AM, JJA and SON versus the DJFM season.

It can be observed that the variability following QM and GESP manipulations is larger than that from DA, mainly during the AM season. The snowmelt process, which is highly active during the AM season, is mostly affected by the temperature, for which forecasts are less variable than for precipitations. The relative ease of forecasting temperature versus precipitation could explain the smaller inter-quartile range in the DA boxplots as compared to the summer seasons, seen in Figure 2.7. As mentioned previously, the median *CRPSS* scores reach below 0 after the sixth day of forecasting during the AM season (e.g., CS and LM catchments with GR4J model and GESP; Figure-A II-2 for the CS catchment). This suggests that the different manipulations face difficulties correcting the variable streamflow of the snowmelt season. These difficulties are also supported by the Talagrand diagrams (Figures 2.3, 2.4 and 2.5) for the AM season, which rarely exhibit a uniform distribution, even after all the manipulations have been performed.

2.6.5 Overall methodology performance analysis

As previously demonstrated, the use of EnKF as a data assimilation method, combined with the post-processing of data with the QM, significantly reduced deviations from uniformity (reduction of the *ABDU* score and more likely to fail to reject the hypothesis of uniformity

with the K-S test) and the bias in ESPs (resulting in an increase in the *CRPSS*) when compared to the OL simulation. The grouping together of the individual hydrological model ESPs into a GESP allowed reducing the under-dispersion problem, leading to the best overall results. These findings were observed over all catchments and for all hydrological seasons analyzed.

In addition, while the Talagrand diagrams showed systematically biased ESP forecasts in the OL simulation, the complete methodology resulted in distributions closer to uniformity. However, the U-shape could not be prevented in all cases, and notably so during the spring freshet over the months of April and May.

The different surface areas of the five catchments do not seem to have significant impacts on the distribution of the GESP forecasts. However, the smallest catchment, which is 1297 km², is arguably still quite large, and catchments smaller than 500 km² for example could see different conclusions.

2.6.6 Limitations

Some limitations to this study must be mentioned and are listed hereafter:

Number and length of NWPEs: In this project, a total of four and a half years of the ECMWF NWPE were available, which was considered sufficient to meet the study objectives. Ideally, the NWPE used to feed the hydrological models should be extended over a period of several years to allow a better assessment of the robustness of the proposed methodology. Furthermore, more than one NWPE could be used to test whether the results are consistent with other models (individually or combined into a multi-NWPE multi-hydrological model GESP).

Forecast lead-time: A lead-time of nine days was used for all forecasts in this study. This means that the findings are limited to this lead-time, and different conclusions could be obtained for forecasts beyond nine days. It should be noted that this choice corresponds to the forecast lead-time needed in the operational context.

Number of hydrological models: The methodology proposed in the study was evaluated using eight lumped hydrological models that also formed the GESP. The use of more or fewer hydrological models could have an impact on the results and on the robustness of the methodology. Moreover, using semi-distributed or distributed hydrological models could provide additional information not available within lumped models.

Number of catchments: Only five catchments were analyzed in the study. Moreover, these large catchments (from 1297 km² to 46,747 km²) are clustered in the same region in Quebec, Canada. Evaluating the methodology on smaller catchments (less than say 500 km²) and on catchments located in other climatic regions could bring different conclusions.

2.7 Conclusion and recommendations

This study presented a new methodology aimed at addressing the problems of bias and under-dispersion of ensemble forecasts (ESP) through the use of a multi-model approach to combine individual hydrological model ESPs into a GESP. To assess the performance and robustness of the proposed methodology, it was compared with each preceding manipulation against an OL simulation baseline. In order, the three main manipulations that were evaluated are data assimilation (DA) using the Ensemble Kalman Filter (EnKF) method, post-processing using quantile mapping (QM), and lastly, an aggregation of eight individual hydrological model ESPs into a GESP. The comparison was performed using several graphs and performance criteria typically used to evaluate forecasts, namely, the Talagrand diagram, the *ABDU* score, the K-S test and the *CRPSS*.

The DA with the EnKF initially improved the individual ESP dispersion (*ABDU* score) and accuracy (*CRPSS* value), as compared to the OL ensemble, but was still mostly undispersed or systematically biased. The QM did solve the systematic bias issue, but further reduced the *ABDU* scores and increased the *CRPSS* value over most individual ESPs. Despite the

improvements brought about by the DA and the QM, most of the ESPs remained under-dispersed over all catchments, and most notably, during the freshet season (AM).

Using a multi-model clustering of individual ESP forecasts into GESP forecasts was found to be effective in further reducing ensemble bias and under-dispersion when compared to the results of the majority of individual ESPs that underwent DA and QM. Some individual hydrological models performed better on a specific catchment and hydrological season, but overall, the GESP forecasts had lower *ABDU* scores, higher *CRPSS* values, and failed to reject the hypothesis of uniformity using the K-S test more often than the individual ESPs. The GESP was also found to perform similarly to or better than the best individual ESP for any of the catchments or seasons tested.

The results obtained in this study support the proposed methodology in reducing bias and ensemble under-dispersion. The methods used are simple and could easily be integrated in an operational context. The use of GESP forecasts can represent a powerful tool to add the variability to the ensembles needed to reach uniformity, while remaining as accurate as or better than the best individual ESP. It should be noted that in all cases, using any or a combination of DA, post-processing and multi-model approaches will significantly improve forecasts by reducing bias and correcting typical ensemble under-dispersions. However, it is important to remember that different data processing methods can add uncertainty to ESP forecasts, which can vary from season to season and from catchment to catchment, and ultimately, may be misinterpreted in a forecasting context. Future research could attempt to test different uncertainty estimations for driving the DA process, depending on the seasons, or try to introduce a probabilistic multi-model method with density forecasting to the proposed methodology.

2.8 Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada (Grant number CRDPJ 522126-17). The authors would like to thank *Rio Tinto* for providing

hydrometeorological and hydrological data of their catchments. We are also grateful to the European Centre for Medium-Range Weather Forecasts (ECMWF) for providing access to the NWEF forecast data from their computing and archiving facilities.

CHAPITRE 3

DISCUSSION

Cette section inclut une discussion plus détaillée des analyses des résultats présentés au chapitre 2. Des analyses additionnelles avec le modèle hydrologique IHACRES et le bassin versant Passes-Dangereuses (PD) sont présentées afin de mieux comparer l'impact de ces différentes composantes sur les résultats. Similairement à l'article présenté au chapitre précédent, chacune des étapes de la méthodologie proposée sera analysée séparément (DA, QM et GESP).

3.1 Analyse de la performance de la méthode d'assimilation de données d'EnKF

La méthode d'EnKF, utilisée pour le DA, vise à réduire le biais lié aux conditions initiales entre les ESPs et les observations en échantillonnant l'incertitude des intrants alimentant les modèles hydrologiques (p. ex. : observations des débits, précipitation et température). Les figures 2.3 et 3.1 présentent respectivement les résultats des simulations obtenus à partir de deux modèles hydrologiques, GR4J et IHACRES sur le bassin versant LSJ. Ces derniers sont représentatifs des six autres modèles hydrologiques sur les cinq bassins versants. Comme le montrent les deux figures, la méthode de DA permet de réduire le biais (valeur *ABDU*) de la distribution des ESPs par rapport à l'ESP de référence (open-loop; OL) pour toutes les périodes étudiées. Les ESPs OL proviennent des simulations hydrologiques initiales par chaque modèle hydrologique sans aucune manipulation de données subséquente (DA, QM, GESP). Les 50 membres météorologiques de l'ECMWF ont donc été fournis aux modèles hydrologiques, ce qui a permis d'obtenir un ensemble de 50 prévisions hydrologiques OL. Toutefois, cette amélioration est relativement marginale et demeure loin de l'objectif d'uniformité. En effet, la majorité des distributions sont encore sous-dispersées ou systématiquement biaisées, avec un rejet de l'hypothèse d'uniformité du test de Kolmogorov-Smirnov ($\alpha = 0,05$) dans le trois quarts des cas. Il est toutefois important de se rappeler que l'objectif premier du DA est de sélectionner les conditions initiales des modèles hydrologiques permettant de se rapprocher

des « vraies » observations. Il n'est donc pas garanti que les conditions initiales assimilées permettent de générer des ESPs plus dispersés et uniformes.

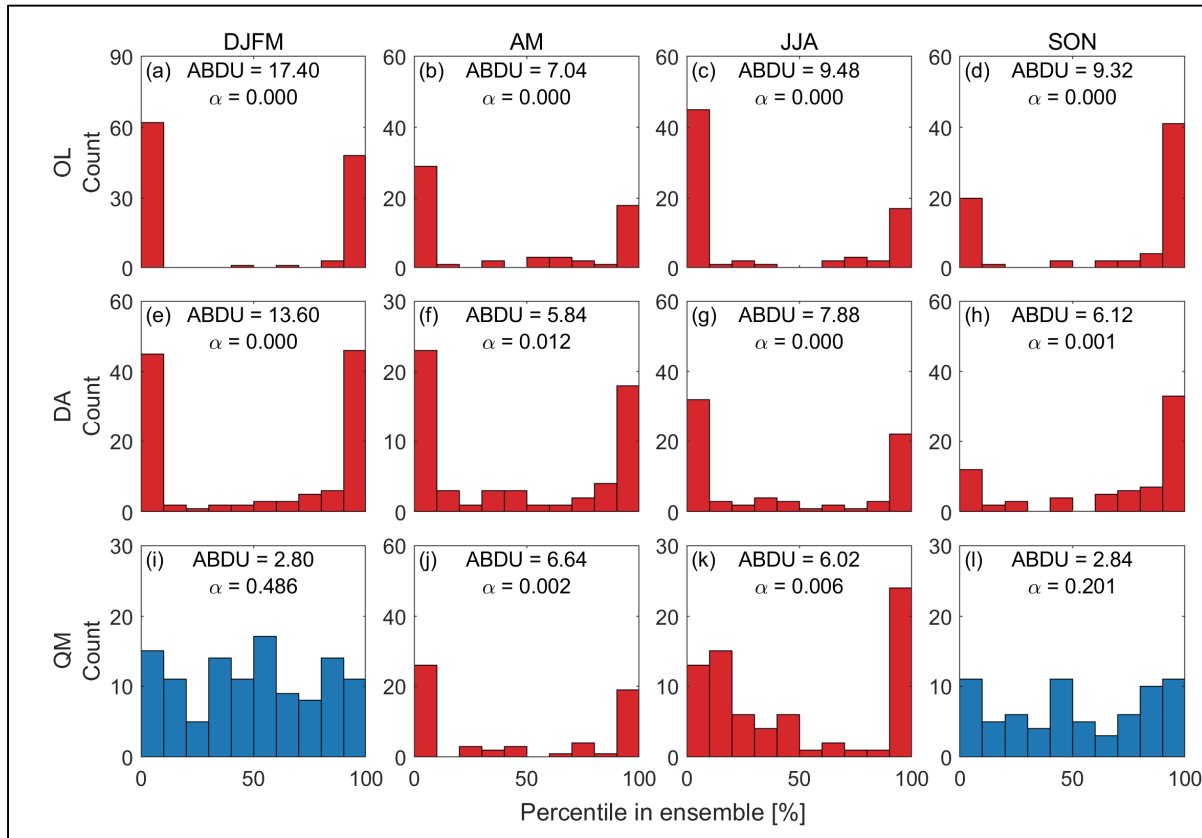


Figure 3.1 Diagrammes de Talagrand représentant le rang (en percentiles) occupé par les débits observés dans l'ensemble des prévisions hydrologiques à l'aide du modèle IHACRES sur le bassin versant LSJ. "OL" fait référence à la simulation de référence. "DA" réfère aux membres de validation des 1250 membres de l'ensemble après avoir effectué l'assimilation des données et "QM" réfère aux mêmes membres après avoir effectué l'assimilation des données et le post-traitement avec la correction par quantiles. Les diagrammes en bleu (rouge) indiquent que l'hypothèse d'uniformité n'est pas rejetée (rejetée) par le test K-S ($\alpha = 0,05$)

L'impact de la méthode de l'EnKF sur la précision des ESPs (évaluée avec le *CRPSS*) a été analysé à l'aide de boîtes à moustaches présentées aux figures 2.7 et 3.2. Ces figures démontrent que la méthode d'assimilation (DA) de données permet d'améliorer la précision de la quasi-totalité des ESPs par rapport à l'ESP OL durant les neuf jours de prévisions. Toutefois, il est possible d'observer sur la figure-A II-2 que la médiane du *CRPSS* obtenue à

la neuvième et dernière journée pour la saison AM du bassin versant CS devient négative, indiquant une baisse de précision de l'ESP, jamais observée sur les figures précédentes. Les figures 2.7 et 3.2 montrent aussi, sans surprise, que les valeurs *CRPSS* diminuent proportionnellement à la durée de la prévision. Visuellement, les valeurs *CRPSS* des ESPs assimilés au neuvième jour de prévision sont habituellement plus petites que celles de ces mêmes ESPs au premier jour de prévision. Ceci qui correspond aux conclusions de Bergeron et al. (2016), qui suggéraient que la méthode d'EnKF était particulièrement efficace sur de courtes périodes de prévision (p. ex. : de 2 à 3 jours).

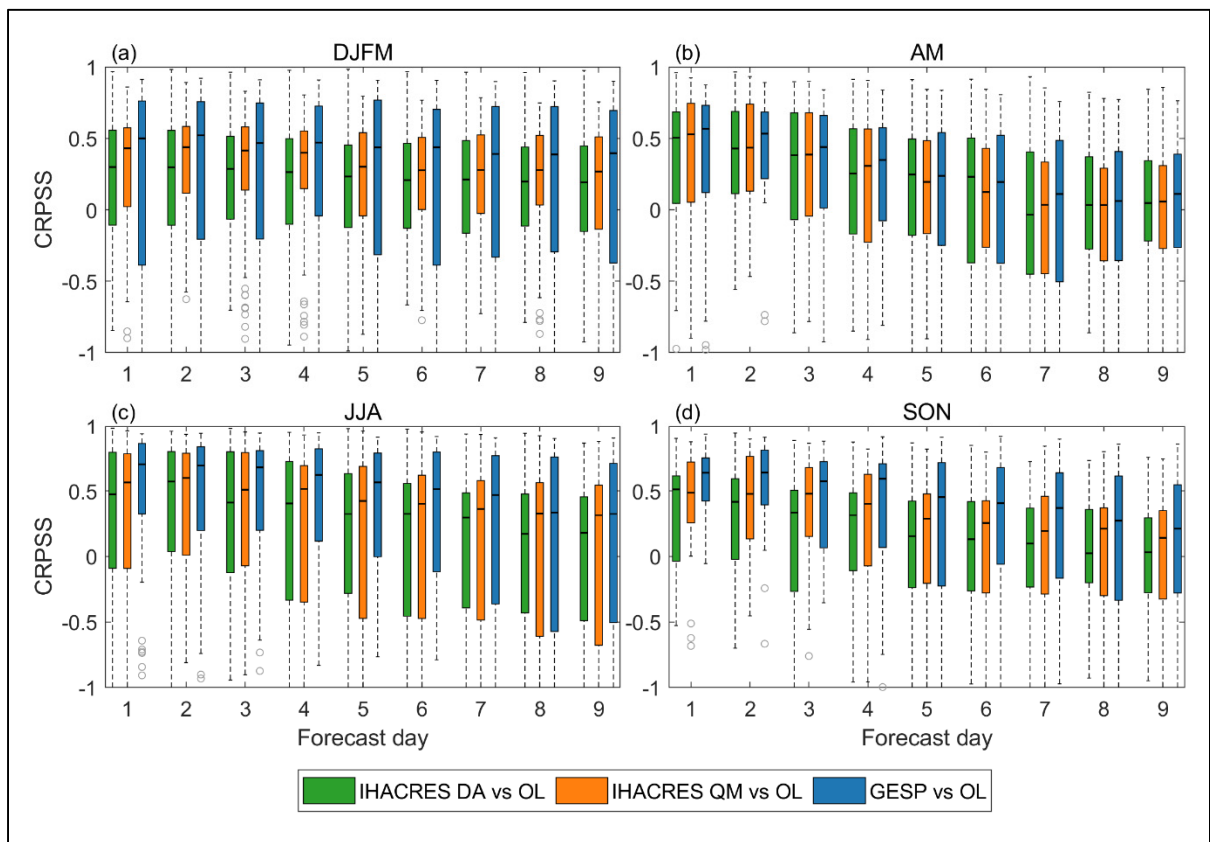


Figure 3.2 Boîtes à moustaches des valeurs CRPSS des membres de l'ensemble hydrologique du modèle IHACRES et de l'ensemble multi-modèles (GESP) sur le bassin versant LSJ. Un CRPSS supérieur à zéro indique une amélioration de la précision globale des prévisions par rapport aux simulations de référence (OL) d'IHACRES. Les membres dont les valeurs du CRPSS sont inférieures à -1 sont considérés comme très mauvais et ne sont pas indiqués afin d'éviter de surcharger le graphique

Finalement, la figure 2.6 présente les diagrammes de Talagrand du GESP sur le bassin versant LSJ selon différents écarts-types appliqués aux paramètres de DA. La figure 2.6 montre qu'il peut être difficile de choisir les paramètres optimaux à appliquer à la méthode d'EnKF étant donné la similitude entre les diagrammes de Talagrand obtenus. Une modification de l'écart-type appliqué aux débits observés (10 % à 15 %), aux précipitations (10 % à 50 %) et aux températures (1 °C à 2 °C) fournis à l'algorithme de DA ne permet pas d'obtenir des améliorations significatives au niveau de l'uniformité des ESPs. Ces écarts-types représentent les paramètres fournis au DA afin d'échantillonner l'incertitude des distributions de ces différentes variables. La valeur observée de ces variables est alors considérée comme la moyenne, alors que l'écart-type choisi représente l'incertitude attendue pour ces mêmes variables. D'ailleurs, Thibault et Anctil (2015) avaient suggéré que la détermination des paramètres à appliquer à la méthode de l'EnKF n'était pas triviale et qu'elle dépendait directement du modèle hydrologique utilisé et des caractéristiques propres à chaque bassin versant. L'impact de l'écart-type appliqué aux températures n'est pas montré sur la figure 2.6 en raison des changements négligeables sur les diagrammes de Talagrand.

3.2 Analyse de la performance de la méthode de post-traitement

La méthode de post-traitement des prévisions hydrologiques utilisée, la correction par quantiles (QM), a été réalisée afin de corriger la distribution des ESPs comparativement à celle des débits observés. Elle consiste à calculer la correction requise entre les percentiles de la distribution des ESPs et celle des débits observés durant chacune des saisons, et ce, pour chaque journée de prévision (jour 1 à jour 9). Seulement 80 % des ESPs d'une même saison sont utilisés afin de calculer les coefficients de correction, mais ceux-ci sont appliqués à tous les ESPs (100 %) de la saison. Les ESPs corrigés restants (20 %) permettent de valider la méthode de correction et sont ainsi présentés sur toutes les figures de résultats. La méthode, appliquée sur les ESPs du bassin versant LSJ, ceux-ci simulés par le modèle hydrologique GR4J (figure 2.3) ou par le modèle hydrologique IHACRES (figure 3.1), permet de réduire la valeur *ABDU*, à l'exception de la période de fonte de la neige (saison AM). Cette exception est probablement due au fait que la méthode de correction par quantiles modifie la distribution

de l'ESP en appliquant les mêmes coefficients de correction d'une journée de prévision à toute la saison, ce qui ne permet pas de conserver la relation temporelle entre les simulations et les observations. Cette faiblesse de la méthode, déjà soulevée par Li et al. (2017), pourrait ainsi être problématique en période de crue en raison du processus physique de fonte de la neige différent d'une année à l'autre (p. ex. : quantité de pluie, températures et ensoleillement durant la période de fonte). Ce processus a alors un impact significatif sur les débits de crue et la méthode QM pourrait ne pas être capable de corriger adéquatement les distributions dans cette situation précise.

Dépendamment de la saison analysée, du modèle hydrologique utilisé et du bassin versant à l'étude, le post-traitement avec le QM peut dans certains cas contribuer à rendre les ESP plus fiables. Considérant que le test de Kolmogorov-Smirnov ne rejette pas l'hypothèse d'uniformité (diagrammes de Talagrand en bleu des figures 2.4 et 3.3), ceci indique que la méthode est parvenue à améliorer la distribution de l'ESP comparativement à l'ESP OL. Toutefois, comme le montrent les figures 2.4 et 3.3, regroupant la distribution des ESPs simulés par chaque modèle hydrologique et pour les bassins versants LSJ et PD respectivement, le QM n'est pas en mesure d'atténuer la problématique de sous-dispersion des ESPs. Les résultats sur les autres bassins versants (figures A I-1 (LM), A I-2 (CD) et A I-3 (CS)) supportent cette conclusion.

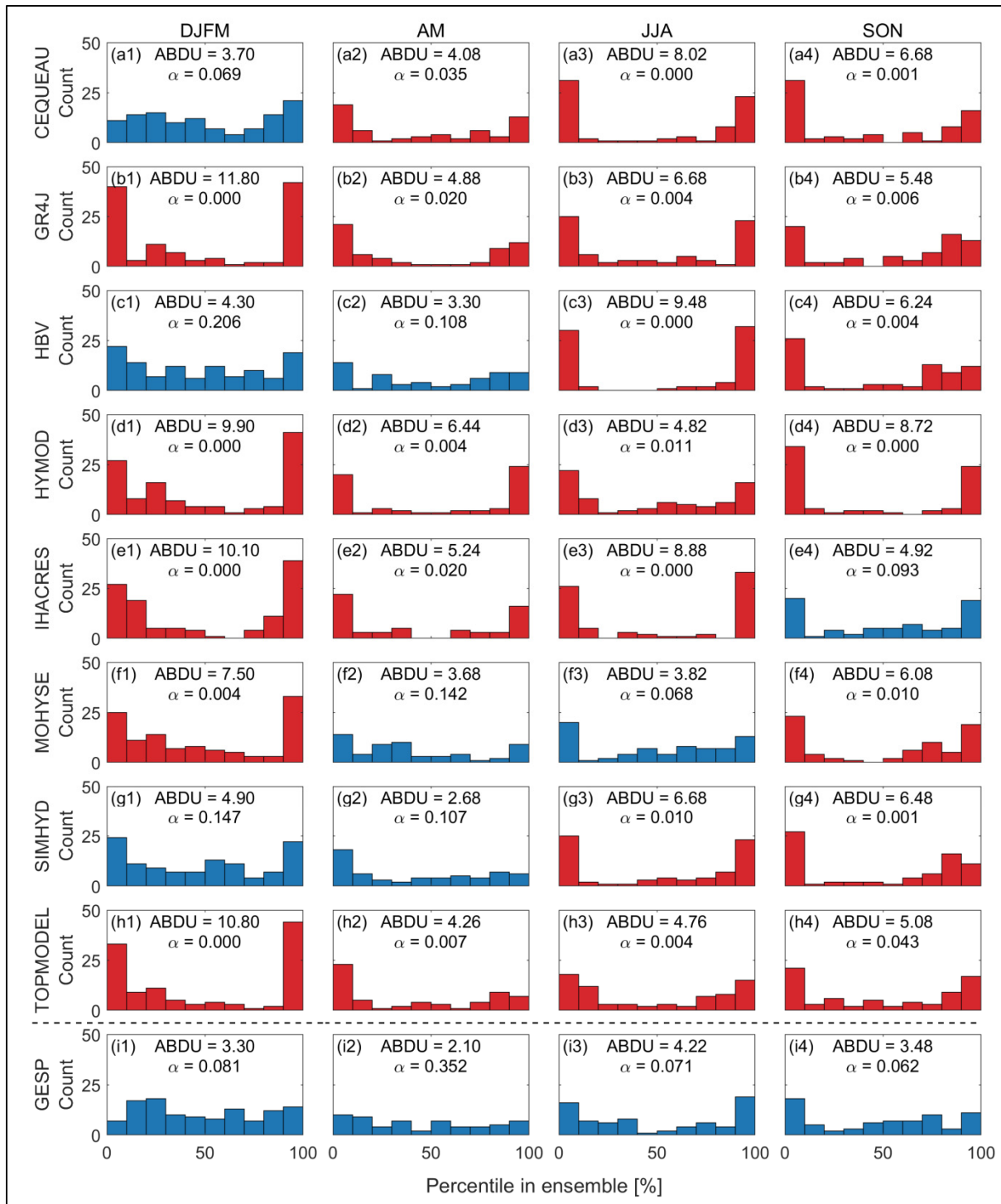


Figure 3.3 Diagrammes de Talagrand représentant le rang (en percentiles) occupé par les débits observés dans les ESPs des huit modèles hydrologiques différents après avoir effectué le DA et le QM sur le bassin versant PD. Les diagrammes en bleu (rouge) indiquent que l'hypothèse d'uniformité n'est pas rejetée (rejetée) par le test K-S ($\alpha = 0,05$)

Finalement, les figures 2.7 et 3.2 indiquent que la précision des ESPs corrigés par QM est améliorée systématiquement par rapport à l'ESP OL puisque la valeur médiane du *CRPSS* est supérieure à zéro dans tous les cas. La méthode QM suivant l'assimilation de données (DA) permet aussi d'obtenir des valeurs *CRPSS* supérieures ou égales à celles obtenues au neuvième jour des prévisions sans l'étape du post-traitement. Ceci indique que QM permet d'améliorer davantage la précision des ESPs à plus long terme. Cette conclusion est supportée sur les quatre saisons lorsque les ESPs sont simulés par les modèles hydrologiques GR4J et IHACRES sur le bassin versant LSJ (figures 2.7 et 3.2), et dans 81 % de toutes les autres situations incluant les huit modèles hydrologiques et les cinq bassins versants.

Comme prévu, les résultats montrent que la méthode QM est capable de réduire le biais des ESPs au pas de temps journalier, ce qui corrobore les conclusions de Hashino et al. (2007) ayant démontrées que le QM était une méthode de post-traitement performante pour réduire les biais des débits mensuels. Toutefois, similairement à l'assimilation de donnée, cette méthode ne permet pas de réduire significativement la problématique de sous-dispersion des ESPs.

3.3 Analyse de la performance de la méthode multi-modèle

En regroupant les ESPs des huit modèles hydrologiques en un grand ensemble (GESP) via une approche multi-modèle, l'objectif était d'uniformiser la distribution de l'ensemble final en regroupant les incertitudes de chaque modèle individuel. En observant les résultats du GESP pour les bassins versants LSJ et PD représentés sur les diagrammes de Talagrand de la figure 2.5, ces derniers suggèrent que l'utilisation du GESP permet d'obtenir des ensembles plus uniformes comparativement à la presque totalité des modèles hydrologiques individuels (voir la figure 2.4 pour le bassin versant LSJ et la figure 3.3 pour le bassin versant PD). Individuellement, certains ESPs individuels assimilés et post-traités permettent quand même d'obtenir des valeurs *ABDU* inférieures aux valeurs obtenues par le GESP, dépendamment de la saison et du bassin versant analysé. Toutefois, puisque le modèle hydrologique considéré comme étant le meilleur selon l'*ABDU* varie selon la saison et le bassin versant, le GESP est

une meilleure option qui offre systématiquement une des meilleures, sinon la meilleure performance. D'ailleurs, dans 17 cas sur 20, le test de Kolmogorov-Smirnov ne rejette pas l'hypothèse d'uniformité ($\alpha = 0.05$). Les figures en annexe I (figures A I-1, A I-2 et A I-3) permettent aussi d'observer que les scores *ABDU* des GESPs sont inférieurs aux scores *ABDU* obtenus par les ESPs individuels sur les trois autres bassins versants (LM, CD et CS) dans la très grande majorité des cas.

La précision des GESPs selon le *CRPSS* est aussi supérieure à celle des ESPs dans la majorité des situations analysées (différentes saisons, bassins versants et modèles hydrologiques). La figure 2.7 montre la performance du GESP par rapport au modèle hydrologique GR4J sur le bassin LSJ. Il est possible d'y observer que les valeurs médianes du *CRPSS* du GESP sont égales ou supérieures à tous les *CRPSS* de l'ESP GR4J assimilé et post-traité durant les saisons DJFM, JJA et SON. Dans le cas de la saison AM, la précision du GESP est légèrement réduite notamment pour le plus petit bassin versant (CS; voir figures A II-2 et A II-4). Dans le cas du bassin versant PD (figure 3.2), les valeurs *CRPSS* du GESP sont égales ou supérieures à l'ESP du modèle IHACRES, ce qui suggère que le GESP est une meilleure option. En tenant compte des quatre périodes à l'étude, des huit modèles hydrologiques et des cinq bassins versants (pour un total de 160 ESPs évalués), la médiane des valeurs *CRPSS* des GESPs est supérieure ou égale aux médianes des ESPs individuels assimilés et post-traités dans 73 % des cas. Ces conclusions sont supportées par Velázquez et al. (2011), qui ont aussi démontré un avantage de regrouper les ESPs en GESP.

Dans la méthodologie proposée, le regroupement des ESPs en GESP s'effectue après l'assimilation de données et le post-traitement de chaque ESP. Le regroupement a aussi été testé avant de faire le post-traitement (sur les ESPs uniquement assimilés) de manière à exécuter une seule correction de biais directement sur le GESP. Cette méthode n'a toutefois pas permis d'obtenir une meilleure performance. Il est possible que le post-traitement du GESP (plutôt que des ESPs) ne permette pas de conserver l'incertitude propre à chaque modèle hydrologique en appliquant le même coefficient de correction à tous les ESPs inclus dans le GESP.

Il est important de noter que dans ce projet de recherche, les GESPs étaient formés de huit ESPs provenant des huit modèles hydrologiques individuels. Il est possible que les conclusions aient été différentes si plus ou moins de modèles hydrologiques avaient été utilisés afin de former les GESPs. L'utilisation de moins de huit modèles pourrait permettre d'obtenir d'aussi bons résultats et d'autres études pourraient être effectuées pour déterminer nombre optimal de modèles hydrologiques à utiliser en contexte de multi-modèle. Toutefois, comme les résultats sont différents de saison en saison et de bassin versant en bassin versant, il pourrait s'avérer difficile de trouver une réponse unique à cette question.

3.4 Impact des différentes saisons sur les prévisions d'ensemble

La séparation des résultats en quatre périodes distinctes (ou saisons) avait pour objectif d'analyser l'impact des différents régimes hydrologiques d'une année complète sur les ESPs et les GESPs. D'emblée, les saisons ne semblent pas avoir d'impact sur l'uniformité des ESPs OL. En effet, les ESPs sont, pour la plupart, sous-dispersés ou biaisés systématiquement sur l'année complète (voir figures 2.3 et 3.1). Sur ces mêmes figures, il est possible d'observer que la méthode DA permet d'améliorer la dispersion des ESPs pour toutes les saisons, en réduisant principalement les biais positifs et négatifs. Cela a pour effet d'abaisser leurs scores *ABDU*, sans toutefois résoudre la problématique de sous-dispersion des ESPs dans la plupart des cas. Ensuite, la méthode QM (résultats présentés sur les figures 2.4, 3.3, A I-1, A I-2 et A I-3) permet à son tour de réduire davantage le score *ABDU* des ESPs comparativement à la méthode DA effectuée seule durant trois saisons sur quatre. En effet, la méthode est performante durant les saisons DJFM, JJA et SON, mais peut parfois augmenter le score *ABDU* des ESPs durant la saison de fonte de la neige (AM). D'ailleurs, les ESPs restent généralement sous-dispersés durant la saison AM suite au post-traitement. Ces mêmes figures permettent finalement d'observer l'impact du GESP sur les différentes saisons. Les scores *ABDU* des GESPs sont rarement plus hauts que tous les scores *ABDU* atteints par les ESPs individuels, et ce, sur tous les bassins versants à l'étude. De plus, la dispersion des GESPs est visuellement beaucoup plus près de l'uniformité que les ESPs, ce que le test statistique de Kolmogorov-Smirnov suggère en ne rejetant pas l'hypothèse d'uniformité ($\alpha = 5 \%$) dans la majorité des situations (17/20).

Il est aussi possible d'analyser, sur les différents bassins versants, l'impact des saisons sur la précision des ESPs à l'aide du *CRPSS* (figures 2.7, 3.2, A II-1, A II-2, A II-3 et A II-4). Dans un premier temps, la saison hivernale (DJFM) présente des *CRPSS* assez stables sur les neuf jours de prévision, sans réduction majeure des *CRPSS* à mesure que la prévision progresse. L'hypothèse émise est que les débits hivernaux habituellement faibles et les précipitations majoritairement sous forme solide rendent plus facile la prévision des débits, permettant de conserver la précision des prévisions sur leurs durées complètes. En contrepartie, les médianes des *CRPSS* des autres saisons diminuent au fur et à mesure que la prévision progresse, probablement dû à la plus grande variabilité des débits durant ces saisons. Les médianes des *CRPSS* obtenues sur les ESPs et les GESPs durant la saison de fonte de neige (AM) peuvent même descendre sous zéro à partir d'un délai de cinq ou six jours dans la prévision (voir exemple sur la figure-A II-2).

3.5 Analyse de la méthodologie globale

Globalement, la combinaison de toutes les méthodes proposées a bel et bien permis de réduire la sous-dispersion et les biais systématiques des ESPs, comme l'indique l'amélioration des critères de performance durant toutes les saisons. La méthode DA combinée à la méthode QM a permis dans un premier temps d'améliorer l'uniformité des ESPs (réduction du score *ABDU*) par rapport aux ESPs OL. Ces ESPs restaient malgré tout sous-dispersés dans la majorité des cas. Le GESP a finalement permis de réduire davantage le score *ABDU* et ainsi de régler la problématique de sous-dispersion des ensembles la plupart des cas. Toutefois, la méthodologie proposée semble moins performante durant la saison de la fonte de la neige (AM). Durant cette période, la méthodologie ne parvenait pas toujours à corriger correctement la distribution des ESPs et des GESPs. D'ailleurs, il est possible visuellement de constater que durant la saison AM, les GESPs sur les bassins versants CD, CS et LSJ demeurent sous-dispersés (figure 2.5). Ceci pourrait être expliqué par la particularité du cycle hydrologique de la saison AM, durant laquelle les débits de crue dépendent directement du processus de fonte de la neige, qui peut être différent d'une année à l'autre. Cette particularité pourrait rendre les corrections

appliquées (DA, QM et GESP) à l'ensemble des prévisions de la saison AM, indépendamment des années, moins efficaces.

Finalement, la superficie des bassins versants ne semble pas avoir un impact significatif sur les résultats puisque l'amélioration des critères de performances obtenue sur le plus grand bassin versant, LSJ (46 747 km²), était semblable à celle obtenue sur tous les autres bassins versants de l'étude, incluant le plus petit bassin, CS (1297 km²).

3.6 Limitations

Cette section présente les limitations de l'étude et l'impact que pourrait avoir l'utilisation de données différentes. Les limites liées aux prévisions météorologiques d'ensemble, aux modèles hydrologiques et aux bassins versants seront énoncées.

3.6.1 Prévisions météorologiques d'ensemble

Le NWPE de l'ECMWF utilisé dans cette étude couvrait la période disponible de janvier 2015 à juin 2019 et chacune des prévisions météorologiques s'étendait sur une durée de neuf jours. Bien que cette période de quatre ans et demi de données ait été jugée suffisante pour vérifier la performance de la méthodologie, des NWPEs couvrant une période plus longue (>10 ans) auraient permis de confirmer sa robustesse à plus long terme. Ceci aurait permis du même coup d'évaluer l'impact de la variabilité naturelle du climat. De plus, leur durée de neuf jours limite les conclusions de l'étude à cette période temporelle et la performance de la méthodologie pourrait s'avérer différente à plus long terme (p. ex. : prévisions sur plusieurs semaines ou même plusieurs mois). Finalement, bien que le GESP dans cette étude soit formé d'un seul NWPE avec plusieurs modèles hydrologiques, l'utilisation de plusieurs NWPEs provenant de plusieurs bases de données (p. ex. : le NWPE d'Environnement et Changements climatiques Canada; ECCC ou encore du *National Oceanic and Atmospheric Administration*; NOAA) pourrait être bénéfique ou défavorable aux GESPs.

3.6.2 Modèles hydrologiques

Dans cette étude, huit modèles hydrologiques ont été sélectionnés afin de valider la méthodologie proposée. L'hypothèse émise était qu'un nombre élevé de modèles hydrologiques pourrait s'avérer bénéfique aux GESPs, qui permettent d'inclure l'incertitude de tous ces modèles individuels. Toutefois, il est aussi important de simplifier autant que possible la méthodologie afin de permettre son utilisation en contexte opérationnel. D'emblée, l'utilisation d'un nombre plus élevé que huit modèles pourrait améliorer davantage la dispersion des GESPs avec un rendement décroissant. Cependant, il est aussi probable qu'un nombre plus restreint de modèles puisse permettre d'obtenir des résultats similaires. Considérant que chaque modèle hydrologique est conçu différemment, l'utilisation de modèles complètement différents pourrait modifier les résultats obtenus (p.ex. deux modèles pourraient générer un GESP aussi performant qu'une dizaine d'autres modèles). Finalement, l'utilisation de modèles hydrologiques semi-distribués ou distribués n'a pas été examinée et pourrait engendrer aussi des conclusions différentes.

3.6.3 Bassins versants

La méthodologie proposée a été vérifiée sur cinq bassins versants au Québec, Canada. Ces bassins versants sont tous situés dans la région du Lac-Saint-Jean et ont conséquemment une climatologie très similaire. Ils sont notamment caractérisés par un régime hydrologique contrôlé par la fonte de neige au printemps. De plus, ils sont tous considérés comme des bassins versants ayant une grande superficie, étant donné que le bassin versant le plus petit (CS) s'étend tout de même sur 1297 km². Les conclusions de l'étude s'appliquent donc à des bassins versants comparables à ceux-ci et la performance de la méthodologie pourrait être différente sur des bassins versants ayant une climatologie différente ou bien une superficie plus petite (< 500 km²).

CONCLUSION

Ce mémoire présente une nouvelle méthodologie visant à améliorer les prévisions hydrologiques d'ensemble (ESPs) dans un contexte opérationnel. Dans un tel contexte, la dispersion et la précision des ESPs jouent un rôle primordial pour la prise de décisions opérationnelles en contexte de gestion de risques. Une mauvaise gestion de la ressource en eau peut engendrer des risques pour la sécurité du public, pour l'environnement et économiques. En plus d'être plus rentable, la gestion efficace de l'eau contribue à la lutte aux changements climatiques en retardant la construction de nouvelles structures de retenues (p. ex. : digues, barrages, etc.) qui modifient considérablement les écosystèmes.

L'objectif de la méthodologie proposée est de réduire le biais et la sous-dispersion des ESPs afin de les rendre plus représentatifs de la réalité. Pour ce faire, une assimilation de données (DA) en utilisant les filtres d'ensemble de Kalman (EnKF) est dans un premier temps effectuée sur chaque ESP. Suite au DA, un post-traitement par correction de biais (QM) des ESPs est utilisé afin de corriger la distribution des différents ensembles. Finalement, une approche multi-modèle est employée afin de regrouper les ESPs dans un grand ensemble hydrologique (GESP). Afin d'évaluer la performance des différentes méthodes énumérées, plusieurs critères de performances ont été choisis : les diagrammes de Talagrand, les scores *ABDU* et *CRPSS*, puis le test Kolmogorov-Smirnov. Tous ces critères évaluaient la performance des ESPs modifiés par la méthodologie comparativement à l'ESP OL.

Avant d'appliquer les différentes méthodes de correction, huit modèles hydrologiques ont été sélectionnés et calibrés sur cinq bassins versants au Lac-Saint-Jean gérés par le partenaire industriel *Rio Tinto*. Une prévision météorologique d'ensemble (NWPE) de 2015 à 2019 a été extraite de la base de données du Centre européen pour les prévisions météorologiques à moyen terme (ECMWF). Les 50 membres de la prévision météorologique d'ensemble ont servi d'intrants aux modèles hydrologiques calibrés afin de constituer les ESPs.

Malgré qu'elles n'aient pas réussi à rendre la dispersion des ESPs uniforme sur les diagrammes de Talagrand, les méthodes DA et QM ont permis de réduire significativement les scores *ABDU* (distance par rapport à l'uniformité) et d'augmenter les *CRPSS* (critère de précision) des ESPs par rapport à l'ESP OL. La performance de ces méthodes a toutefois été réduite durant la saison AM, pendant laquelle la fonte de neige se produit. Le regroupement multi-modèle des ESPs en GESPs a pour sa part démontré son efficacité afin de réduire le biais et la sous-dispersion des ensembles. La dispersion des GESPs était, dans la plupart des cas, plus près de l'uniformité que tous les ESPs individuels. De plus, les GESPs présentaient les scores *ABDU* les plus faibles et des *CRPSS* égaux ou plus élevés que la plupart des ESPs et ce, peu importe la saison ou le bassin versant analysé.

Les résultats obtenus montrent donc qu'il est possible de réduire le biais et d'uniformiser les ESPs en appliquant la méthodologie proposée. Comme la performance des ESPs individuels n'est pas constante de saison en saison et de bassin versant en bassin versant, les GESPs deviennent la solution à privilégier. Il est toutefois important de noter quelques limites à la méthodologie proposée (aussi discutées dans le chapitre 3):

- 1) le nombre et la durée des prévisions météorologiques d'ensemble (NWPEs);
- 2) le nombre et le type de modèles hydrologiques utilisés;
- 3) la superficie et la climatologie des bassins versants analysés.

En terminant, l'étude n'a pas permis d'identifier les paramètres optimaux à appliquer à l'assimilation de données par EnKF, étant donné le peu d'impacts que les différents paramètres avaient sur les résultats.

RECOMMANDATIONS

Dans un premier temps, plusieurs recommandations sont émises afin d'adresser les limites de la méthodologie détaillées au chapitre 3 :

- 1) utiliser plusieurs sources de prévisions météorologiques d'ensemble (NWPEs) et inclure les prévisions hydrologiques d'ensemble (ESPs) formés à partir de ces nouvelles prévisions aux grands ensembles multi-modèles déjà utilisés (GESPs);
- 2) analyser la performance de la méthodologie proposée lorsqu'un nombre différent de huit modèles hydrologiques est utilisé et évaluer l'apport de modèles hydrologiques paramétrés différemment pour échantillonner l'incertitude paramétrique;
- 3) vérifier les conclusions de l'étude sur des bassins versants ayant des caractéristiques (superficies, régimes hydrologiques, localisation, topographie) différentes des bassins versants analysés dans la présente recherche.

Il a été observé que les GESPs pouvaient demeurer sous-dispersés principalement durant la période de fonte de la neige (saison AM). L'algorithme CEMANEIGE de modélisation du processus de fonte de la neige a été utilisé dans cette recherche, mais il est possible qu'un autre algorithme puisse ajouter la variabilité manquante aux GESPs afin que leurs distributions soient uniformes. Il est donc recommandé d'ajouter de nouveaux ESPs qui auront utilisé un autre algorithme de modélisation du processus de fonte de la neige.

Alors que les paramètres d'assimilation de donnée étaient identiques d'une saison à l'autre, d'autres études pourraient tenter de déterminer les paramètres optimaux à utiliser pour chacune des saisons individuelles afin de tenir compte de l'impact des différents régimes hydrologiques sur la dispersion des ensembles. Finalement, bien qu'un regroupement multi-modèle simple ait été choisi dans la méthodologie pour des raisons pratiques, de prochains projets de recherches pourraient être effectués en appliquant une méthode multi-modèle plus complexe, permettant d'ajouter un poids différent aux ESPs individuels lors de l'étape du regroupement pour former le GESP.

ANNEXE I

DIAGRAMMES DE TALAGRAND EN MATÉRIEL SUPPLÉMENTAIRE

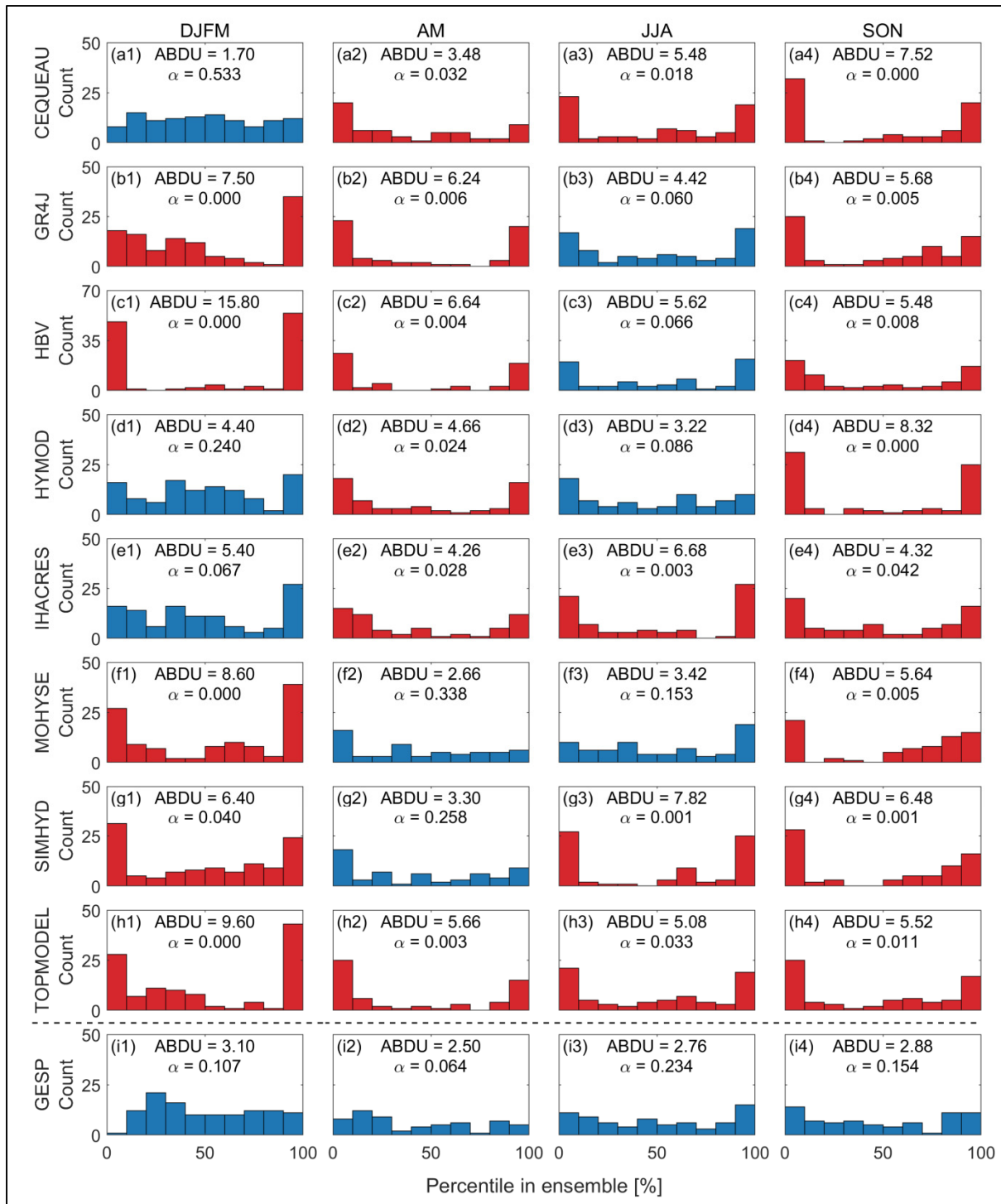


Figure-A I-1 Diagrammes de Talagrand représentant le rang (en percentiles) occupé par les débits observés dans les ESPs des huit modèles hydrologiques différents après avoir effectué le DA et le QM sur le bassin versant LM. Les diagrammes en bleu (rouge) indiquent que l'hypothèse d'uniformité n'est pas rejetée (rejetée) par le test K-S ($\alpha = 0,05$)

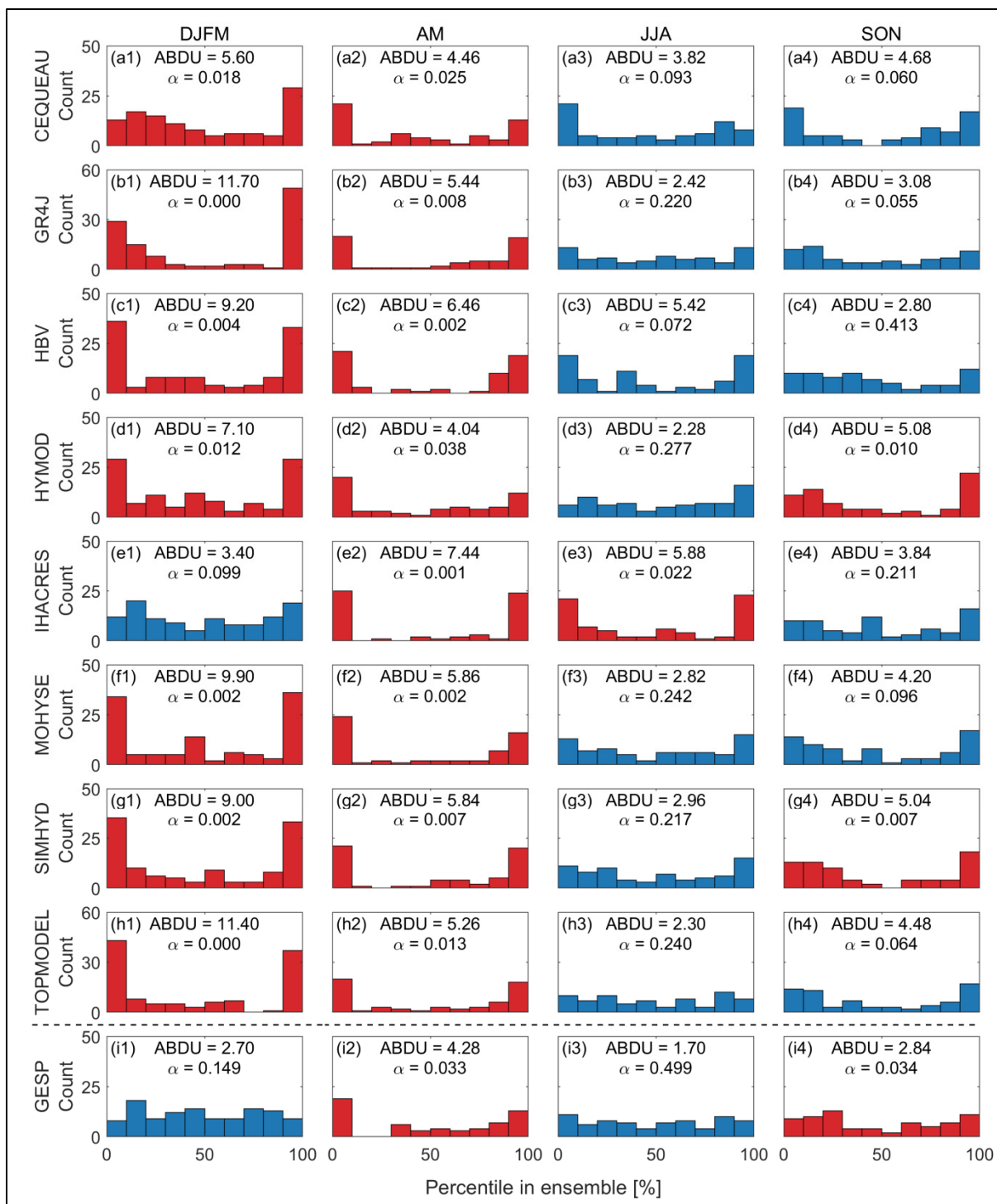


Figure-A I-2 Diagrammes de Talagrand représentant le rang (en percentiles) occupé par les débits observés dans les ESPs des huit modèles hydrologiques différents après avoir effectué le DA et le QM sur le bassin versant CD. Les diagrammes en bleu (rouge) indiquent que l'hypothèse d'uniformité n'est pas rejetée (rejetée) par le test K-S ($\alpha = 0,05$)

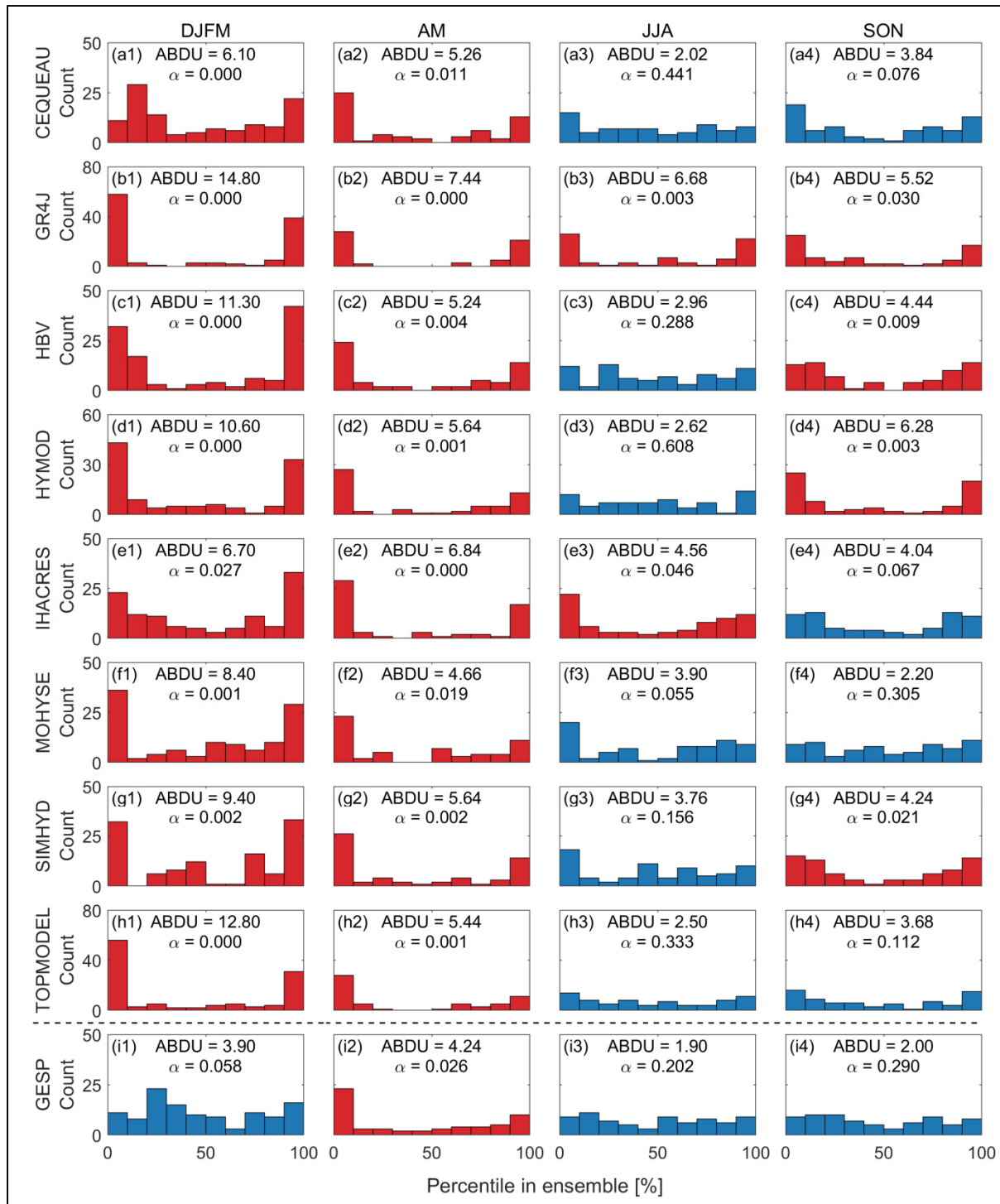


Figure-A I-3 Diagrammes de Talagrand représentant le rang (en percentiles) occupé par les débits observés dans les ESPs des huit modèles hydrologiques différents après avoir effectué le DA et le QM sur le bassin versant CS. Les diagrammes en bleu (rouge) indiquent que l'hypothèse d'uniformité n'est pas rejetée (rejetée) par le test K-S ($\alpha = 0,05$)

ANNEXE II

BOÎTES À MOUSTACHES DES VALEURS CRPSS EN MATÉRIEL SUPPLÉMENTAIRE

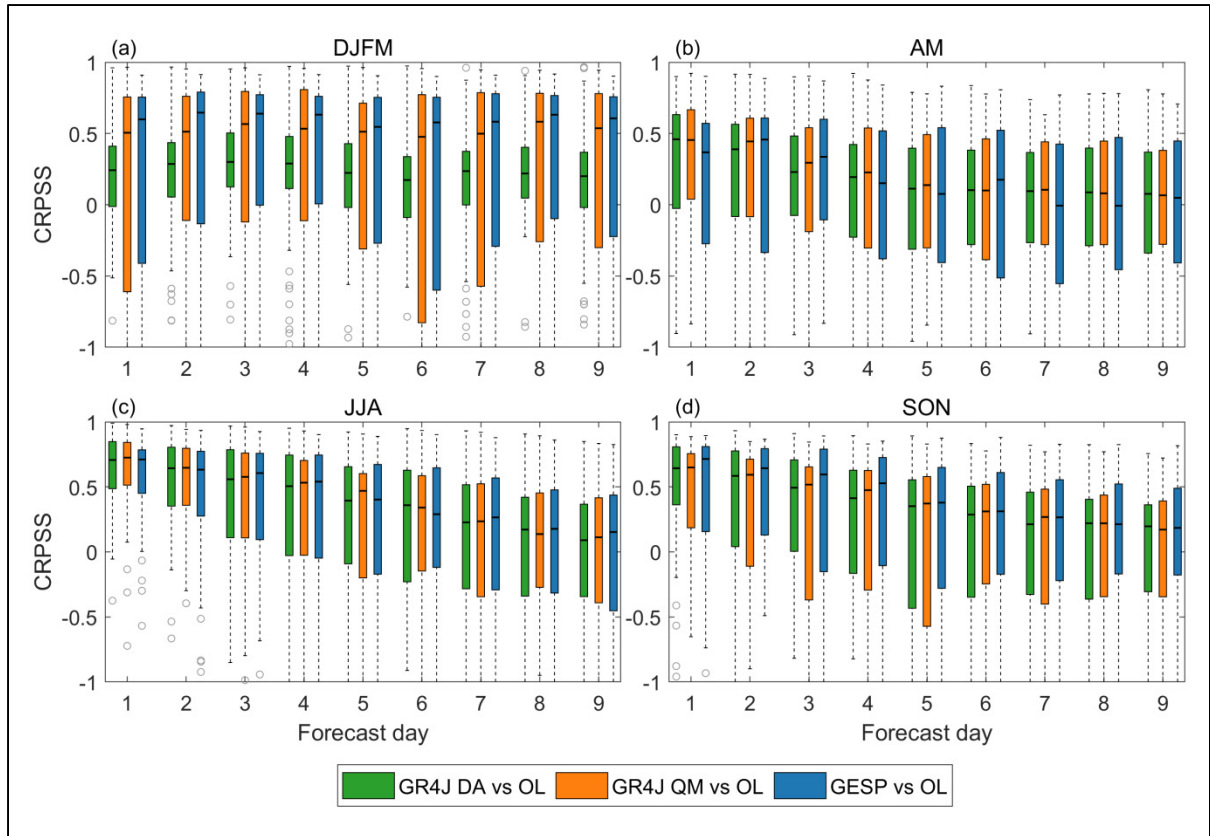


Figure-A II-1 Boîtes à moustaches des valeurs CRPSS des membres de l'ensemble hydrologique du modèle GR4J et de l'ensemble multi-modèles (GESP) sur le bassin versant CD. Un CRPSS supérieur à zéro indique une amélioration de la précision globale des prévisions par rapport aux simulations de référence (OL) de GR4J. Les membres dont les valeurs du CRPSS sont inférieures à -1 sont considérés comme très mauvais et ne sont pas indiqués afin d'éviter de surcharger le graphique

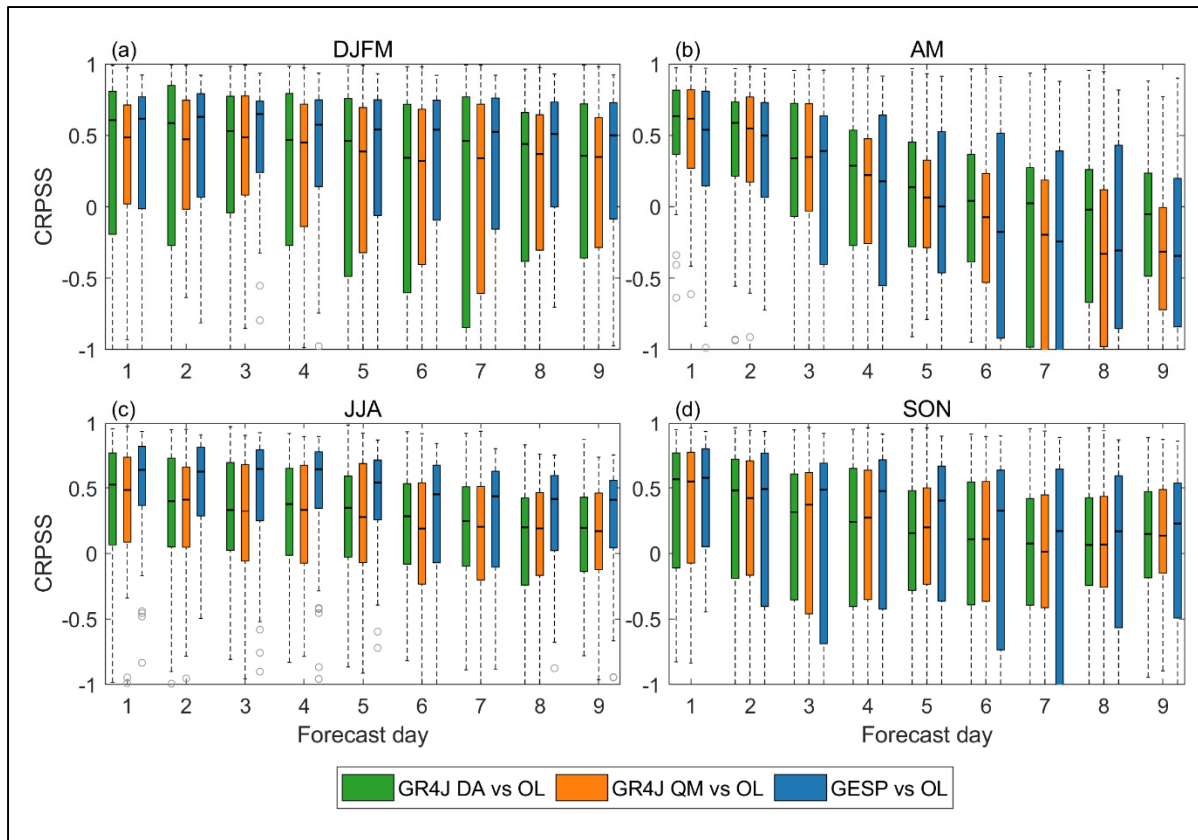


Figure-A II-2 Boîtes à moustaches des valeurs CRPSS des membres de l'ensemble hydrologique du modèle GR4J et de l'ensemble multi-modèles (GESP) sur le bassin versant CS. Un CRPSS supérieur à zéro indique une amélioration de la précision globale des prévisions par rapport aux simulations de référence (OL) de GR4J. Les membres dont les valeurs du CRPSS sont inférieures à -1 sont considérés comme très mauvais et ne sont pas indiqués afin d'éviter de surcharger le graphique

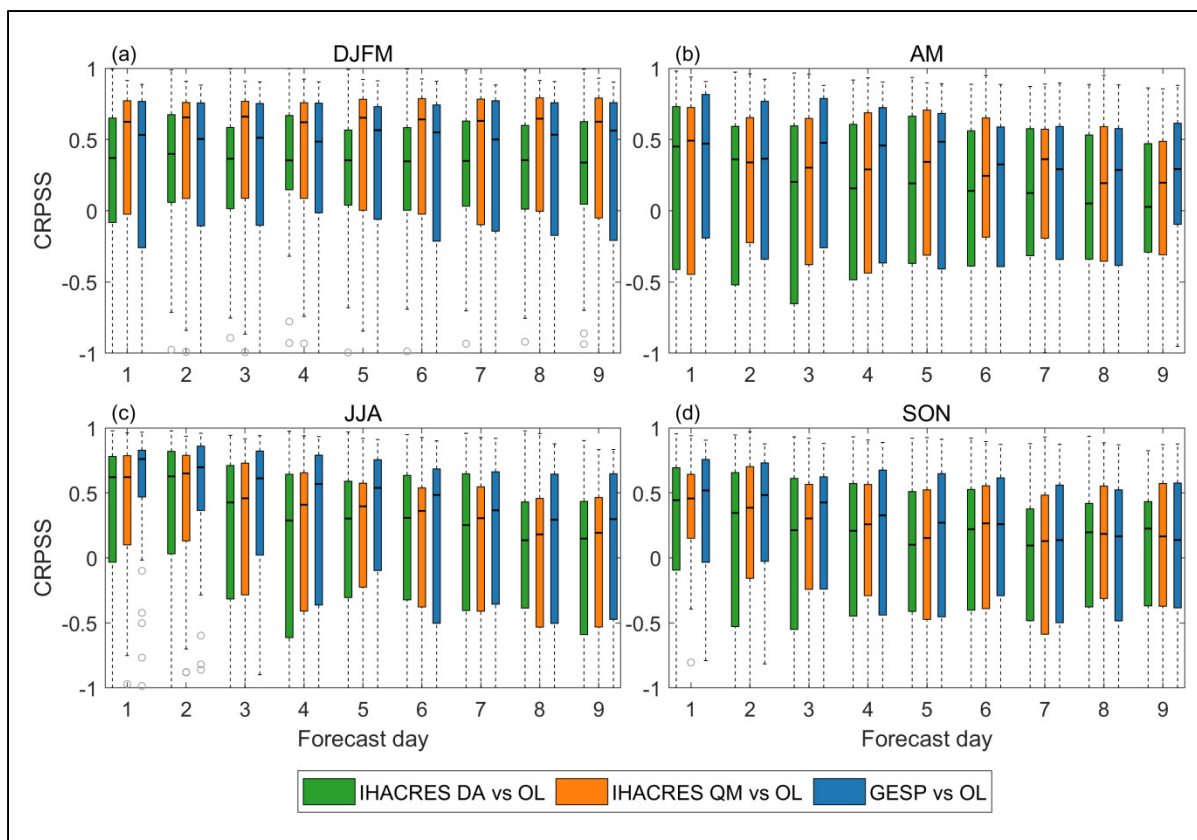


Figure-A II-3 Boîtes à moustaches des valeurs CRPSS des membres de l'ensemble hydrologique du modèle IHACRES et de l'ensemble multi-modèles (GESP) sur le bassin versant LM. Un CRPSS supérieur à zéro indique une amélioration de la précision globale des prévisions par rapport aux simulations de référence (OL) d'IHACRES. Les membres dont les valeurs du CRPSS sont inférieures à -1 sont considérés comme très mauvais et ne sont pas indiqués afin d'éviter de surcharger le graphique

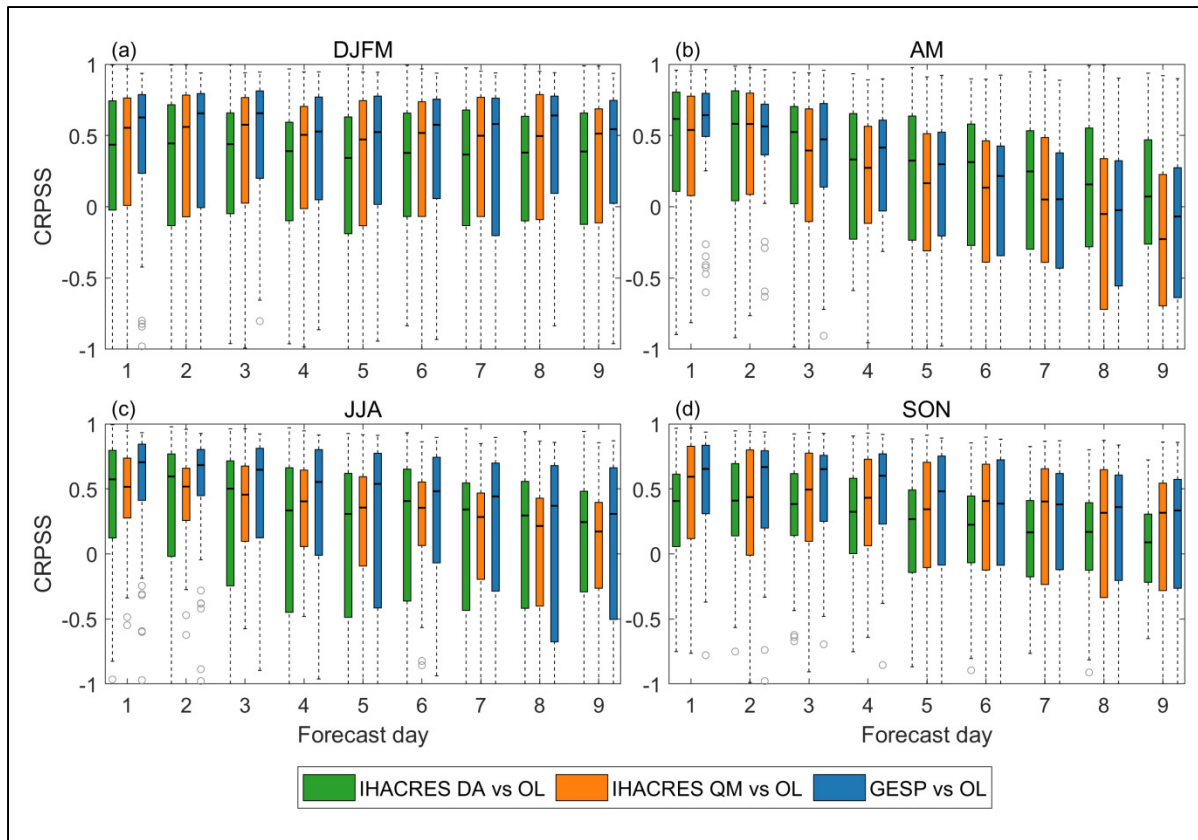


Figure-A II-4 Boîtes à moustaches des valeurs CRPSS des membres de l'ensemble hydrologique du modèle IHACRES et de l'ensemble multi-modèles (GESP) sur le bassin versant CS. Un CRPSS supérieur à zéro indique une amélioration de la précision globale des prévisions par rapport aux simulations de référence (OL) d'IHACRES. Les membres dont les valeurs du CRPSS sont inférieures à -1 sont considérés comme très mauvais et ne sont pas indiqués afin d'éviter de surcharger le graphique

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

- Abaza, M., Anctil, F., Fortin, V., & Turcotte, R. (2013). A Comparison of the Canadian Global and Regional Meteorological Ensemble Prediction Systems for Short-Term Hydrological Forecasting. *Monthly Weather Review*, 141(10), 3462-3476. doi: 10.1175/mwr-d-12-00206.1. Repéré à <https://doi.org/10.1175/MWR-D-12-00206.1>
- Abaza, M., Anctil, F., Fortin, V., & Turcotte, R. (2014). Sequential streamflow assimilation for short-term hydrological ensemble forecasting. *Journal of Hydrology*, 519, 2692-2706. doi: <https://doi.org/10.1016/j.jhydrol.2014.08.038>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169414006489>
- Ajami, N. K., Hornberger, G. M., & Sunding, D. L. (2008). Sustainable water resource management under hydrological uncertainty. *Water Resources Research*, 44(11). doi: 10.1029/2007wr006736. Repéré à <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007WR006736>
- Anctil, F., & Ramos, M.-H. (2019). Verification Metrics for Hydrological Ensemble Forecasts. Dans Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke & J. C. Schaake (Éds.), *Handbook of Hydrometeorological Ensemble Forecasting* (pp. 893-922). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-39925-1_3. Repéré à https://doi.org/10.1007/978-3-642-39925-1_3
- Andreadis, K. M., & Lettenmaier, D. P. (2006). Assimilating remotely sensed snow observations into a macroscale hydrology model. *Advances in Water Resources*, 29(6), 872-886. doi: <https://doi.org/10.1016/j.advwatres.2005.08.004>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0309170805002058>
- Arheimer, B., Lindström, G., & Olsson, J. (2011). A systematic review of sensitivities in the Swedish flood-forecasting system. *Atmospheric Research*, 100(2), 275-284. doi: <https://doi.org/10.1016/j.atmosres.2010.09.013>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0169809510002516>
- Arsenault, R., Brisette, F., Malo, J.-S., Minville, M., & Leconte, R. (2013). Structural and Non-Structural Climate Change Adaptation Strategies for the Péribonka Water Resource System. *Water Resources Management*, 27(7), 2075-2087. doi: 10.1007/s11269-013-0275-6. Repéré à <https://doi.org/10.1007/s11269-013-0275-6>
- Arsenault, R., Gatien, P., Renaud, B., Brisette, F., & Martel, J.-L. (2015). A comparative analysis of 9 multi-model averaging approaches in hydrological continuous streamflow simulation. *Journal of Hydrology*, 529, 754-767. doi: <https://doi.org/10.1016/j.jhydrol.2015.09.001>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169415006848>

- Arsenault, R., Latraverse, M., & Duchesne, T. (2016). An Efficient Method to Correct Under-Dispersion in Ensemble Streamflow Prediction of Inflow Volumes for Reservoir Optimization. *Water Resources Management*, 30(12), 4363-4380. doi: 10.1007/s11269-016-1425-4. Repéré à <https://doi.org/10.1007/s11269-016-1425-4>
- Arsenault, R., Poulin, A., Côté, P., & Brissette, F. (2014). Comparison of Stochastic Optimization Algorithms in Hydrological Model Calibration. *Journal of Hydrologic Engineering*, 19(7), 1374-1384. doi: doi:10.1061/(ASCE)HE.1943-5584.0000938. Repéré à <https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29HE.1943-5584.0000938>
- Bennett, J. C., Robertson, D. E., Shrestha, D. L., Wang, Q. J., Enever, D., Hapuarachchi, P., & Tuteja, N. K. (2014). A System for Continuous Hydrological Ensemble Forecasting (SCHEF) to lead times of 9days. *Journal of Hydrology*, 519, 2832-2846. doi: <https://doi.org/10.1016/j.jhydrol.2014.08.010>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169414006088>
- Bergeron, J. M., Trudel, M., & Leconte, R. (2016). Combined assimilation of streamflow and snow water equivalent for mid-term ensemble streamflow forecasts in snow-dominated regions. *Hydrol. Earth Syst. Sci.*, 20(10), 4375-4389. doi: 10.5194/hess-20-4375-2016. Repéré à <https://www.hydrol-earth-syst-sci.net/20/4375/2016/>
- Bergström, S., & Forsman, A. (1973). Development of a conceptual deterministic rainfall-runoff model. *Nordic Hydrol.*, 4, 147-170. doi: 10.2166/nh.1973.0012
- Beven, K. J., Kirkby, M. J., Schofield, N., & Tagg, A. F. (1984). Testing a physically-based flood forecasting model (TOPMODEL) for three U.K. catchments. *Journal of Hydrology*, 69(1), 119-143. doi: [https://doi.org/10.1016/0022-1694\(84\)90159-8](https://doi.org/10.1016/0022-1694(84)90159-8). Repéré à <http://www.sciencedirect.com/science/article/pii/0022169484901598>
- Boucher, M.-A., Roulin, E., & Fortin, V. (2019). Short-Range Ensemble Forecast Post-processing. In Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke & J. C. Schaake (Éds.), *Handbook of Hydrometeorological Ensemble Forecasting* (pp. 795-818). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Boucher, M.-A., Tremblay, D., Delorme, L., Perreault, L., & Anctil, F. (2012). Hydro-economic assessment of hydrological forecasting systems. *Journal of Hydrology*, 416-417, 133-144. doi: <https://doi.org/10.1016/j.jhydrol.2011.11.042>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169411008274>
- Bourdin, D. R., & Stull, R. B. (2013). Bias-corrected short-range Member-to-Member ensemble forecasts of reservoir inflow. *Journal of Hydrology*, 502, 77-88. doi: <https://doi.org/10.1016/j.jhydrol.2013.08.028>. Repéré à <http://www.sciencedirect.com/science/article/pii/S002216941300615X>

- Bourgin, F., Ramos, M. H., Thirel, G., & Andréassian, V. (2014). Investigating the interactions between data assimilation and post-processing in hydrological ensemble forecasting. *Journal of Hydrology*, 519, 2775-2784. doi: <https://doi.org/10.1016/j.jhydrol.2014.07.054>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169414005885>
- Brochero, D., Anctil, F., & Gagné, C. (2011). Simplifying a hydrological ensemble prediction system with a backward greedy selection of members – Part 1: Optimization criteria. *Hydrol. Earth Syst. Sci.*, 15(11), 3307-3325. doi: 10.5194/hess-15-3307-2011. Repéré à <https://www.hydrol-earth-syst-sci.net/15/3307/2011/>
- Chiew, F. H. S., Peel, M., C., & Western, A. W. (2002). Application and testing of the simple rainfall-runoff model SIMHYD. Dans *Mathematical Models of Small Watershed Hydrology and Applications* (pp. 335-367). Littleton: Water resources Publications. Repéré à <https://www.wrpllc.com/books/mmsw.html>
- Cloke, H. L., & Pappenberger, F. (2008). Evaluating forecasts of extreme events for hydrological applications: an approach for screening unfamiliar performance measures. *Meteorological Applications*, 15(1), 181-197. doi: 10.1002/met.58. Repéré à <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/met.58>
- Cloke, H. L., & Pappenberger, F. (2009). Ensemble flood forecasting: A review. *Journal of Hydrology*, 375(3), 613-626. doi: <https://doi.org/10.1016/j.jhydrol.2009.06.005>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169409003291>
- Council, N. R. (2006). *Completing the forecast: Characterizing and communicating uncertainty for better decisions using weather and climate forecasts*. National Academies Press.
- Cuo, L., Pagano, T. C., & Wang, Q. J. (2011). A Review of Quantitative Precipitation Forecasts and Their Use in Short- to Medium-Range Streamflow Forecasting. *Journal of Hydrometeorology*, 12(5), 713-728. doi: 10.1175/2011jhm1347.1. Repéré à <https://doi.org/10.1175/2011JHM1347.1>
- DeChant, C. M., & Moradkhani, H. (2011). Improving the characterization of initial condition for ensemble streamflow prediction using data assimilation. *Hydrol. Earth Syst. Sci.*, 15(11), 3399-3410. doi: 10.5194/hess-15-3399-2011. Repéré à <https://www.hydrol-earth-syst-sci.net/15/3399/2011/>
- DeChant, C. M., & Moradkhani, H. (2012). Examining the effectiveness and robustness of sequential data assimilation methods for quantification of uncertainty in hydrologic forecasting. *Water Resources Research*, 48(4). doi: 10.1029/2011wr011011. Repéré à <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011WR011011>

- Demargne, J., Wu, L., Regonda, S. K., Brown, J. D., Lee, H., He, M., . . . Zhu, Y. (2014). The Science of NOAA's Operational Hydrologic Ensemble Forecast Service. *Bulletin of the American Meteorological Society*, 95(1), 79-98. doi: 10.1175/bams-d-12-00081.1. Repéré à <https://doi.org/10.1175/BAMS-D-12-00081.1>
- Demeritt, D., Cloke, H., Pappenberger, F., Thielen, J., Bartholmes, J., & Ramos, M.-H. (2007). Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting. *Environmental Hazards*, 7(2), 115-127. doi: 10.1016/j.envhaz.2007.05.001. Repéré à <https://www.tandfonline.com/doi/abs/10.1016/j.envhaz.2007.05.001>
- Devineni, N., Sankarasubramanian, A., & Ghosh, S. (2008). Multimodel ensembles of streamflow forecasts: Role of predictor state in developing optimal combinations. *Water Resources Research*, 44(9). doi: 10.1029/2006wr005855. Repéré à <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006WR005855>
- Duan, Q., Sorooshian, S., & Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, 28(4), 1015-1031. doi: 10.1029/91wr02985. Repéré à <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/91WR02985>
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5), 10143-10162. doi: 10.1029/94jc00572. Repéré à <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/94JC00572>
- Evensen, G. (2003). The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4), 343-367. doi: 10.1007/s10236-003-0036-9. Repéré à <https://doi.org/10.1007/s10236-003-0036-9>
- Fan, F. M., Collischonn, W., Meller, A., & Botelho, L. C. M. (2014). Ensemble streamflow forecasting experiments in a tropical basin: The São Francisco river case study. *Journal of Hydrology*, 519, 2906-2919. doi: <https://doi.org/10.1016/j.jhydrol.2014.04.038>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169414003126>
- Fortin, V., & Turcotte, R. (2007). *Le modèle hydrologique mohyse*. Montreal, Canada: Département des sciences de la Terre et de l'atmosphère, Université du Québec à Montréal.
- Georgakakos, K. P., Graham, N. E., Modrick, T. M., Murphy, M. J., Shamir, E., Spencer, C. R., & Sperflage, J. A. (2014). Evaluation of real-time hydrometeorological ensemble prediction on hydrologic scales in Northern California. *Journal of Hydrology*, 519, 2978-3000. doi: <https://doi.org/10.1016/j.jhydrol.2014.05.032>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169414003965>

- Georgakakos, K. P., Seo, D.-J., Gupta, H., Schaake, J., & Butts, M. B. (2004). Towards the characterization of streamflow simulation uncertainty through multimodel ensembles. *Journal of Hydrology*, 298(1), 222-241. doi: <https://doi.org/10.1016/j.jhydrol.2004.03.037>. Repéré à <http://www.sciencedirect.com/science/article/pii/S002216940400246X>
- Girard, G., Morin, G., & Charbonneau, R. (1972). Modèle précipitations-débits à discrétisation spatiale. *Cahiers ORSTOM, série hydrologie*, 9(4), 35-52.
- Granger, C. W. J., & Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of Forecasting*, 3(2), 197-204. doi: 10.1002/for.3980030207. Repéré à <https://onlinelibrary.wiley.com/doi/abs/10.1002/for.3980030207>
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1), 80-91. doi: <https://doi.org/10.1016/j.jhydrol.2009.08.003>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169409004843>
- Hamill, T. M. (2001). Interpretation of Rank Histograms for Verifying Ensemble Forecasts. *Monthly Weather Review*, 129(3), 550-560. doi: 10.1175/1520-0493(2001)129. Repéré à [https://doi.org/10.1175/1520-0493\(2001\)129](https://doi.org/10.1175/1520-0493(2001)129)
- Hashino, T., Bradley, A. A., & Schwartz, S. S. (2007). Evaluation of bias-correction methods for ensemble streamflow volume forecasts. *Hydrol. Earth Syst. Sci.*, 11(2), 939-950. doi: 10.5194/hess-11-939-2007. Repéré à <https://www.hydrol-earth-syst-sci.net/11/939/2007/>
- Hemri, S. (2019). Multi-model Combination and Seamless Prediction. Dans Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke & J. C. Schaake (Éds.), *Handbook of Hydrometeorological Ensemble Forecasting* (pp. 285-307). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-39925-1_19. Repéré à https://doi.org/10.1007/978-3-642-39925-1_19
- Hemri, S., Fundel, F., & Zappa, M. (2013). Simultaneous calibration of ensemble river flow predictions over an entire range of lead times. *Water Resources Research*, 49(10), 6744-6755. doi: 10.1002/wrcr.20542. Repéré à <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/wrcr.20542>
- Hemri, S., & Klein, B. (2017). Analog-Based Postprocessing of Navigation-Related Hydrological Ensemble Forecasts. *Water Resources Research*, 53(11), 9059-9077. doi: 10.1002/2017wr020684. Repéré à <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017WR020684>

- Hopson, T. M., & Webster, P. J. (2010). A 1–10-Day Ensemble Forecasting Scheme for the Major River Basins of Bangladesh: Forecasting Severe Floods of 2003–07*. *Journal of Hydrometeorology*, 11(3), 618-641. doi: 10.1175/2009jhm1006.1. Repéré à <https://doi.org/10.1175/2009JHM1006.1>
- Hopson, T. M., Wood, A., & Weerts, A. H. (2019). Motivation and Overview of Hydrological Ensemble Post-processing. Dans Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke & J. C. Schaake (Éds.), *Handbook of Hydrometeorological Ensemble Forecasting* (pp. 783-793). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-39925-1_36. Repéré à https://doi.org/10.1007/978-3-642-39925-1_36
- IPCC, P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.- O. Pörtner, . . . Malley, J. (2019). *Summary for Policymakers: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*. The Intergovernmental Panel on Climate Change. Repéré à <https://www.ipcc.ch/srccl/chapter/summary-for-policymakers/>
- IPCC, Team, C. W., R.K., P., & L.A., M. (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland: The Intergovernmental Panel on Climate Change. Repéré à <https://www.ipcc.ch/srccl/chapter/summary-for-policymakers/>
- Jakeman, A. J., Littlewood, I. G., & Whitehead, P. G. (1990). Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. *Journal of Hydrology*, 117(1), 275-300. doi: [https://doi.org/10.1016/0022-1694\(90\)90097-H](https://doi.org/10.1016/0022-1694(90)90097-H). Repéré à <http://www.sciencedirect.com/science/article/pii/002216949090097H>
- Jiménez, K. Q., Collischonn, W., & Paiva, R. C. D. d. (2019). Data assimilation using the ensemble Kalman filter in a distributed hydrological model on the Tocantins River, Brasil. *RBRH*, 24. Repéré à http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2318-03312019000100214&nrm=iso
- Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrol. Earth Syst. Sci.*, 23(10), 4323-4331. doi: 10.5194/hess-23-4323-2019. Repéré à <https://hess.copernicus.org/articles/23/4323/2019/>
- Krzysztofowicz, R. (2001). The case for probabilistic forecasting in hydrology. *Journal of Hydrology*, 249(1), 2-9. doi: [https://doi.org/10.1016/S0022-1694\(01\)00420-6](https://doi.org/10.1016/S0022-1694(01)00420-6). Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169401004206>

- Krzysztofowicz, R., & Maranzano, C. J. (2004). Hydrologic uncertainty processor for probabilistic stage transition forecasting. *Journal of Hydrology*, 293(1), 57-73. doi: <https://doi.org/10.1016/j.jhydrol.2004.01.003>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169404000137>
- Leandro, J., Gander, A., Beg, M. N. A., Bhola, P., Konnerth, I., Willems, W., . . . Disse, M. (2019). Forecasting upper and lower uncertainty bands of river flood discharges with high predictive skill. *Journal of Hydrology*, 576, 749-763. doi: <https://doi.org/10.1016/j.jhydrol.2019.06.052>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169419305992>
- Leutbecher, M., Lock, S.-J., Ollinaho, P., Lang, S. T. K., Balsamo, G., Bechtold, P., . . . Weisheimer, A. (2017). Stochastic representations of model uncertainties at ECMWF: state of the art and future vision. *Quarterly Journal of the Royal Meteorological Society*, 143(707), 2315-2339. doi: 10.1002/qj.3094. Repéré à <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3094>
- Li, W., Duan, Q., Miao, C., Ye, A., Gong, W., & Di, Z. (2017). A review on statistical postprocessing methods for hydrometeorological ensemble forecasting. *WIREs Water*, 4(6), e1246. doi: 10.1002/wat2.1246. Repéré à <https://onlinelibrary.wiley.com/doi/abs/10.1002/wat2.1246>
- Liu, Y., Brown, J., Demargne, J., & Seo, D.-J. (2011). A wavelet-based approach to assessing timing errors in hydrologic predictions. *Journal of Hydrology*, 397(3), 210-224. doi: <https://doi.org/10.1016/j.jhydrol.2010.11.040>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169410007365>
- Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H. J., Kumar, S., Moradkhani, H., . . . Restrepo, P. (2012). Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. *Hydrol. Earth Syst. Sci.*, 16(10), 3863-3887. doi: 10.5194/hess-16-3863-2012. Repéré à <https://www.hydrol-earth-syst-sci.net/16/3863/2012/>
- Maxwell, D. H., Jackson, B. M., & McGregor, J. (2018). Constraining the ensemble Kalman filter for improved streamflow forecasting. *Journal of Hydrology*, 560, 127-140. doi: <https://doi.org/10.1016/j.jhydrol.2018.03.015>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169418301781>
- Mazrooei, A., Sinha, T., Sankarasubramanian, A., Kumar, S., & Peters-Lidard, C. D. (2015). Decomposition of sources of errors in seasonal streamflow forecasting over the U.S. Sunbelt. *Journal of Geophysical Research: Atmospheres*, 120(23), 11,809-811,825. doi: 10.1002/2015jd023687. Repéré à <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD023687>

- McCollor, D., & Stull, R. (2008). Hydrometeorological Short-Range Ensemble Forecasts in Complex Terrain. Part II: Economic Evaluation. *Weather and Forecasting*, 23(4), 557-574. doi: 10.1175/2007waf2007064.1. Repéré à <https://doi.org/10.1175/2007WAF2007064.1>
- Montanari, A., & Grossi, G. (2008). Estimating the uncertainty of hydrological forecasts: A statistical approach. *Water Resources Research*, 44(12). doi: 10.1029/2008wr006897. Repéré à <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008WR006897>
- Moradkhani, H., Sorooshian, S., Gupta, H. V., & Houser, P. R. (2005). Dual state–parameter estimation of hydrological models using ensemble Kalman filter. *Advances in Water Resources*, 28(2), 135-147. doi: <https://doi.org/10.1016/j.advwatres.2004.09.002>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0309170804001605>
- Muhammad, A., Stadnyk, T. A., Unduche, F., & Coulibaly, P. (2018). Multi-Model Approaches for Improving Seasonal Ensemble Streamflow Prediction Scheme with Various Statistical Post-Processing Techniques in the Canadian Prairie Region. *Water*, 10(11), 1604. Repéré à <https://www.mdpi.com/2073-4441/10/11/1604>
- Muthusamy, M., Godiksen, P. N., & Madsen, H. (2016). Comparison of Different Configurations of Quantile Regression in Estimating Predictive Hydrological Uncertainty. *Procedia Engineering*, 154, 513-520. doi: <https://doi.org/10.1016/j.proeng.2016.07.546>. Repéré à <http://www.sciencedirect.com/science/article/pii/S187770581631935X>
- Neuman, S. P. (2003). Maximum likelihood Bayesian averaging of uncertain model predictions. *Stochastic Environmental Research and Risk Assessment*, 17(5), 291-305. doi: 10.1007/s00477-003-0151-7. Repéré à <https://doi.org/10.1007/s00477-003-0151-7>
- Noh, S. J., Weerts, A. H., Rakovec, O., Lee, H., & Seo, D.-J. (2019). Assimilation of Streamflow Observations. Dans Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke & J. C. Schaake (Éds.), *Handbook of Hydrometeorological Ensemble Forecasting* (pp. 745-780). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-39925-1_33. Repéré à https://doi.org/10.1007/978-3-642-39925-1_33
- Olsson, J., & Lindström, G. (2008). Evaluation and calibration of operational hydrological ensemble forecasts in Sweden. *Journal of Hydrology*, 350(1), 14-24. doi: <https://doi.org/10.1016/j.jhydrol.2007.11.010>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169407006889>
- Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., & Loumagne, C. (2005). Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–

- runoff modelling. *Journal of Hydrology*, 303(1), 290-306. doi: <https://doi.org/10.1016/j.jhydrol.2004.08.026>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169404004056>
- Pagano, T. C., Shrestha, D. L., Wang, Q. J., Robertson, D., & Hapuarachchi, P. (2013). Ensemble dressing for hydrological applications. *Hydrological Processes*, 27(1), 106-116. doi: 10.1002/hyp.9313. Repéré à <https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.9313>
- Pappenberger, F., Thielen, J., & Del Medico, M. (2011). The impact of weather forecast improvements on large scale hydrology: analysing a decade of forecasts of the European Flood Alert System. *Hydrological Processes*, 25(7), 1091-1113. doi: 10.1002/hyp.7772. Repéré à <https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.7772>
- Perrin, C., Michel, C., & Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, 279(1), 275-289. doi: [https://doi.org/10.1016/S0022-1694\(03\)00225-7](https://doi.org/10.1016/S0022-1694(03)00225-7). Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169403002257>
- Poissant, D., Arsenault, R., & Brissette, F. (2017). Impact of parameter set dimensionality and calibration procedures on streamflow prediction at ungauged catchments. *Journal of Hydrology: Regional Studies*, 12, 220-237. doi: <https://doi.org/10.1016/j.ejrh.2017.05.005>. Repéré à <http://www.sciencedirect.com/science/article/pii/S221458181730191X>
- R. Lanthier, M. (2018). *Analyse des approches multi-modèle en prévision hydrologique* (Master, École de Technologie Supérieure, Université du Québec, Montréal).
- Ramos, M.-H., Mathevet, T., Thielen, J., & Pappenberger, F. (2010). Communicating uncertainty in hydro-meteorological forecasts: mission impossible? *Meteorological Applications*, 17(2), 223-235. doi: 10.1002/met.202. Repéré à <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/met.202>
- Reggiani, P., & Weerts, A. H. (2008). A Bayesian approach to decision-making under uncertainty: An application to real-time forecasting in the river Rhine. *Journal of Hydrology*, 356(1), 56-69. doi: <https://doi.org/10.1016/j.jhydrol.2008.03.027>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169408001698>
- Reichle, R. H., McLaughlin, D. B., & Entekhabi, D. (2002). Hydrologic Data Assimilation with the Ensemble Kalman Filter. *Monthly Weather Review*, 130(1), 103-114. doi: 10.1175/1520-0493(2002)130. Repéré à [https://doi.org/10.1175/1520-0493\(2002\)130](https://doi.org/10.1175/1520-0493(2002)130)

- Roulin, E. (2007). Skill and relative economic value of medium-range hydrological ensemble predictions. *Hydrol. Earth Syst. Sci.*, 11(2), 725-737. doi: 10.5194/hess-11-725-2007. Repéré à <https://www.hydrol-earth-syst-sci.net/11/725/2007/>
- Roulin, E., & Vannitsem, S. (2015). Post-processing of medium-range probabilistic hydrological forecasting: impact of forcing, initial conditions and model errors. *Hydrological Processes*, 29(6), 1434-1449. doi: 10.1002/hyp.10259. Repéré à <https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.10259>
- Schaake, J. C., Demargne, J., Hartman, R., Mullusky, M., Welles, E., Wu, L., . . . Seo, D. J. (2007). Precipitation and temperature ensemble forecasts from single-value forecasts. *Hydrol. Earth Syst. Sci. Discuss.*, 2007, 655-717. doi: 10.5194/hessd-4-655-2007. Repéré à <https://www.hydrol-earth-syst-sci-discuss.net/4/655/2007/>
- Schaake, J. C., Hamill, T. M., Buizza, R., & Clark, M. (2007). HEPEX: The Hydrological Ensemble Prediction Experiment. *Bulletin of the American Meteorological Society*, 88(10), 1541-1548. doi: 10.1175/bams-88-10-1541. Repéré à <https://doi.org/10.1175/BAMS-88-10-1541>
- Schwanenberg, D., Fan, F. M., Naumann, S., Kuwajima, J. I., Montero, R. A., & Assis dos Reis, A. (2015). Short-Term Reservoir Optimization for Flood Mitigation under Meteorological and Hydrological Forecast Uncertainty. *Water Resources Management*, 29(5), 1635-1651. doi: 10.1007/s11269-014-0899-1. Repéré à <https://doi.org/10.1007/s11269-014-0899-1>
- Seo, D.-J., Liu, Y., Moradkhani, H., & Weerts, A. (2014). Ensemble prediction and data assimilation for operational hydrology. *Journal of Hydrology*, 519, 2661-2662. doi: <https://doi.org/10.1016/j.jhydrol.2014.11.035>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169414009470>
- Thibault, A., & Anctil, F. (2015). On the difficulty to optimally implement the Ensemble Kalman filter: An experiment based on many hydrological models and catchments. *Journal of Hydrology*, 529, 1147-1160. doi: <https://doi.org/10.1016/j.jhydrol.2015.09.036>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169415007192>
- Thibault, A., Anctil, F., & Boucher, M. A. (2016). Accounting for three sources of uncertainty in ensemble hydrological forecasting. *Hydrol. Earth Syst. Sci.*, 20(5), 1809-1825. doi: 10.5194/hess-20-1809-2016. Repéré à <https://www.hydrol-earth-syst-sci.net/20/1809/2016/>
- Thibault, A., Seiller, G., Poncelet, C., & Anctil, F. (2019, 2019/05/15/). *HOOPLA: un outil multifonction pour la modélisation hydrologique* présentée à Colloque AQT/RHQ - La télédétection de l'eau, dans tous leurs états, Sherbrooke, Canada. Repéré à <https://aqtrhq2019.sciencesconf.org/>

- Valéry, A., Andréassian, V., & Perrin, C. (2014). 'As simple as possible but not simpler': What is useful in a temperature-based snow-accounting routine? Part 2 – Sensitivity analysis of the Cemaneige snow accounting routine on 380 catchments. *Journal of Hydrology*, 517, 1176-1187. doi: <https://doi.org/10.1016/j.jhydrol.2014.04.058>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169414003321>
- Velázquez, J. A., Anctil, F., Ramos, M. H., & Perrin, C. (2011). Can a multi-model approach improve hydrological ensemble forecasting? A study on 29 French catchments using 16 hydrological model structures. *Adv. Geosci.*, 29, 33-42. doi: 10.5194/adgeo-29-33-2011. Repéré à <https://adgeo.copernicus.org/articles/29/33/2011/>
- Verkade, J. S., Brown, J. D., Reggiani, P., & Weerts, A. H. (2013). Post-processing ECMWF precipitation and temperature ensemble reforecasts for operational hydrologic forecasting at various spatial scales. *Journal of Hydrology*, 501, 73-91. doi: <https://doi.org/10.1016/j.jhydrol.2013.07.039>. Repéré à <http://www.sciencedirect.com/science/article/pii/S0022169413005660>
- Voisin, N., Pappenberger, F., Lettenmaier, D. P., Buizza, R., & Schaake, J. C. (2011). Application of a Medium-Range Global Hydrologic Probabilistic Forecast Scheme to the Ohio River Basin. *Weather and Forecasting*, 26(4), 425-446. doi: 10.1175/waf-d-10-05032.1. Repéré à <https://doi.org/10.1175/WAF-D-10-05032.1>
- Wagener, T., Boyle, D. P., Lees, M. J., Wheeler, H. S., Gupta, H. V., & Sorooshian, S. (2001). A framework for development and application of hydrological models. *Hydrology and Earth System Sciences Discussions*, 5(1), 13-26. Repéré à <https://hal.archives-ouvertes.fr/hal-00304565>
- Wetterhall, F., Pappenberger, F., Alfieri, L., Cloke, H., & Thielen, J. (2014, May 01, 2014). *Forecaster priorities for improving probabilistic flood forecasts* présentée à EGU General Assembly Conference Abstracts. Repéré à <https://ui.adsabs.harvard.edu/abs/2014EGUGA..16.8159W>
- Wetterhall, F., Pappenberger, F., Alfieri, L., Cloke, H. L., Thielen-del Pozo, J., Balabanova, S., . . . Holubecka, M. (2013). HESS Opinions "Forecaster priorities for improving probabilistic flood forecasts". *Hydrol. Earth Syst. Sci.*, 17(11), 4389-4399. doi: 10.5194/hess-17-4389-2013. Repéré à <https://www.hydrol-earth-syst-sci.net/17/4389/2013/>
- Wetterhall, F., & Smith, P. (2019). Hydrological Challenges in Meteorological Post-processing. Dans Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke & J. C. Schaake (Éds.), *Handbook of Hydrometeorological Ensemble Forecasting* (pp. 239-253). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-39925-1_15. Repéré à https://doi.org/10.1007/978-3-642-39925-1_15

- Wood, A. W., & Schaake, J. C. (2008). Correcting Errors in Streamflow Forecast Ensemble Mean and Spread. *Journal of Hydrometeorology*, 9(1), 132-148. doi: 10.1175/2007jhm862.1. Repéré à <https://doi.org/10.1175/2007JHM862.1>
- Zappa, M., Fundel, F., & Jaun, S. (2013). A 'Peak-Box' approach for supporting interpretation and verification of operational ensemble peak-flow forecasts. *Hydrological Processes*, 27(1), 117-131. doi: 10.1002/hyp.9521. Repéré à <https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.9521>
- Zhao, T., Bennett, J. C., Wang, Q. J., Schepen, A., Wood, A. W., Robertson, D. E., & Ramos, M.-H. (2017). How Suitable is Quantile Mapping For Postprocessing GCM Precipitation Forecasts? *Journal of Climate*, 30(9), 3185-3196. doi: 10.1175/jcli-d-16-0652.1. Repéré à <https://doi.org/10.1175/JCLI-D-16-0652.1>