

Université du Québec
Institut National de la Recherche Scientifique
Centre Eau Terre Environnement

Nouvelles approches de calibration automatique appliquées à un modèle hydrologique conceptuel semi-distribué

Par

Samah Larabi

Thèse présentée pour l'obtention du grade de Philosophiae doctor (Ph.D.)
en sciences de l'eau

Jury d'évaluation

Président du jury et Examineur interne	Taha B.M.J. Ouarda INRS-ETE
Examineur externe	Wayne Jenkinson NRC
Examineur externe	Robert Leconte Université de Sherbrooke
Directeur de recherche	André St-Hilaire INRS-ETE
Codirecteurs de recherche	Fateh Chebana INRS-ETE Marco Latraverse Rio Tinto

REMERCIEMENTS

La réalisation de ce projet a été possible grâce à plusieurs personnes à qui je voudrais témoigner toute ma reconnaissance.

Tout d’abord, je tiens à adresser ma gratitude à mon directeur de recherche, André St-Hilaire, pour cette opportunité de toucher de près et de développer le sens de la recherche ainsi que toutes les opportunités qu’il m’a offertes au cours de mon doctorat. Son appui, ses précieux conseils et encouragements m’ont aidée à avancer et améliorer ce projet.

Je remercie aussi mon codirecteur Fateh Chebana pour ses judicieux conseils et son œil critique qui ont contribué à alimenter ma réflexion.

Je remercie mon codirecteur Marco Latraverse pour son aide technique et ses éclaircissements.

Mes remerciements vont aussi à Yves Trambly, mon tuteur de stage de fin d’études au cycle d’ingénieur, pour m’avoir encouragée d’aller plus loin dans mon parcours académique.

Je remercie mes collègues d’avoir partagé les données et enrichir ce projet.

Un profond merci au Conseil de Recherches en Sciences Naturelles et en Génie (CRNSG) et Rio Tinto pour avoir financé ma thèse.

Enfin, merci à mes parents qui m’ont aidée et soutenue tout au long de ce projet. Je leur suis redevable d’une éducation dont je suis fière.

RÉSUMÉ

Les modèles hydrologiques sont des outils indispensables et largement utilisés pour la prévision des variables hydrologiques, gestion de réservoirs, prise de décision ou encore pour améliorer la compréhension des processus régissant le cycle hydrologique d'un bassin versant. Ces outils sont des ensembles d'équations mathématiques qui décrivent et représentent de manière relativement simple les processus hydrologiques occurrents dans un bassin versant. Chaque modèle inclut plusieurs paramètres pour décrire l'état du système en question. Certains de ces paramètres sont conceptuels ou indisponibles par mesure directe et nécessitent une estimation à partir de données historiques avant de mettre en place le modèle. Ce processus d'estimation des paramètres est appelé calibration. Cette tâche est complexe et présente beaucoup de difficultés à cause des incertitudes liées à la structure du modèle, aux intrants du modèle et aux données historiques de débits. La calibration du modèle impacte sa fiabilité, sa précision et sa performance à reproduire les mécanismes du système étudié ainsi que ses prédictions des variables hydrologiques. Étant donné ces impacts, il y a une nécessité de définir une stratégie de calibration efficace et performante. C'est dans cet esprit que s'inscrit l'objectif de cette thèse qui consiste à améliorer et proposer de nouvelles approches pour la calibration automatique des modèles hydrologiques. Les méthodes développées sont appliquées au modèle hydrologique conceptuel semi-distribué CEQUEAU et pourraient être appliquées à n'importe quel autre modèle. Trois bassins versants au Canada sont considérés à savoir le bassin du lac St-Jean situé au Québec, la rivière du Nechako en Colombie-Britannique et le ruisseau Catamaran au Nouveau-Brunswick.

Les travaux de recherches menés visent à atteindre cinq objectifs spécifiques reliés à différents axes de la calibration: 1) définir une procédure d'optimisation robuste, 2) définir une mesure de performance fiable incluant toute l'information contenue dans les données historiques, 3) comparer l'efficacité de cette mesure contre celle utilisées dans des stratégies de calibration avancées, 4) définir une stratégie de calibration pour améliorer le réalisme et la fiabilité des estimations des réponses du système (ex. les débits), 5) définir une stratégie de calibration multisite afin d'améliorer la performance du modèle à l'échelle du bassin.

Le premier objectif s'intéresse à la réduction du risque d'équifinalité et l'estimation de paramètres conceptuellement réalistes. Une procédure d'estimation de paramètres combinant des simulations Monte Carlo avec fonctions de profondeur et un algorithme d'optimisation globale

appelé ‘Tabu Search’ est proposée. La simulation Monte Carlo est destinée au raffinement de l’espace des paramètres pour trouver la région d’attraction de cet espace. L’utilisation d’une fonction de profondeur avec la simulation Monte Carlo permet un échantillonnage stratégique et restreint dans la région contenant des combinaisons de paramètres performantes. Par conséquent, cette étape fournit un espace optimal qui est exploré davantage par le ‘Tabu Search’ dans le but de trouver un seul optimum global. La procédure permet une réduction importante de l’espace des paramètres en particulier les paramètres sensibles (jusqu’à 90%). L’algorithme d’optimisation globale converge vers des solutions performantes et plus réalistes après optimisation de l’espace. En effet, l’ajustement des paramètres dans l’espace borné selon des valeurs physiquement plausibles de paramètres mène à des jeux de paramètres simulant des comportements hydrologiques irréalistes (par exemple coefficient d’infiltration entre les couches de sols conceptualisés par le modèle fixé à 0 i.e. aucune infiltration) bien qu’ils soient performants (reproduisent adéquatement le débit). La méthodologie a été appliquée avec une validation croisée (‘split sample’) et une validation croisée différentielle (‘Differential split-sample’). Dans les deux cas, le modèle calibré est robuste et performant.

Le deuxième objectif a pour but de définir une méthode de calibration prenant en compte toute l’information sur la variabilité, forme et occurrence des états hydrologiques (étiage, crue, récession, etc.) plutôt que de mettre uniquement l’accent sur la crue printanière. À cet effet, l’analyse de données fonctionnelles (FDA) est intégrée dans le processus de calibration et validation du modèle hydrologique. Cet outil mathématique permet d’analyser des données observées sur un intervalle régulier (espace, temps,...) comme des fonctions continues. Dans notre cas, cet outil permet d’analyser les hydrogrammes comme une seule entité ou observation au lieu de 365 observations journalières. Dans ce contexte, la fonction objectif est une distance entre les hydrogrammes simulés et observés. La procédure de calibration développée lors de la première partie de la thèse est utilisée et deux fonctions objectives adaptées à chaque étape ont été définies. L’utilisation de la FDA fournit la même qualité d’estimation de débits de crue qu’une calibration avec le critère Nash-Sutcliffe Efficiency NSE (ex. $NSE = 0.9$ en période de calibration et validation au bassin du Lac St-Jean). Toutefois, elle permet d’améliorer l’estimation des débits d’étiage, l’estimation de certaines signatures hydrologiques ainsi que de considérer toute la variabilité de l’hydrogramme. Évaluer la performance du modèle à simuler les signatures hydrologiques qui caractérisent l’écoulement dans le bassin versant (distribution, fréquence et

variabilité de débits de crue et d'étiages) permet de détecter la composante du modèle qui pourrait être déficiente.

Le troisième objectif explore une formulation différente de mesure de performance utilisant la FDA et compare l'efficacité de la calibration avec celle-ci avec celle produite en appliquant les autres critères conventionnellement utilisés dans la littérature (calibration avec une seule fonction objectif avec le critère NSE, calibration multiobjective avec le critère Kling Gupta Efficiency (KGE) et calibration en utilisant les signatures hydrologiques). Dans cet objectif, la fonction objectif basée sur la FDA est le ratio entre le volume intégré (aire sous la courbe de l'hydrogramme) annuel simulé et observé. La procédure combinant simulation Monte Carlo et fonction de profondeur développée lors du premier objectif est utilisée pour comparer la performance des jeux de paramètres engendrés par les différentes calibrations considérées. Les résultats montrent que la calibration basée sur la FDA permet une meilleure estimation de l'hydrogramme en entier. Elle permet une meilleure représentation des processus hydrologiques du bassin (i.e. reproduire adéquatement plusieurs ou toutes les signatures hydrologiques considérées) par rapport aux autres méthodes. Les calibrations conventionnelles assurent une bonne performance du modèle pour simuler la crue avec une très faible performance sur les débits d'étiage. Néanmoins, la calibration avec le NSE est la meilleure méthode conventionnelle pour capter l'occurrence et le volume de la crue printanière dans un bassin où la crue est dominée par la fonte de neige. Quelle que soit la méthode de calage, l'incertitude sur la pointe et le volume de la crue printanière est significative.

Le quatrième objectif vise à améliorer la qualité de la calibration du modèle pour mieux reproduire les observations du bassin versant (dans notre cas, les débits) ainsi que de mieux capturer les changements des états internes et processus hydrologiques. La stratégie de calibration développée est une procédure de calibration multicritère par étape et par processus basée sur la procédure de calibration définie lors de la première étape et la FDA. Pour reproduire le volume d'eau observé dans le bassin versant, le modèle a tendance à compenser entre les volumes d'eau provenant de ses différentes composantes simulant les processus internes (ex. fonte de neige, évaporation, apports d'eau souterraine). Pour réduire cet effet de compensation parfois erronée, les composantes du modèle régissant chaque processus hydrologique sont calibrées séparément. L'ajustement de leurs paramètres est fait selon un critère spécifique défini à partir d'une partie de l'hydrogramme qui est directement reliée au processus hydrologique en

question. Cette stratégie de calibration a été appliquée uniquement au bassin versant du lac St-Jean où l'on dispose de plusieurs stations hydrométriques. L'estimation de l'hydrogramme en entier (débits de crue et d'étiage) est améliorée avec une valeur de NSE de 0.86 et un biais d'étiage hivernal de 16% comparativement à une calibration classique sur les observations de débits uniquement (NSE de 0.71 et biais d'étiage hivernal de 21%). La validation du modèle sur différentes stations a révélé la bonne performance du modèle sur des sous-bassins ayant des caractéristiques similaires aux stations utilisées pour la calibration avec une valeur de NSE supérieure à 0.7. Globalement, le NSE moyen sur toutes les stations du bassin du lac St-Jean n'ayant pas servi à la calibration est de 0.61.

Le cinquième objectif consiste à améliorer la performance du modèle à une échelle spatiale plus importante que celle d'une seule station, c'est-à-dire à l'échelle du bassin. En effet, le but d'un modèle hydrologique distribué/semi-distribué, contrairement à un modèle global, est de simuler le débit sur n'importe quel point du bassin. Généralement, les modèles hydrologiques sont calibrés à l'aide d'observations de débits à l'exutoire uniquement. Le calage du modèle en un seul point n'est pas une garantie d'un modèle performant à l'échelle d'un grand bassin versant. Dans le cadre de cette thèse, une calibration locale réfère à une calibration à l'échelle d'une station et une calibration globale réfère à une calibration à l'échelle du bassin versant. Deux procédures de calibration multisite sont proposées. La première consiste en une calibration globale du modèle sur toutes les stations simultanément. La deuxième consiste en une calibration locale par station suivie d'une calibration globale en évaluant la performance du modèle sur toutes les stations simultanément. La calibration locale permet d'explorer et analyser la variabilité spatiale des paramètres à l'échelle du bassin. Les paramètres qui présentent une variabilité spatiale négligeable sont fixés à la valeur moyenne des valeurs résultantes du calage local sur les différentes stations hydrométriques. La deuxième étape, calage global, met l'accent sur les paramètres qui varient le plus à l'échelle du bassin. Les deux procédures sont appliquées avec une validation croisée. L'application des procédures sur deux périodes différentes permet d'explorer la variabilité temporelle des paramètres. Les résultats des calibrations locales et globales montrent que certains paramètres de la fonction de production du modèle CEQUEAU, tel que les hauteurs des réservoirs conceptualisés et le coefficient de vidange basse du réservoir Sol (CVSB), ont une variabilité spatiale et temporelle négligeable sur le bassin versant du lac St-Jean. Ces paramètres peuvent être fixés sans calage global et sans

perte de performance. Les paramètres de fonte et les coefficients de vidange : haute du réservoir Nappe (CVNH), haute du réservoir Sol (CVSI) et du réservoir Lac et Marrais (CVMAR) sont variables spatialement et dépendent aussi de la période de calibration. Les deux procédures mènent à des performances globalement similaires (ex NSE de 0.86 à la station de l'exutoire du bassin du Lac St-Jean, NSE de 0.83 à la station ASHA située plus en amont dans le bassin du lac St-Jean).

Globalement les travaux de cette thèse proposent une combinaison de méthodes incluant algorithmes, fonctions objectifs et stratégies de calibration qui visent à définir des paramètres d'un modèle réaliste (qui reproduit les processus hydrologiques internes avec des valeurs de paramètres conceptuellement réalistes) et performant dans l'ensemble (autant sur les débits d'étiages que la crue). De plus, les travaux ont permis une exploration de la variabilité des paramètres du modèle CEQUEAU dans le bassin du lac St-Jean où l'on dispose de plusieurs stations hydrométriques et sur lequel le modèle est utilisé pour faire de la prévision opérationnelle.

Mots clés : Analyse de données fonctionnelles, Calibration, CEQUEAU, Fonction de profondeur, Simulation Monte Carlo, Tabu Search.

ABSTRACT

Hydrological models are important tools widely used for flow forecasting, reservoir management, or to enhance the knowledge about processes of the hydrological cycle within a watershed. These tools are an ensemble of mathematical equations to describe and represent relatively in a simple way the hydrological processes occurring within a watershed. Each model has multiple of parameters to describe the states of the considered system. Some of these parameters are conceptual and/or unavailable by direct measurement and need to be estimated using historical data before setting up the model. This process of estimating model parameters is called calibration. This task can be complex and difficult because of the uncertainty associated with the model structure, inputs of the model and the historical data. Model calibration impacts the accuracy, the precision and the performance of the model to reproduce the mechanism of the system and its response. Therefore, an efficient and performant calibration strategy must be defined carefully. In this context, the objective of this thesis is to improve and develop novel approaches for automatic calibration. The developed methods are applied to the CEQUEAU conceptual semi-distributed hydrological model and can be applied to any model. Three Canadian watersheds are considered, namely the Lac St-Jean drainage basin located in the province of Québec, Nechako River in British-Columbia and the Catamaran Brook in New-Brunswick.

The objectives of the research work conducted for this thesis are related to different themes, all related to model calibration: 1) defining a robust optimization procedure for parameter estimation, 2) defining a reliable method to asses model performance reflecting all information contained in the historical data, 3) Comparing this method to the conventional methods used with uncertainty analysis, 4) defining a calibration strategy to improve the realism and the reliability of the response system estimation, and 5) defining a multi-site calibration strategy in order to improve the model performance at the basin scale.

The first objective of the thesis is to reduce the equifinality risk and estimate realistic conceptual model parameters. A procedure combining Monte Carlo simulations coupled with depth functions and a global optimization algorithm called ‘Tabu Search’ is proposed for estimating model parameters. The Monte Carlo simulations aim at refining the parameter space in order to find the attraction region of the space. The use of a depth function with the Monte Carlo

simulations allows for a strategic sampling in the region of the space including performing sets of parameters. Consequently, this step defines an optimal parameter space that is explored deeply by the Tabu Search in order to find a global optimum. This procedure considerably reduces the parameter space particularly for the most sensitive parameters (up to 90%). The global optimization algorithm converges to a more realistic and performant solution when the space parameter is optimized first. Adjusting model parameters within a space limited to the range of feasible parameter values, leads to a performant model that does not necessarily simulate a realistic hydrological behaviour of the modelled watershed (e.g. percolation coefficient set to 0 which means there is no infiltration to supply the groundwater). The proposed procedure was applied with split sample and a differential split sample procedure and in both cases the calibrated model is robust and performant.

The second objective of the thesis aims at defining a new performance measure that allows conserving all the information about the variability, form and timing of hydrological states. To do so, functional data analysis (FDA) is integrated within the process of model calibration and validation. This mathematical tool allows analyzing observed data within a framework as continuous functions. In our case, it allows us to analyze the annual hydrograph as a single entity instead of 365 daily observations. In this context, the proposed objective function is a distance between the simulated and observed hydrographs. The calibration procedure defined in the first objective of the thesis is used and two objective functions adapted for each step are defined. Using FDA during calibration yields to similar results on high flows as using the Nash-Sutcliffe efficiency criterion (e.g. similar NSE values, around 0.9 during calibration and validation periods for the Lac St-Jean drainage basin). It improves the simulation of low flows, the simulation of some hydrological signatures and conserves the entire variability of the hydrograph. Evaluating a number of hydrological signatures, i.e. descriptive statistics reflecting different flow regime characteristics to investigate the performance of the calibrated model helps to identify which model component is the source of possible inadequacies in the simulations. These inadequacies are most likely caused by elements such as the poor performance to simulate the hydrological processes (e.g. snowmelt, evapotranspiration) not explicitly or specifically considered at this step.

The third objective explores a new definition of a performance measure based on FDA and compares the efficiency of calibration with this measure versus other measures used in conventional calibrations such as using the NSE in a single criterion calibration, using a Kling Gupta Efficiency (KGE) as a multi-objective criterion and using hydrological signatures in a diagnostic approach. The objective function based on FDA is defined as a ratio between the integrated annual observed and simulated volume. The procedure developed in the first part of the thesis that combines Monte Carlo simulations with a depth function is used to compare the performance and the consistency of the parameters sets resulting from the different calibrations approaches. The results show that calibration based on FDA allows a better estimation of the entire hydrograph and can achieve high consistency levels (i.e. finding parameters sets adequately reproducing all hydrological signatures considered). The conventional calibrations showed efficiency and high performance on simulating the spring flood at the expense of low flows. However, the NSE remains the best conventional method to catch the timing and the flood volume in a snow-dominated watershed. The uncertainty analysis shows that, regardless of the calibration method used, the uncertainty is more pronounced during the spring flood than in other periods of the year.

This fourth objective aims at improving the quality of the calibration in terms of reproducing the response of the system and capturing the changes of the internal states and hydrological processes. The proposed calibration strategy is a stepwise multi-criteria and process-based calibration. This strategy uses the calibration procedure defined in the first objective and the performance measures defined are based on the same principle as for the second objective of the thesis. To reproduce the same observed flows, the model tends to compensate between the water volumes coming from the different sources (groundwater flow, interflow, and runoff). To reduce this effect, model components controlling each hydrological process are calibrated separately. For each component, a specific criterion is defined based on the part of the hydrograph directly related to the specific hydrological process. This strategy is applied to calibrate the model only at the Lac St-Jean drainage basin, where different gauging stations are available. The estimation of both low and high flows simultaneously is improved with a NSE value of 0.86 and a winter flow relative bias of 16% compared to a classical strategy using only the observed flows for calibration (NSE is 0.71 and winter flow relative bias is 21%). The validation of the calibrated model at different gauging stations confirms its good performance (NSE value above 0.7) in sub-

basins showing similar hydrological characteristics as those of the sub-basin used for calibration. Overall, the mean NSE value at all stations of the Lac St-Jean drainage basin excluded for this third calibration strategy is 0.61.

The purpose of the fifth objective is to improve the performance of the model at a broader basin scale. The aim of any distributed/semi-distributed hydrological model, unlike lumped models, is to simulate the discharge at any point of the modelled watershed. Usually, the models are calibrated against flows observed from only one station. Hence, the calibrated model is performant to reproduce adequately flows at this station but not necessarily at a basin scale. Here, we propose two calibration strategies. By a global calibration, we mean a calibration at a basin scale (i.e. multiple gauging stations) and local calibration implies only one station. The first strategy is a global calibration by adjusting all model parameters at the different stations simultaneously. The second strategy is a local calibration followed by a global calibration. The local calibration aims at exploring the spatial variability of parameters. Only parameters that have high variability are considered for the global calibration. The other parameters are fixed to an average of values obtained at all stations. Both strategies are applied with a split sample test. Calibration of the model using two different data periods allows exploring the temporal variability of the parameters. The results show that only the snow parameters and the drainage coefficients except the Upper-Zone Lower drainage coefficient (CVSB) are spatially and temporally variable. All other parameters were fixed. The overall global calibration is as performant as a local calibration (e.g. NSE value is 0.86 at the outlet of the watershed and NSE is 0.83 at the ASHA station, located further upstream).

Overall the work of this thesis proposes a combination of methods for automatic calibration including algorithms, objective functions and calibration strategies to set a realistic (conceptually realistic parameters and representing the internal states adequately and hydrological processes), performant and reliable model for both low and high flows simulations simultaneously. In addition, the thesis offers an exploration of the CEQUEAU parameters values deemed adequate in the Lac St-Jean drainage basin where different hydrometric stations are available and where the model is used in operational forecasts.

Keywords: Calibration, CEQUEAU, Depth Function, Functional Data Analysis, Monte Carlo Simulation, Tabu Search.

TABLE DES MATIÈRES

RÉSUMÉ	V
ABSTRACT	X
LISTE DES TABLEAUX.....	XIX
LISTE DES FIGURES.....	XXI
CHAPITRE 1 INTRODUCTION GENERALE.....	1
1. MISE EN CONTEXTE	1
2. STRUCTURE DE LA THESE.....	3
CHAPITRE 2 SYNTHÈSE.....	5
1. REVUE DE LITTÉRATURE.....	5
1.1 Modélisation hydrologique	5
1.2 Estimation des paramètres.....	6
2. PROBLÉMATIQUES ET OBJECTIFS DE LA THESE.....	12
3. MODÈLE HYDROLOGIQUE CEQUEAU	17
3.1 Schématisation du bassin versant et données météorologiques	17
3.2 Simulation du débit	19
4. ZONES D'ÉTUDE.....	23
4.1. Bassin versant du Lac St-Jean	23
4.2. Bassin versant de la rivière Nechako.....	24
4.3. Bassin versant du ruisseau Catamaran	26
5. OUTILS STATISTIQUES.....	28
5.1. Fonction de profondeur	28
5.2. Analyse de données fonctionnelles.....	30
6. SYNTHÈSE DES TRAVAUX DE RECHERCHE.....	34
6.1. Objectif 1 : Simulation Monte Carlo couplé au Tabu Search pour estimer les paramètres d'un modèle hydrologique	35
6.2. Objectif 2 : Analyse de données fonctionnelles pour calibrer et évaluer un modèle hydrologique	44
6.3. Objectif 3 : Comparaison d'approches de calibration avec analyse d'incertitude	56
6.4. Objectif 4 : Calibration par processus multicritère pour un modèle hydrologique conceptuel	66
6.5. Objectif 5 : Calibration multisite du modèle CEQUEAU au bassin versant du lac St-Jean	76

CHAPITRE 3 SIMULATION MONTE CARLO COUPLE AU TABU SEARCH POUR ESTIMER LES PARAMETRES D'UN MODELE HYDROLOGIQUE.....89

ABSTRACT.....	90
1 INTRODUCTION	91
2 MODEL, STUDY AREAS AND DATA	94
2.1 <i>CEQUEAU model</i>	94
2.2 <i>Study areas and data</i>	97
2.3 <i>Calibration/Evaluation structure</i>	99
3 METHODOLOGY	100
3.1 <i>Refining the parameters space</i>	101
3.2 <i>Sensitivity analysis</i>	105
3.3 <i>Tabu Search algorithm</i>	105
3.4 <i>Covariance Matrix Adaptation-Evolution strategy (CMA-ES)</i>	108
3.5 <i>Performance criteria</i>	108
4 Results.....	109
4.1 <i>Identification of the optimal space</i>	109
4.2 <i>Calibration without refining the parameter space</i>	114
4.3 <i>Calibration after refining the parameter space</i>	115
4.4 <i>Calibration with differential split sample</i>	119
5 DISCUSSION	120
6 SUMMARY AND CONCLUSION	124
ACKNOWLEDGMENTS.....	125
REFERENCES	125

CHAPITRE 4 ANALYSE DE DONNEES FONCTIONNELLES POUR CALIBRER ET EVALUER UN MODELE HYDROLOGIQUE.....129

ABSTRACT.....	130
1 INTRODUCTION	131
2 METHODOLOGY	133
2.1 <i>Data smoothing and functional statistics</i>	133
2.2 <i>Calibration procedure</i>	137
2.3 <i>Proposed objective function</i>	139
2.4 <i>Evaluation criteria</i>	141
3. CALIBRATION OF THE HYDROLOGICAL CEQUEAU MODEL.....	142
3.1 <i>CEQUEAU model</i>	142
3.2 <i>Study areas and data</i>	145
3.3 <i>Results</i>	147

4.	DISCUSSION	154
5.	CONCLUSIONS.....	156
	ACKNOWLEDGEMENT.....	157
	REFERENCES	157
CHAPITRE 5 COMPARAISON D'APPROCHES DE CALIBRATION AVEC ANALYSE		
D'INCERTITUDE.....		161
	ABSTRACT.....	162
1	INTRODUCTION	163
2	METHODOLOGY	168
2.1	<i>Calibration with classical approaches.....</i>	<i>168</i>
2.2	<i>Calibration with functional data analysis.....</i>	<i>172</i>
2.3	<i>Parameter search procedure</i>	<i>174</i>
3	CASE STUDY	175
3.1	<i>Hydrological model description.....</i>	<i>175</i>
3.2	<i>Study area and data</i>	<i>177</i>
4	RESULTS	178
4.1	<i>Convergence of the algorithm.....</i>	<i>179</i>
4.2	<i>Comparison of the global performance of the four approaches</i>	<i>180</i>
4.3	<i>Comparison of the parameter distribution and its performance.....</i>	<i>183</i>
4.4	<i>Uncertainty Evaluation.....</i>	<i>188</i>
5	DISCUSSION	189
6	CONCLUSIONS.....	192
	ACKNOWLEDGMENT	193
	REFERENCES	193
CHAPITRE 6 CALIBRATION MULTI-CRITERE ET PAR PROCESSUS POUR UN MODELE		
HYDROLOGIQUE CONCEPTUEL.....		197
	ABSTRACT.....	198
1.	INTRODUCTION	199
2.	HYDROLOGICAL CEQUEAU MODEL AND STUDY AREA	201
2.1	<i>Hydrological CEQUEAU model.....</i>	<i>201</i>
2.2	<i>Study area</i>	<i>203</i>
3.	METHODOLOGY	204
3.1	<i>Smoothing data with functional data analysis</i>	<i>205</i>
3.2	<i>Global performance of the model</i>	<i>206</i>
3.3	<i>Adjusting snow parameters.....</i>	<i>209</i>
3.4	<i>Adjusting physical parameters.....</i>	<i>210</i>

3.5	<i>Adjusting routing parameters</i>	211
4.	RESULTS AND DISCUSSION.....	211
4.1	<i>Performance at the outlet station used for calibration</i>	211
4.2	<i>Basin-wide performance</i>	213
4.3	<i>The realism of the simulated internal processes</i>	214
5.	CONCLUSIONS.....	216
	SUPPLEMENTARY MATERIALS	217
	ACKNOWLEDGEMENTS.....	219
	REFERENCES	219
	CHAPITRE 7 CONCLUSION GENERALE ET PERSPECTIVES.....	223
	LISTE DE REFERENCES BIBLIOGRAPHIQUES	229

LISTE DES TABLEAUX

TABLE 2-1	EXEMPLES DE MODELES HYDROLOGIQUES.....	6
TABLE 2-2	LISTE DES PARAMETRES DU MODELE CEQUEAU ET LEURS DESCRIPTIONS.....	22
TABLE 2-3	CARACTERISTIQUES DES BASSINS VERSANTS CONSIDERES.	27
TABLE 2-4	CARACTERISTIQUES ANNUELLES DES SOUS-BASSINS DU LAC ST-JEAN.	27
TABLE 2-5	BORNES OPTIMALES OBTENUES APRES OPTIMISATION DE L'ESPACE INITIAL AVEC LA PROCEDURE PROPOSEE.....	41
TABLE 2-6	PERFORMANCE DES JEUX DE PARAMETRES DE L'ESPACE OPTIMAL AVEC LES QUANTILES 0.05 ET 0.95 ET L'ECART TYPE AUX QUATRE BASSINS CHUTE-DU-DIABLE (CD), PASSES-DANGEREUSES(PD), RIVIERE NECHAKO (NECH) ET RUISSEAU CATAMARAN (CAT)	42
TABLE 2-7	PERFORMANCE DU MODELE CALIBRE AVEC LA VALIDATION CROISEE DIFFERENTIELLE AUX TROIS BASSINS (CHUTE-DU-DIABLE (CD), PASSES-DANGEREUSES(PD), RIVIERE NECHAKO (NECH).	43
TABLE 2-8	BIAIS SAISONNIER RELATIF ABSOLU DES DEUX CONFIGURATIONS OPTFDA ET OPTNSE.	53
TABLE 2-9	PARAMETRES AJUSTES AVEC LA CALIBRATION BASEE SUR LA FDA (OPTFDA) ET AVEC LE NSE (OPTNSE) AU BASSIN DU LAC ST-JEAN (LSJ), AU RUISSEAU CATAMARAN (CAT) ET A LA RIVIERE DU NECHAKO (NECH).	54
TABLE 2-10	DISTRIBUTION DES JEUX DE PARAMETRES OPTIMAUX OBTENUS AVEC LES APPROCHES DE CALIBRATION AVEC NSE (OPTNSE) AVEC KGE (OPTKGE), BASEE SUR LES SIGNATURES HYDROLOGIQUES AU SEUIL D'ACCEPTATION DE 10% (OPTSIG10) ET DE 5% (OPTSIG5) ET L'APPROCHE PROPOSEE BASEE SUR LA FDA (OPTFDA).	60
TABLE 2-11	VALEURS DES PARAMETRES DU MODELE CEQUEAU AJUSTEES CONTRE LES DEBITS (OPTGLOB) ET AVEC L'APPROCHE PROPOSEE PAR PROCESSUS (OPTPROCESS).	74
TABLE 2-12	RESULTATS DE LA CALIBRATION MULTISITE SUR LA PERIODE 1998-2006.....	81
TABLE 2-13	RESULTATS DE LA CALIBRATION MULTISITE SUR LA PERIODE 2003-2011.....	82
TABLE 2-14	PERFORMANCE DES DEUX STRATEGIES DE CALIBRATION MULTISITE.	84
TABLE 2-15	NSE DE LA CALIBRATION GLOBALE ET LA CALIBRATION LOCALE.	87
TABLE 3-1	CEQUEAU PARAMETERS DESCRIPTION AND RANGES.	96
TABLE 3-2	SOME CHARACTERISTICS OF THE BASINS STUDIED.	97
TABLE 3-3	PERFORMANCE OF THE SET OF THE DEEPEST POINTS Y* ACCORDING TO DIFFERENT THRESHOLDS FIXED AS QUANTILES.	103
TABLE 3-4	PARAMETERS OPTIMAL RANGES.	111
TABLE 3-5	SENSITIVITY OF PARAMETERS IN THE INITIAL SPACE (S) AND IN THE OPTIMAL SPACE (SOPT).	113
TABLE 3-6	PERFORMANCE OF THE INFERRED OPTIMAL SET WITH THE 0.05 AND 0.95 QUANTILES BY THE FIRST STEP OF THE PROPOSED METHODOLOGY.	114
TABLE 3-7	PERFORMANCE OF THE SET OF PARAMETERS OBTAINED WITH TS AND CMA-ES EXECUTED IN THE INITIAL SPACE.	115

TABLE 3-8	PERFORMANCES OF THE SET OF PARAMETERS RESULTED OF THE CALIBRATION IN THE OPTIMAL SPACE.	116
TABLE 3-9	SEASONAL NSE AND BIAS OF THE SET OF PARAMETERS RESULTED OF THE CALIBRATION IN THE OPTIMAL SPACE.....	119
TABLE 3-10	PERFORMANCE OF THE SET OF PARAMETER RESULTED OF THE CALIBRATION WITH THE DIFFERENTIAL SPLIT SAMPLE.....	120
TABLE 4-1	MODEL KEY PARAMETERS.....	144
TABLE 4-2	CHARACTERISTICS OF THE BASINS USED.....	146
TABLE 4-3	NSE AND SIGNATURE RELATIVE BIAS (%) VALUES FOR CALIBRATION AND EVALUATION PERIODS FOR BOTH CALIBRATIONS OPTNSE AND OPTFDA.....	148
TABLE 4-4	CALIBRATED PARAMETER VALUES BY THE PROPOSED APPROACH (OPTFDA) AND USING THE NSE (OPTNSE) FOR THE THREE DRAINAGE BASINS (LAC ST-JEAN, CATAMARAN BROOK, NECHAKO RIVER).....	153
TABLE 5-1	HYDROLOGICAL SIGNATURES USED IN THIS STUDY.	170
TABLE 5-2	MODEL KEY PARAMETERS.....	177
TABLE 5-3	MEAN AND STANDARD DEVIATION OF PARAMETER VALUES OBTAINED BY THE DIFFERENT CALIBRATION APPROACHES. HIGHLIGHTED VALUES ARE THE LOWEST STANDARD DEVIATION.	183
TABLE 5-4	PERFORMANCE OF THE MEAN PARAMETER VALUES FOR EACH CALIBRATION APPROACH. VARIABILITY A , BIAS B AND CORRELATION R ARE THE KGE COMPONENTS; OF-SIG10 AND OF-SIG5 ARE THE NUMBER OF SIGNATURES SATISFIED OF THE 12 SIGNATURES CONSIDERED IN THIS STUDY AT THE 10% AND 5% ACCEPTANCE THRESHOLD RESP.	185
TABLE 6-1	KEY PARAMETERS OF THE CEQUEAU MODEL WITH CALIBRATED VALUES AGAINST OVERALL PERFORMANCE AND USING THE PROPOSED APPROACH.	202
TABLE 6-2	HYDROMETRIC STATIONS IN THE LAC ST-JEAN DRAINAGE BASIN.	203
TABLE 6-3	CALIBRATION CRITERIA TO EVALUATE MODEL COMPONENTS CONSIDERING DATA (FLOWS, RUNOFF AND BASEFLOW) AS FUNCTIONAL DATA.	207
TABLE 6-4	PERFORMANCE OF THE CALIBRATED MODEL (%).	212
TABLE 6-5	RESULTS OF THE CALCULATION OF THE CALIBRATION CRITERIA.	217

LISTE DES FIGURES

FIGURE 1-1	MODELISATION PLUIE-DEBIT CONVENTIONNELLE.	2
FIGURE 2-1	ORGANIGRAMME DE LA CALIBRATION AUTOMATIQUE COMME PROBLEME D'OPTIMISATION INVERSE. EN ROUGE SONT LES DIFFICULTES ASSOCIEES A LA DEFINITION DE CHAQUE ELEMENT. EN BLEU SONT LES OBJECTIFS DE LA THESE.	15
FIGURE 2-2	SUBDIVISION D'UN BASSIN VERSANT EN CARREAUX ENTIERS ET CARREAUX PARTIELS SELON LES LIGNES DE PARTAGES DES EAUX (MORIN ET PAQUET, 2007)	18
FIGURE 2-3	SCHEMA DE LA FONCTION DE PRODUCTION DU MODELE CEQUEAU (MORIN ET PAQUET, 2007).	19
FIGURE 2-4	SCHEMA DE LA FONCTION DE TRANSFERT DU MODELE CEQUEAU (MORIN ET PAQUET, 2007).	21
FIGURE 2-5	CARTE DU BASSIN VERSANT DU LAC ST-JEAN ET LA LOCALISATION DES SOUS-BASSINS PASSES-DANGEREUSES ET CHUTE-DU-DIABLE.	24
FIGURE 2-6	CARTE DU BASSIN DE LA RIVIERE NECHAKO.	25
FIGURE 2-7	CARTE DU BASSIN DU RUISSEAU CATAMARAN (ADAPTEE DE CAISSIE ET AL., 1998).	26
FIGURE 2-8	FONCTIONS MOYENNES (A GAUCHE) ET TEST DE STUDENT ASSOCIE (A DROITE).	33
FIGURE 2-9	SCHEMA DE LA CALIBRATION AUTOMATIQUE D'UN MODELE HYDROLOGIQUE ET LES METHODES PROPOSEES (EN ROUGE) POUR DEFINIR CHAQUE ELEMENT.	34
FIGURE 2-10	ORGANIGRAMME DU TABU SEARCH.	37
FIGURE 2-11	EXEMPLE DE LISSAGE DES HYDROGRAMMES AU LAC ST-JEAN POUR LES DEBITS OBSERVES ENTRE 2002 ET 2009. A) DEBITS JOURNALIERS OBSERVES. B) SCATTER PLOT ENTRE LES DEBITS OBSERVES ET LES DEBITS LISSES. C)-F) HYDROGRAMMES OBSERVES ET LISSES POUR LES ANNEES SELECTIONNEES.	46
FIGURE 2-12	ILLUSTRATION DE L'EXTRACTION DES EVENEMENTS HYDROLOGIQUES A PARTIR DE LA DERIVEE DE L'HYDROGRAMME.	47
FIGURE 2-13	TEST DE STUDENT FONCTIONNEL UTILISE POUR COMPARER LES HYDROGRAMMES OBSERVES ET SIMULES.	50
FIGURE 2-14	ANALYSE DES DEBITS MAXIMUMS EXTREMES DE COURBES DE DEBITS CLASSES OBSERVES ET SIMULES AVEC LE MODELE PARAMETRE AVEC OPTFDA (MODELE CALIBRE AVEC FDA) ET OPTNSE (MODELE CALIBRE AVEC NSE) AUX TROIS BASSINS : LAC ST-JEAN (LSJ), RUISSEAU CATAMARAN (CAT) ET A LA RIVIERE NECHAKO (NECH).	51
FIGURE 2-15	ANALYSE DES DEBITS MOYENS DE LA COURBE DE DEBITS CLASSES OBSERVES ET SIMULES AVEC OPTFDA (MODELE CALIBRE AVEC FDA) ET OPTNSE (MODELE CALIBRE AVEC NSE) AUX TROIS BASSINS : LAC ST-JEAN (LSJ) , RUISSEAU CATAMARAN (CAT) ET LA RIVIERE NECHAKO (NECH).	52
FIGURE 2-16	ANALYSE DES DEBITS FAIBLES DE LA COURBE DE DEBITS CLASSES OBSERVES ET SIMULES AVEC OPTFDA (MODELE CALIBRE AVEC FDA) ET OPTNSE (MODELE CALIBRE AVEC NSE) AUX TROIS BASSINS : LAC ST-JEAN (LSJ) , RUISSEAU CATAMARAN (CAT) ET LA RIVIERE NECHAKO (NECH).	52
FIGURE 2-17	CONVERGENCE DE L'ALGORITHME SELON L'APPROCHE DE CALIBRATION ET ANALYSE DE LA DISTRIBUTION DE LA FONCTION OBJECTIF CONSIDEREE.	59

FIGURE 2-18	VALEURS DE NSE EN CALIBRATION VS EN VALIDATION POUR LES JEUX DE PARAMETRES OPTIMAUX OBTENUS AVEC LA CALIBRATION AVEC NSE (OptNSE), AVEC KGE (OptKGE), BASEE SUR LES SIGNATURES HYDROLOGIQUES AU SEUIL D'ACCEPTATION DE 10% (OptSIG10) ET DE 5% (OptSIG5) ET L'APPROCHE PROPOSEE BASEE SUR LA FDA (OptFDA).	61
FIGURE 2-19	NOMBRE DE SIGNATURES SATISFAITES AU SEUIL D'ACCEPTATION 10% EN CALIBRATION VS EN VALIDATION AVEC LES JEUX DE PARAMETRES OPTIMAUX OBTENUS AVEC LA CALIBRATION AVEC NSE (OptNSE), AVEC KGE (OptKGE), BASEE SUR LES SIGNATURES HYDROLOGIQUES AU SEUIL D'ACCEPTATION DE 10% (OptSIG10) ET DE 5% (OptSIG5) ET L'APPROCHE PROPOSEE BASEE SUR LA FDA (OptFDA).	62
FIGURE 2-20	NOMBRE DE SIGNATURES HYDROLOGIQUES SATISFAITES AU SEUIL D'ACCEPTATION 5% EN CALIBRATION VS EN VALIDATION AVEC LES JEUX DE PARAMETRES OPTIMAUX OBTENUS AVEC LA CALIBRATION AVEC NSE (OptNSE), AVEC KGE (OptKGE), BASEE SUR LES SIGNATURES HYDROLOGIQUES AU SEUIL D'ACCEPTATION DE 10% (OptSIG10) ET DE 5% (OptSIG5) ET L'APPROCHE PROPOSEE BASEE SUR LA FDA (OptFDA).	63
FIGURE 2-21	HYDROGRAMME MOYEN OBSERVE AVEC LES HYDROGRAMMES MOYENS MINIMAUX ET MAXIMAUX SIMULES EN PERIODE DE CALIBRATION AVEC LE MODELE CALIBRE AVEC NSE (OptNSE), AVEC KGE (OptKGE), AVEC LES SIGNATURES HYDROLOGIQUES AU SEUIL D'ACCEPTATION DE 10% (OptSIG10) ET DE 5% (OptSIG5) ET AVEC L'APPROCHE PROPOSEE BASEE SUR LA FDA (OptFDA).	65
FIGURE 2-22	MOYENNE INTERANNUELLE DES HYDROGRAMMES LISSES OBSERVES ET SIMULES A LA STATION LSJ.	70
FIGURE 2-23	COURBE DE DEBITS CLASSES OBSERVES ET SIMULES PAR OPTPROCESS ET OPTGLOB.	71
FIGURE 2-24	STOCK DE NEIGE SIMULE AVEC LE MODELE CALIBRE CONTRE LES DEBITS (OPTGLOB) ET AVEC L'APPROCHE PAR PROCESSUS (OPTPROCESS) AUX DEUX STATIONS AU BASSIN DU LAC ST-JEAN: A L'EXUTOIRE (LSJ) ET ASHAPSHUAM (ASHA).	72
FIGURE 2-25	NIVEAU D'EAU SIMULE AVEC LE MODELE CALIBRE CONTRE LES DEBITS (OPTGLOB) ET AVEC L'APPROCHE PAR PROCESSUS (OPTPROCESS) DANS LES DEUX RESERVOIRS CONCEPTUALISE DANS CEQUEAU AUX DEUX STATIONS AU BASSIN LAC ST-JEAN: A L'EXUTOIRE (LSJ) ET ASHAPSHUAM (ASHA).	72
FIGURE 2-26	ORGANIGRAMME DE LA CALIBRATION MULTISITE LOCALE SUIVIE PAR UNE CALIBRATION GLOBALE.	77
FIGURE 2-27	SERIE CHRONOLOGIQUE DES DEBITS OBSERVES A L'EXUTOIRE DU BASSIN VERSANT DU LAC ST-JEAN.	79
FIGURE 2-28	SERIE CHRONOLOGIQUE DES DEBITS OBSERVES A LA STATION ASHA.	79
FIGURE 2-29	SERIE CHRONOLOGIQUE DES DEBITS OBSERVES A LA STATION MISNI2.	80
FIGURE 2-30	COURBE DE DEBITS CLASSES OBSERVES ET SIMULES AVEC LA CALIBRATION MULTISITE SUR LA STATION LSJ.	85
FIGURE 2-31	COURBE DE DEBITS CLASSES OBSERVES ET SIMULES AVEC LA CALIBRATION MULTISITE SUR LA STATION ASHA.	86

FIGURE 2-32	COURBE DE DEBITS CLASSES OBSERVES ET SIMULES AVEC LA CALIBRATION MULTISITE SUR LA STATION MISNI2.	86
FIGURE 3-1	DIAGRAM OF THE PRODUCTION FUNCTION (ADAPTED FROM MORIN AND PAQUET 1995 AND ST-HILAIRE ET AL; 2000).	95
FIGURE 3-2	MAP OF THE NECHAKO RIVER IN THE WESTERN CANADA, CHUTE-DU-DIABLE AND PASSES-DANGEREUSES IN EASTERN CANADA.	98
FIGURE 3-3	MAP OF THE CATAMARAN BROOK BASIN IN THE EASTERN CANADA.	99
FIGURE 3-4	CHRONOLOGICAL FLOW SERIES USED TO CALIBRATE THE MODEL WITH THE DIFFERENTIAL SPLIT SAMPLE TEST ON CHUTE-DU-DIABLE BASIN.	100
FIGURE 3-5	SCATTER PLOT OF MODEL PARAMETERS, OF THE SET X (N RANDOM SNOWMELT PARAMETERS), THE SET X^* (TOP 10% OF X), AND THE SET Y^* (THE DEEPEST POINTS WITHIN THE SUBSET X^*).	102
FIGURE 3-6	TABU SEARCH ALGORITHM FLOW CHART.	107
FIGURE 3-7	THE EVOLUTION OF THE MEAN NASH OF THE ENSEMBLE Y^* IN THE IDENTIFICATION OF THE OPTIMAL SPACE PROCESS.	110
FIGURE 3-8	PERCENTAGE (G) OF THE INITIAL SPACE RETAINED BY PARAMETER AND BY BASIN.	112
FIGURE 3-9	INTERANNUAL MEAN FLOW SIMULATED AND OBSERVED DURING THE CALIBRATION PERIOD IN ALL BASINS.	117
FIGURE 3-10	INTERANNUAL MEAN FLOW SIMULATED AND OBSERVED DURING VALIDATION PERIOD IN ALL BASINS.	118
FIGURE 3-11	DEEPEST POINTS Y^* WITHIN THE SET X^* SELECTED BY MAHALANOBIS AND HALFSPACE DEPTH FUNCTIONS.	121
FIGURE 4-1	TRANSFORMATION OF FLOW TIME SERIES INTO CURVES REPRESENTING SMOOTHED HYDROGRAPHS. DAILY OBSERVED FLOWS IN THE LAC ST-JEAN BASIN FROM 2002 TO 2009 ARE PLOTTED IN THE UPPER FIGURE AND IN THE LOWER FIGURE, THE ASSOCIATED SMOOTHED HYDROGRAPHS ARE PLOTTED, WHERE EACH HYDROGRAPH IS A TEMPORAL CURVE.	135
FIGURE 4-2	EXAMPLE OF MEAN FUNCTIONS (LEFT FIGURE) WITH THE ASSOCIATED FUNCTIONAL STUDENT TEST (RIGHT FIGURE).	136
FIGURE 4-3	FLOW CHART OF THE CALIBRATION PROCEDURE.	138
FIGURE 4-4	EXTRACTING THE TIMING OF HYDROLOGICAL EVENTS THROUGH A CHANGE OF SLOPE FROM THE FIRST DERIVATIVE.	139
FIGURE 4-5	DIAGRAM OF THE CEQUEAU MODEL PRODUCTION FUNCTION (ADAPTED FROM MORIN AND PAQUET 2007 AND ST-HILAIRE ET AL; 2000).	143
FIGURE 4-6	LAC ST-JEAN DRAINAGE BASIN AND THE NECHAKO RIVER BASIN.	145
FIGURE 4-7	CATAMARAN BROOK BASIN (ADAPTED FROM (CAISSIE ET AL., 1998)).	146
FIGURE 4-8	RESULTS OF THE FUNCTIONAL STUDENT TEST FOR MODEL SIMULATIONS DURING THE CALIBRATION PERIOD. RIGHT PANEL: RESULTS OF OPTNSE. LEFT PANEL, RESULTS OF OPTFDA.	149

FIGURE 4-9	THE UPPER PART OF THE FLOW DURATION CURVE REPRESENTING THE RESPONSE OF THE CATCHMENT TO EXTREME HYDROLOGICAL EVENTS.	150
FIGURE 4-10	THE MID-SEGMENT OF THE FLOW DURATION CURVE REPRESENTING THE RESPONSE OF THE CATCHMENT TO MODERATE HYDROLOGICAL EVENTS.	151
FIGURE 4-11	THE LOWER PORTION OF THE FLOW DURATION CURVE REPRESENTING FLOWS DURING DROUGHT EVENTS.	151
FIGURE 5-1	CONVERTING RAW DATA TO CONTINUOUS FUNCTIONS TO BE USED WITH FDA. IN THE UPPER PANEL ARE THE ORIGINAL DAILY FLOW SERIES. IN THE LOWER PANEL ARE THE ASSOCIATED ANNUAL SMOOTHED HYDROGRAPHS WHERE EACH HYDROGRAPH IS A SINGLE OBSERVATION.	167
FIGURE 5-2	PRODUCTION FUNCTION OF CEQUEAU MODEL (ADAPTED FROM MORIN ET PAQUET 1995).	176
FIGURE 5-3	THE LOCALIZATION OF THE CHUTE-DU-DIABLE WATERSHED WITHIN THE GREAT LAC ST-JEAN DRAINAGE BASIN.	178
FIGURE 5-4	CONVERGENCE OF THE ALGORITHM USING DIFFERENT CALIBRATION APPROACHES (OPTNSE USING NSE, OPTKGE USING KGE, OPTSIG5 AND OPTSIG10 USING THE SIGNATURE-BASED CALIBRATION AT THE 5% AND 10% ACCEPTANCE THRESHOLD RESP., AND OPTFDA USING THE FUNCTIONAL APPROACH).	180
FIGURE 5-5	MODEL PERFORMANCE DURING CALIBRATION PERIOD VS MODEL PERFORMANCE DURING THE EVALUATION PERIOD.	181
FIGURE 5-6	MODEL CONSISTENCY DURING CALIBRATION PERIOD VS MODEL CONSISTENCY DURING THE EVALUATION PERIOD. SIG5 IS THE NUMBER OF SIGNATURES SATISFIED AT THE 5% ACCEPTANCE THRESHOLD AND SIG10 AT 10% ACCEPTANCE THRESHOLD.	182
FIGURE 5-7	THE OBSERVED AND SIMULATED MEAN HYDROGRAPHS DURING THE CALIBRATION PERIOD.	185
FIGURE 5-8	INTERANNUAL COEFFICIENT OF VARIATION OF FLOWS DURING THE CALIBRATION PERIOD.	186
FIGURE 5-9	THE RESULTS OF THE STUDENT TEST APPLIED FOR THE PARAMETERIZED MODEL WITH PARAMETERS MEAN VALUES OBTAINED WITH THE DIFFERENT APPROACHES. DAYS WHEN THE OBSERVED STUDENT STATISTIC (BLACK LINE) IS OVER THE P-VALUE (DOTTED LINE) ARE THE DAYS WHEN THE SIMULATION AND OBSERVATIONS ARE SIGNIFICANTLY DIFFERENT.	187
FIGURE 5-10	INTERANNUAL MEAN HYDROGRAPH WITH UNCERTAINTY BOUNDS DURING THE CALIBRATION PERIOD.	188
FIGURE 5-11	INTERANNUAL MEAN HYDROGRAPH WITH UNCERTAINTY BOUNDS DURING THE EVALUATION PERIOD.	189
FIGURE 6-1	FLOW CHART OF THE PRODUCTION FUNCTION OF CEQUEAU MODEL (ADAPTED FROM MORIN ET PAQUET 1995, ST-HILAIRE ET AL., 2000).	201
FIGURE 6-2	THE LAC ST-JEAN DRAINAGE BASIN.	204
FIGURE 6-3	INTERPRETATION OF THE FIRST DERIVATIVE OF A SMOOTHED HYDROGRAPH. UPPER FIGURE PRESENTS THE HYDROGRAPH FEATURES. BOTTOM FIGURE PROVIDE EXPLANATION OF EXTRACTING TIMING OF THESE FEATURES THROUGH A CHANGE OF SLOPE OF THE FIRST DERIVATIVE OF THE SMOOTHED HYDROGRAPH.	208

FIGURE 6-4	EXAMPLE OF SMOOTHED HYDROGRAPHS. IN THE UPPER FIGURE ARE THE ORIGINAL SMOOTHED HYDROGRAPHS. IN THE BOTTOM FIGURE ARE THE SHIFTED HYDROGRAPHS ACCORDING TO THE AVERAGE HYDROGRAPH.	209
FIGURE 6-5	EVALUATION OF MODEL ACCURACY USING THE FUNCTIONAL STUDENT TEST. UPPER FIGURE COMPARES THE OBSERVED HYDROGRAPHS WITH SIMULATED HYDROGRAPHS WITH OPTGLOB AND THE BOTTOM FIGURE WITH OPTPROCESS. DAYS WITH STUDENT STATISTIC (BLUE LINE) OVER P-VALUE (DOTTED RED LINES) ARE THE DAYS WHEN SIMULATIONS AND OBSERVATIONS ARE SIGNIFICANTLY DIFFERENT.	212
FIGURE 6-6	OBSERVED AND SIMULATED FLOW SERIES IN THE OUTLET (UPPER FIGURE) AND AT THE ASHA STATION (BOTTOM FIGURE) FOR SELECTED YEARS.	213
FIGURE 6-7	SIMULATED SNOWPACK BY THE CALIBRATED MODEL WITH THE PROPOSED PROCESS-BASED CALIBRATION (OPTPROCESS) AND AGAINST STREAMFLOW SOLELY (OPTGLOB) AT THE STATIONS LSJ (OUTLET) AND ASHA (UPSTREAM).	218
FIGURE 6-8	SIMULATED WATER VOLUME IN THE SATURATED (UPPER) AND UNSATURATED (LOWER) ZONES BY THE CALIBRATED MODEL WITH THE PROPOSED PROCESS-BASED CALIBRATION (OPTPROCESS) AND AGAINST STREAMFLOW SOLELY (OPTGLOB) AT STATIONS LSJ (OUTLET) AND ASHA (UPSTREAM).	218
FIGURE 7-1	PROCESSUS GENERAL DE LA CALIBRATION PROPOSE DANS CETTE THESE.	223
FIGURE 7-2	PRINCIPE DE LA CALIBRATION AVEC LA FDA PROPOSEE DANS CETTE THESE AINSI QUE LE PRINCIPE CLASSIQUE UTILISE.	225

CHAPITRE 1 INTRODUCTION GÉNÉRALE

1. Mise en contexte

Les modèles hydrologiques sont des ensembles d'équations mathématiques destinés à décrire de manière plus ou moins simple (selon le niveau de complexité du modèle) le cycle hydrologique d'un bassin versant. Ces modèles sont utilisés pour traiter plusieurs problématiques liées aux ressources en eau (Singh et Woolhiser, 2002; Moradkhani et Sorooshian, 2009). Ces problématiques incluent la prévision de certaines variables hydrologiques, gestion/conception de réservoirs, planification, ainsi que la compréhension de la dynamique et l'interaction entre climat et surface terrestre. Chaque modèle est conçu selon les objectifs pour lesquels il a été développé et son utilisation prévue (Singh et Woolhiser, 2002). Les modèles conceptuels sont en particulier favorisés à cause de leur simplicité à capturer approximativement la dynamique et le mécanisme hydrologique général dominant dans un bassin versant (Duan et al., 1992; Kavetski et al., 2006). Ils sont aussi favorisés pour leur rapidité d'exécution. Ces modèles sont paramétriques et utilisent le plus souvent une représentation des éléments de stockage, souvent appelés réservoirs comme composantes principales du modèle (Wagner et al, 2004; Pechlivanidis et al., 2011). Ces réservoirs ont pour but de simuler l'emménagement et permettre l'écoulement de l'eau sur un bassin versant. Ils sont alors alimentés par les précipitations, infiltration ou percolation et vidés par les flux d'évapotranspiration et de ruissellement. À cause de leur nature conceptuelle, l'identification de ces modèles (ex. définir la dimension (hauteur) des réservoirs, attribuer une valeur aux paramètres distribuant les volumes d'eau entre les réservoirs tels les paramètres de stockage, de vidange ou d'évapotranspiration) dépend de la comparaison entre les valeurs simulées et les observations historiques de débits dans le bassin modélisé. Ce processus d'identification est communément appelé calibration. En effet, la calibration consiste à ajuster les paramètres du modèle, alimenté par des données météorologiques contre les observations de débits en utilisant une mesure d'adéquation. Cette dernière, souvent appelée mesure de performance ou fonction objectif, quantifie la différence entre les simulations du modèle et les

observations réelles du système en question. Le schéma du processus de calibration est présenté dans la Figure 1-1.

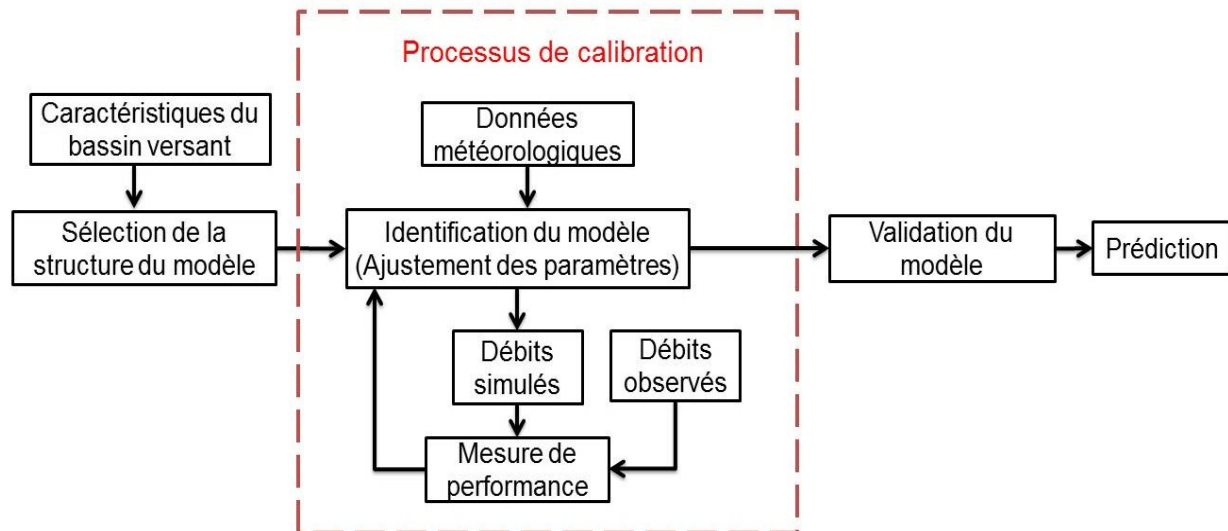


Figure 1-1 Modélisation pluie-débit conventionnelle.

Cette procédure implique la définition des bornes de paramètres (intervalle de valeurs que peuvent prendre les paramètres), le choix d’une stratégie de calibration (méthode d’ajustement), d’une fonction objectif ou mesure de performance, et le choix d’une période de calibration avec la quantité de données historiques suffisantes (Gupta et al., 1999; Singh et Woolhiser, 2002). La qualité et la performance d’une procédure de calibration impactent fortement la robustesse d’un modèle hydrologique (i.e. performant en conditions de calibration et en d’autres conditions possibles). Cette robustesse est essentielle pour produire des résultats fiables que ce soit dans un contexte d’étude d’impact de changement climatique ou autres applications opérationnelles tels la prévision, ou conception/gestion d’ouvrages hydrauliques (Coron et al., 2012). Dans ce contexte, le but de cette thèse est de développer des techniques et des stratégies de calibration efficaces pour améliorer la fiabilité, et la performance d’un modèle hydrologique conceptuel réaliste.

2. Structure de la thèse

La thèse est séparée en sept chapitres. Après cette introduction, on expose une synthèse des travaux réalisés. Les quatre chapitres suivants présentés sous forme d'articles scientifiques, exposent en détail la méthodologie et les résultats de chaque objectif spécifique. À noter que le dernier objectif spécifique ne fait pas l'objet d'un chapitre en entier, mais la méthodologie détaillée et les résultats sont présentés dans la partie synthèse 6.5. Finalement, une conclusion générale ainsi que les perspectives sont présentées.

Le chapitre 2 présente une revue de littérature sur la modélisation hydrologique ainsi que les outils de la calibration automatique. Il expose aussi les problématiques traitées dans cette thèse et les objectifs associés. On présente le modèle hydrologique, les zones d'études et les outils statistiques utilisés suivis par une synthèse des résultats des travaux réalisés.

Le chapitre 3 présente l'article intitulé 'Estimating hydrological model parameters by combining Monte Carlo simulations with Tabu Search'. Ce chapitre propose une nouvelle procédure de calage qui combine les simulations Monte Carlo avec le Tabu Search. Ce chapitre vise à réduire le risque d'équifinalité et estimer des paramètres hydrologiques réalistes et robustes.

Le chapitre 4 présente l'article intitulé 'Using functional Data Analysis to calibrate and evaluate hydrological model performance'. Ce chapitre vise à définir une méthode de calage qui conserve la forme de l'hydrogramme et permet d'évaluer les débits faibles et élevés simultanément.

Le chapitre 5 présente l'article intitulé 'Comparison of different calibration approaches with uncertainty analysis and introducing the functional data analysis approach'. Ce chapitre propose une nouvelle méthode de calage avec l'analyse de données fonctionnelles. Le but de cet article est de comparer la performance de la calibration avec l'analyse de données fonctionnelles et les approches classiques utilisées en hydrologie, à savoir l'approche basée sur une seule fonction objectif, l'approche multiobjective et l'approche basée sur les signatures hydrologiques.

Le chapitre 6 présente l'article intitulé 'A multi-criteria process-based calibration using functional data analysis to improve hydrological model realism'. Cet article propose une méthode de calibration par étape et par processus pour améliorer le degré de réalisme d'un modèle hydrologique conceptuel.

Le chapitre 7 présente une conclusion générale des travaux effectués avec les perspectives.

1. Revue de littérature

Afin de mieux mettre en contexte les travaux de cette thèse et les problématiques liées à l'estimation des paramètres de modèle hydrologique, une brève revue de littérature sur la modélisation hydrologique, et les différents éléments de la calibration (stratégie, fonction objectif et algorithmes) est présentée.

1.1 Modélisation hydrologique

Un modèle hydrologique est un ensemble d'équations mathématiques décrivant de manière simplifiée les processus hydrologiques dans un bassin versant et le bilan hydrique qui en découle à un pas de temps donné. Dans la littérature, il existe une multitude de modèles hydrologiques de natures différentes conçus pour répondre à une problématique et un usage précis (Singh et Woolhiser, 2002). Ces modèles peuvent être catégorisés selon la distribution spatiale des données et des sorties du modèle: on parle d'un modèle global, semi-distribué ou distribué. Les modèles peuvent aussi être définis selon le niveau de sophistication de la représentation des processus hydrologiques : conceptuel ou physique (Moradkhani et Sorooshian, 2009; Pechlivanidis et al., 2011). Un modèle global considère le bassin versant comme une seule entité et simule le débit uniquement en un point, soit à l'exutoire. En contraste, un modèle distribué discrétise le bassin en plusieurs unités hydrologiques pour simuler le débit sur n'importe quelle région du bassin en tenant compte des processus de transfert entre les différentes unités. Les modèles conceptuels quant à eux représentent de manière simple les processus dominant un bassin versant alors que les modèles physiques sont basés sur des lois physiques (ex. : loi de Darcy) pour décrire les processus hydrologiques. Ces modèles sont plus susceptibles de produire les bonnes réponses pour les bonnes raisons que les modèles conceptuels (Kirchner, 2006). Par contre, ils sont exhaustifs et requièrent une spatialisation des données et des paramètres ainsi que beaucoup de données en intrants qui sont fréquemment indisponibles. Quant aux modèles conceptuels, ils sont pratiques et requièrent généralement moins d'intrants que les modèles

distribués, se limitant souvent aux températures de l'air et précipitations. Le tableau 2-1 présente des exemples de modèles hydrologiques.

Table 2-1 Exemples de modèles hydrologiques

Modèle	Type	Référence
University of British Columbia watershed model (UBCWM)	conceptuel	Quick and Pipes(1977)
Hydrological Simulation (HBV) model	conceptuel, semi-distribué	Bergstrom (1976)
National Hydrology Research Institute (NHRI) model	physique, paramètres globaux	Vandenberg (1989)
Institute of Hydrology distributed Model (IHDM)	physique, distribué	Beven et al. (1987)
Physically Based Runoff Production Model (TOPMODEL)	physique, distribué	Beven (1995)
Generalized River Modeling Package-System Hydrologue European (MIKE-SHE)	physique, distribué	Refsgaard and Storm (1995)
Soil Water Assessment Tool (SWAT)	conceptuel, distribué	Arnold et al., (1998), Neitsch et al., (2005)
HYDROTEL	physique, distribué	Fortin et al. (2001a,b)
CEQUEAU	conceptuel, semi-distribué	Morin et Paquet (2007)

1.2 Estimation des paramètres

L'estimation des paramètres d'un modèle hydrologique nécessite la définition des conditions initiales, des bornes limites des valeurs possibles des paramètres, d'une procédure d'ajustement (manuelle ou automatique) ainsi que le choix d'une fonction objectif (mesure de performance). La qualité de la calibration dépend alors de chaque élément choisi. Elle est aussi limitée par la quantité et la qualité des données disponibles en intrants (ex. précipitation, température) et d'observations de débits. Les recherches pour améliorer les techniques de calibration s'intéressent alors à définir des mesures de performance fiables, des algorithmes d'optimisation efficaces et rapides (en particulier pour les modèles avec des temps de calculs importants) ainsi que la définition de la quantité de données nécessaires pour le calage. Dans le cadre de cette thèse, on s'intéresse aux deux premiers axes. La sous-section suivante présente une brève description des outils existant dans la littérature.

1.2.1 Procédure de calibration

Il existe deux approches de procédure de calibration : manuelle ou automatique. La première approche consiste à augmenter ou diminuer les valeurs des paramètres selon une comparaison graphique entre débits observés et simulés ou selon un critère d'ajustement jusqu'à obtention d'un résultat satisfaisant. Cette approche, bien qu'encore utilisée, est subjective et demande de l'expertise, qui est souvent difficile à transmettre et prend beaucoup de temps. En plus, elle devient plus complexe si le nombre de paramètres est important à cause de la non-linéarité des processus hydrologiques. La calibration automatique se formalise sous forme d'un problème d'optimisation inverse ou l'on cherche à optimiser (maximiser ou minimiser) une fonction objectif sachant les sorties du modèle via un algorithme d'optimisation (Moradkhani et Sorooshian, 2009). Elle est plus optimale, objective et efficace. Cependant, cette approche n'est pas si sophistiquée au point de remplacer entièrement une méthode manuelle en raison de la difficulté de construire les fonctions objectifs et les algorithmes d'optimisation qui reproduisent le jugement humain. Par conséquent, elle est plus efficace si elle utilisée conjointement avec une procédure manuelle (Pechlivanidis et al., 2011).

Des approches hybrides ont été proposées pour profiter des avantages qu'offrent les deux méthodes de calibration. Liew et Garbrecht (2003) proposent une méthode par étape pour calibrer le modèle SWAT. La première étape est destinée à ajuster les paramètres du modèle de manière à avoir une simulation adéquate et des biais annuels et mensuels les plus faibles possible. La deuxième étape est une inspection visuelle des hydrogrammes et les courbes de débits classés. Boyle et al. (2000) proposent une stratégie de calibration multicritères sur deux étapes qui simule les techniques et stratégies d'évaluation utilisées lors d'une calibration manuelle. Les critères utilisés reflètent différentes caractéristiques de l'hydrogramme en décomposant celui-ci en trois composantes (écoulement rapide lors d'épisode pluvieux, écoulement retardé, écoulement de base). Suivant la même philosophie, différentes stratégies de calibration automatique basées sur différents critères ont été proposées. Certaines stratégies consistent en une calibration sur différentes portions des séries chronologiques des débits comme Boyle et al. (2000), Bekele et Nicklow (2007), Choi et Beven (2007). D'autres utilisent différents critères (ex. NSE, pente de la courbe de débits classés, 'Root Mean Square Error' (RMSE)) qui reflètent la qualité d'estimation de ces différentes portions de l'hydrogramme (Moussa et Chahinian, 2009; Li et al., 2010; Kollat et al., 2012).

Récemment, le principe de diagnostic basé sur l'analyse des signatures hydrologiques est utilisé comme stratégie de calibration (ex. Yilmaz et al., 2008 ; Hingray et al., 2010; Euser et al., 2013; Shafii et Tolson, 2015). Ces signatures sont des indices calculés à partir des séries chronologiques de débits simulés et observés et caractérisent de multiples aspects de l'écoulement dans un bassin versant. Ce principe vise à examiner à quelle mesure un modèle peut être réconcilié aux données et détecter la composante qui nécessite une amélioration (Gupta et al., 2008). D'autres stratégies consistent à des calibrations sur des variables autres que les débits de crue (ex. ETP, débit de base) pour évaluer les états internes (Cao et al., 2006). Ce type de stratégie est toutefois limité par la disponibilité d'observations de ces variables. Il y a aussi des stratégies qui s'intéressent à des calibrations multisite en calibrant simultanément le modèle avec des observations de débit provenant de différents sites du bassin versant. Ces stratégies considèrent soit un critère défini comme une moyenne pondérée de la performance du modèle sur les différents sites (Cao et al., 2006; Li et al., 2010), ou un critère à optimiser simultanément sur les différents sites de manière à trouver le front de Pareto avec un algorithme destiné à l'optimisation multiobjective (Zhang et al., 2008).

1.2.2 Mesures de performance

Une fonction objectif ou mesure de performance est une mesure qui quantifie la différence entre les simulations du modèle et les observations mesurées dans un bassin versant. Il existe deux types de mesures de performances : dérivées des statistiques (les méthodes du maximum de vraisemblance) ou basées sur les aspects de performance du modèle (Wagener et al., 2004). La méthode GLUE (*Generalised Likelihood Uncertainty Analysis*) introduite par Beven et Binley (1992) est la méthode statistique la plus utilisée en hydrologie pour l'estimation des paramètres tout en tenant compte de l'incertitude de prédiction du modèle résultant. Toutefois, ces méthodes du maximum de vraisemblance sont subjectives, nécessitant la sélection préalable d'une distribution de paramètres, un critère du maximum de vraisemblance, et sont optimales sous certaines conditions (Gupta et al., 1998; Bardossy et Singh, 2008). Le deuxième type de mesure de performance est généralement défini comme la somme des résidus telle que la racine de la somme des carrés des résidus (RMSE) ou le critère du Nash-Sutcliffe (Nash-Sutcliffe Efficiency, ou NSE) qui est une normalisation du RMSE (Nash et Sutcliffe, 1970). Ce type de mesures est

plus utilisé par les hydrologues et semble plus intuitif pour ces derniers (Penchlivanidis et al., 2011).

Une multitude de mesures de performance basées sur une agrégation des résidus existent dans la littérature notamment des variations du RMSE et NSE. Krause et al. (2005) comparent plusieurs mesures (NSE, log NSE, RMSE, R^2 , ...) et concluent qu'aucune mesure de performance n'est optimale. En fait, aucune mesure de performance, peu importe sa définition, n'est capable d'identifier tous les aspects et caractéristiques importantes de l'hydrogramme (Yapo et al., 1998; Vrugt et al., 2003). Par exemple, le NSE est performant pour capter l'occurrence du débit de pointe et la corrélation linéaire entre les simulations et les observations, mais sous-estime la variabilité et le débit moyen (Krause et al., 2005; Gupta et al., 2009). Gupta et al. (2009) proposent une décomposition de ce critère, nommée *Kling-Gupta Efficiency* (KGE), pour tenir compte de la variabilité, de la corrélation et de la moyenne des débits. Il a été démontré que le KGE est capable de capter la variabilité, le débit de pointe ainsi que le débit moyen avec une corrélation linéaire élevée entre les simulations du modèle et les observations.

Gupta et al. (1998) soulignent l'importance d'utiliser simultanément une combinaison de mesures de performance dans le but de capter différentes caractéristiques de l'hydrogramme et introduisent ainsi l'approche multiobjective. Il existe deux stratégies de calibration avec l'approche multiobjective: agrégation de plusieurs fonctions ou le concept de Pareto. Dans le premier cas, il s'agit d'optimiser une seule fonction objectif exprimée comme une agrégation de plusieurs fonctions objectifs. Dans ce cas, les résultats dépendent fortement de la méthode d'agrégation. Quant au second cas, il s'agit d'optimiser simultanément plusieurs fonctions objectifs ($F(\theta) = \{f_1(\theta), \dots, f_n(\theta)\}$) pour déterminer le front de Pareto. Ce dernier est le sous-espace de solutions non-dominées tel que chaque solution θ_i satisfait les deux conditions suivantes : 1) toute solution θ_j n'appartenant pas au front de Pareto, il existe au moins une solution θ_i tel que $F(\theta_i) < F(\theta_j)$; 2) il n'existe aucune solution θ_j du front de Pareto tel que $f_k(\theta_j) < f_k(\theta_i) \forall k \in \{1, \dots, n\}$ (Gupta et al., 1998). La stratégie du Front de Pareto permet de considérer explicitement le compromis entre les différentes fonctions objectifs contrairement à la première stratégie qui consiste en une agrégation. Néanmoins, cette stratégie implique un temps de calcul exhaustif qui augmente avec le nombre de fonctions objectifs (Efstratidis and Koutsoyiannis, 2010; Pokhrel et al., 2012).

L'utilisation de l'approche multiobjective nécessite la sélection préalable de fonctions objectifs différentes qui mesurent différents aspects de l'hydrogramme jugés importants dépendamment du problème en question. Par conséquent, même si cette méthode est plus efficace que l'approche mono-objective, les jeux de paramètres résultants ne peuvent représenter adéquatement l'hydrogramme en entier. Le choix d'une méthode ou d'une autre reste subjectif et dépend de l'hydrologue (Krause et al., 2005).

1.2.3 Algorithme d'optimisation

Les premiers algorithmes utilisés pour la calibration automatique des modèles hydrologiques sont de type local tel que la méthode du gradient, méthode du simplex, ou '*Pattern Search*'. Ces méthodes se sont révélées incapables de trouver un optimum global et sont inadaptées dans le cadre de l'ajustement des paramètres des modèles hydrologiques à cause de la non-linéarité des processus hydrologiques et/ou la multimodalité des fonctions objectifs (Duan et al. 1992). Les algorithmes d'optimisation globale sont plus efficaces et plus susceptibles de trouver un optimum global.

Les algorithmes utilisés sont généralement de type évolutionnaire (ensemble de solutions) ou génétique (Duan et al., 1992; Sorooshian et al., 1993; Seibert, 2000; Zhang et al., 2008). Ces algorithmes consistent à échantillonner un ensemble initial aléatoire et ensuite améliorer ses membres selon différentes stratégies (Wagner et al., 2004). L'algorithme '*Shuffled Complex Evolution*' (SCE-UA, Duan et al., 1992) est sans contredit l'algorithme le plus populaire en hydrologie et est reconnu pour être efficace. La majorité des études prouvant l'efficacité de SCE-UA impliquent des modèles hydrologiques conceptuels avec des temps de simulations de l'ordre de quelques secondes ou moins, ce qui permet d'utiliser un nombre maximal d'évaluations de modèle élevé (qui dépassent 10 000 évaluations de modèle) (Tolson et Shoemaker, 2007). Franchini et al. (1998) utilisent 250 000 évaluations de modèle pour comparer entre SCE-UA, un algorithme génétique (GA) couplé à la Programmation Quadratique Séquentielle (SQP) et l'algorithme '*Pattern Search*' (PS) couplé aussi à la SQP pour un problème de dimension 37.

Permettre de tels nombres d'évaluations de modèle pour calibrer un modèle hydrologique distribué ou semi-distribué entraîne automatiquement des temps de calcul très longs. Tolson et Shoemaker (2007) proposent une approche nommée '*Dynamically Dimensioned Search*

Algorithm’ (DDS), qui est un algorithme d’optimisation globale stochastique basé sur une solution unique. Cet algorithme a la particularité d’ajuster la dimension de l’espace en fonction du nombre maximal d’évaluations de modèle permis. Cet algorithme est plus performant et plus efficace que le SCE-UA dans le cas d’une calibration d’un modèle hydrologique distribué avec un nombre de paramètres supérieur à 10 et un nombre maximal d’évaluations de modèle limité. En effet, Arsenault et al. (2013), dans leur étude comparant dix algorithmes (*‘Adaptive Simulated Annealing’* (ASA), *‘Covariance Matrix Evolution Strategy’* (CMA-ES), *‘Cucko Search’*, DDS, *‘Differential Evolution’*, GA, *‘Harmony Search’*, PS, *‘Particle Swarm Optimisation’* (PSO) et SCE-UA), concluent que le SCE-UA est très efficace quand l’espace des paramètres est petit, mais il n’est pas capable d’éviter un maximum/minimum local quand l’espace est très large. Ils concluent aussi que CMA-ES, ASA et DDS sont les trois meilleures méthodes parmi les algorithmes considérés pour calibrer un modèle hydrologique complexe (i.e. ayant un nombre important de paramètres à estimer). Dans un contexte opérationnel, ils recommandent les algorithmes CMA-ES et DDS. Ces algorithmes sont capables d’aboutir à une performance acceptable dans un nombre limité (ex. 1000) d’évaluations du modèle.

Ces algorithmes sont destinés à l’optimisation mono-objective et avec l’intérêt que gagne l’approche multiobjective, certains de ces algorithmes ont été adaptés pour l’optimisation multiobjective pour trouver le front de Pareto. Yapo et al. (1998) proposent une extension de l’algorithme SCE-UA : *Multi-objective Complex Evolution* (MOCOM-UA). Vrugt et al. 2003 proposent l’algorithme *Mutliobjective Shuffled Complex Evolution Metropolis* (MOSCEM) une amélioration du *Shuffled Complex Evolution Metropolis* SCEM-UA qui utilise le concept de la dominance de Pareto pour faire évoluer la population. Parmi les algorithmes adaptés pour l’optimisation multiobjective et appliqués pour l’ajustement de paramètres, on cite *‘Multiobjective Particle Swarm (MOPSO)* (Gill et al., 2006), *‘Non Dominated Sorting Genetic Algorithm II (NSGAI)*’ (Shafii et De Smedt, 2009), *Epsilon Dominance Nondominated Sorted genetic Algorithm II (ϵ -NSGAI)*, *‘Strength Pareto evolutionary algorithm 2 (SPEA 2)’* (Tang et al., 2006) et le *‘Pareto Archived Dynamically Dimensioned Search (PA-DDS)’* (Asadzadeh et al., 2014).

2. Problématiques et objectifs de la thèse

2.1 Problématiques

La fiabilité d'un modèle hydrologique est étroitement liée à la qualité de sa calibration (Duan et al., 1992; Gupta et al., 1999). Avec le développement des ressources informatiques et l'émergence de modèles hydrologiques plus complexes, les algorithmes d'optimisation globale sont de plus en plus utilisés pour la calibration automatique. La plupart de ces algorithmes sont basés sur l'évolution d'une population (ensemble de solutions) tel que les algorithmes génétiques (Holland, 1975; Seibert, 2000; Shafii et De Smedt, 2009; Boisvert et al., 2016) ou la méthode *Shuffled Complex Evolution* (SCE-UA, Duan et al., 1992) largement utilisée en hydrologie (ex. Gupta et al., 1999 ; Madsen, 2000; Madsen et al., 2002; Francés et al., 2007). Ces algorithmes consistent à explorer aléatoirement l'espace des paramètres borné par les valeurs limites que peuvent prendre ces paramètres (Arsenault et al., 2013). Ces algorithmes convergent vers des ensembles de solutions équivalentes, i.e. jeux de paramètres ayant la même performance au lieu d'un optimum global unique. L'existence de jeux de paramètres de performances équivalentes, appelée équi-finalité, est causée principalement par les incertitudes liées à la structure du modèle et la non linéarité des modèles hydrologiques (Beven et Freer, 2001). Par conséquent, ces algorithmes peuvent converger vers des jeux de paramètres conceptuellement ou physiquement irréalistes ou moins performants sur d'autres périodes de données (Gupta et al., 1999; Andréassian et al., 2012; Brigode et al., 2013; Gharari et al., 2013).

Le choix de la fonction objectif ou la mesure de performance lors de la calibration impacte aussi la performance du modèle calibré. En effet, les fonctions objectifs classiques réduisent toute l'information contenue dans les données en une seule valeur (Gupta et al., 2008). En conséquence, chaque fonction objectif est particulièrement sensible à un seul aspect (débit élevé ou étiage) de l'hydrogramme dépendamment de sa définition. Conséquemment, utiliser une seule fonction objectif est inadéquat pour représenter toute la variabilité d'un hydrogramme (Vrugt et al., 2003; Pechlivanidis et al., 2011). On cite par exemple le critère de Nash Sutcliffe (Nash et Sutcliffe, 1970), critère qui est encore le plus utilisé en hydrologie pour le calage de modèles hydrologiques. Ce critère est souvent jugé le choix optimal pour évaluer la simulation dans un contexte où on met l'emphasis sur capter le débit de pointe et son occurrence. En revanche, il est reconnu être inadapté pour représenter les débits d'étiage. L'approche multiobjective est alors

favorisée pour inclure plusieurs fonctions objectives reflétant l'adéquation de plusieurs aspects de l'hydrogramme. Dans ce cas, la calibration consiste soit à optimiser une agrégation de plusieurs fonctions objectifs ou utiliser le concept de Pareto (Pechlivanidis et al., 2011; Azadzadeh et al., 2014). Le concept de Pareto consiste à optimiser différentes fonctions objectives simultanément dans le but de trouver le sous-espace non-dominé (front de Pareto optimal) où améliorer une fonction objectif mène à une détérioration d'une autre (Yapo et al., 1998). La méthode multiobjective reste néanmoins complexe pour trouver un jeu de paramètres optimal. En fait, un jeu de paramètres optimal selon une fonction objectif est rarement optimal pour une autre fonction objectif ; alors si des compromis entre ces différentes fonctions sont nécessaires pour satisfaire des objectifs multiples (représenter la crue, débits d'étiage), un jeu de paramètres optimal devient difficile à définir (Beven, 2001).

Les mesures de performance reflètent généralement l'adéquation entre les simulations de débits et les observations. Comparer les débits observés et simulés ne révèle aucune information directe sur la performance du modèle à reproduire les processus hydrologiques internes dans le bassin versant considéré (Gupta et al., 2008). En effet, le modèle a tendance à compenser entre les volumes provenant de ses différentes composantes et conséquemment entre les différents paramètres, pour reproduire le plus fidèlement possible le volume final observé. Pour cette raison, un modèle peut simuler adéquatement le débit observé sans tenir compte de la réalité du changement des états et processus hydrologiques. Certes, il est important d'obtenir les bonnes réponses pour les bonnes raisons (Kirchner, 2006). L'identification d'un tel modèle réaliste est toutefois limitée par de nombreux facteurs tels : le choix de critères utilisés lors de la calibration, la nature du modèle hydrologique, la quantité d'information et de données, etc...

Les modèles distribués ou semi-distribués visent à représenter la variabilité spatiale des processus hydrologiques qui est due entre autres à l'hétérogénéité des conditions physiographiques et météorologiques dans un bassin versant. Généralement ces modèles sont calibrés contre les observations provenant d'un site jaugé, idéalement à l'exutoire du bassin versant. Ce type de calibration en un site particulier ne tient pas compte de la performance du modèle à simuler le débit sur tous les points du bassin versant (White et Chaubey, 2005). En effet, quelle que soit la qualité de la calibration du modèle sur un site, l'extrapolation des paramètres sur d'autres sites est problématique à cause de la non-unicité des paramètres. La calibration avec des observations de débits provenant de différents sites permet de contraindre

davantage l'ajustement des paramètres et augmenter la fiabilité spatiale du modèle. Ces modèles requièrent un temps de simulation important à cause de leur nature distribuée ce qui entraîne des temps de calcul importants pour procéder à de multiples calibrations par sous-bassins. Une calibration globale en considérant tous les sous-bassins simultanément pourrait augmenter la performance du bassin et réduire le temps de calcul.

2.2 Objectifs de la thèse

L'objectif global de la thèse est de développer de nouvelles approches efficaces pour améliorer la qualité de l'ajustement automatique des paramètres d'un modèle hydrologique ainsi que sa performance et son degré de réalisme. Les méthodes développées sont appliquées au modèle hydrologique conceptuel semi-distribué CEQUEAU sur différents bassins. L'application sur trois bassins localisés dans différentes régions du Canada permet de valider la robustesse des méthodes face à différentes caractéristiques hydrologiques et physiographiques. Cinq objectifs spécifiques visant les différents axes de la calibration sont définis. Les quatre premiers objectifs spécifiques traitent les éléments clés du processus de calibration à savoir algorithme d'optimisation et fonctions objectifs comme illustrés dans la Figure 2-1. Le cinquième objectif s'intéresse à la stratégie de calibration pour améliorer la performance du modèle pour capter la variabilité spatiale dans un bassin versant.

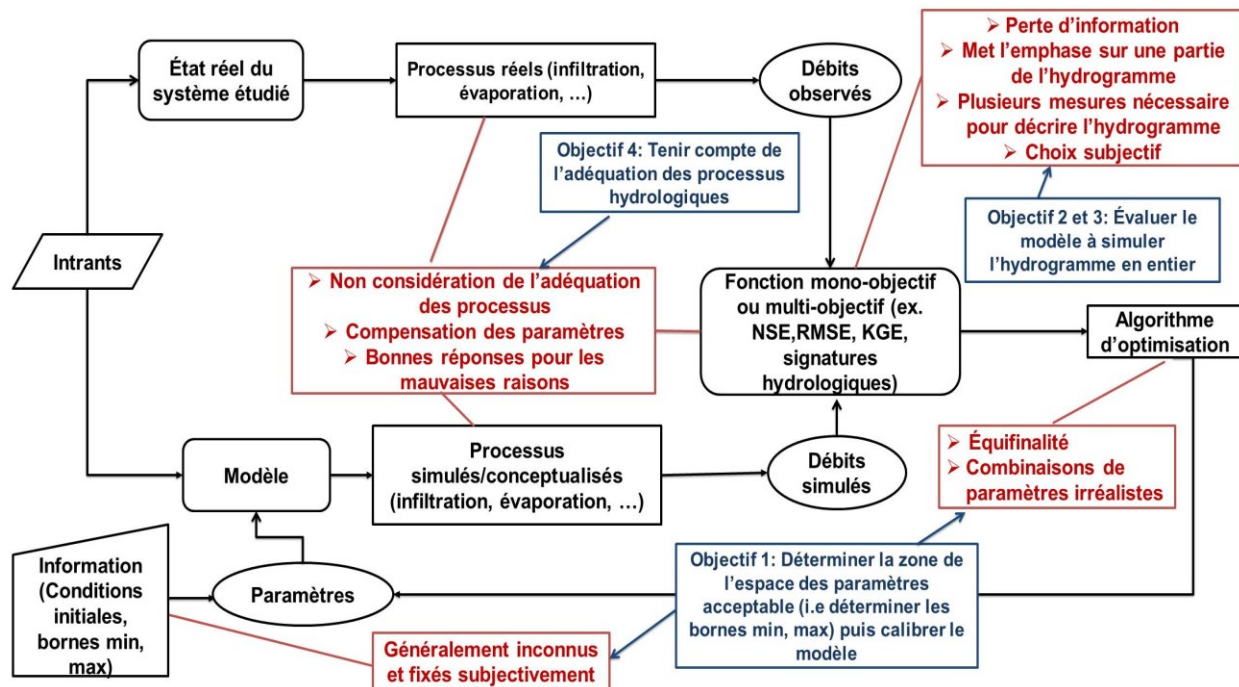


Figure 2-1 Organigramme de la calibration automatique comme problème d'optimisation inverse. En rouge sont les difficultés associées à la définition de chaque élément. En bleu sont les objectifs de la thèse.

Le premier objectif spécifique traite la problématique reliée au risque d'équifinalité et de la convergence des algorithmes d'optimisation vers des combinaisons de paramètres irréalistes, bien que performantes. Cet objectif vise alors à définir la région optimale de l'espace des paramètres par bassin selon une mesure de performance définie, pour ensuite ajuster les paramètres du modèle; contrairement aux approches classiques où les paramètres sont ajustés dans un espace défini uniquement selon les valeurs plausibles de ces paramètres. La procédure définissant l'espace optimal est basée sur une simulation Monte Carlo utilisant une fonction de profondeur pour restreindre l'échantillonnage des paramètres. La procédure est inspirée de l'algorithme ROPE (Robust Parameter Estimation) proposé par Bardossy et Singh (2008). Cependant, une adaptation de l'algorithme d'optimisation globale 'Tabu Search', rarement utilisé en hydrologie, est utilisée pour ajuster les paramètres du modèle dans l'espace restreint. Le but d'utiliser le Tabu Search est de reproduire les techniques de la calibration manuelle à savoir la modification d'un paramètre à la fois. Cette stratégie vise à améliorer une solution initiale en conservant la performance du modèle gagnée sur les itérations précédentes. Contrairement, aux

autres approches d'optimisation comme les algorithmes évolutionnaires qui consistent en une modification de l'ensemble des paramètres à la fois qui est associé à un risque de perte de performance sur les itérations suivantes. Toutefois, l'utilisation d'un algorithme ou un autre dépend du choix de l'utilisateur.

Le deuxième objectif spécifique s'intéresse aux problématiques et limites des fonctions objectifs existantes. Cet objectif vise à définir une fonction objectif adéquate tenant compte de toute la variabilité et l'occurrence des événements hydrologiques. Pour ce faire, l'analyse de données fonctionnelles (FDA) est utilisée pour calibrer et évaluer la performance du modèle hydrologique. À travers cet objectif, on expose les avantages à exploiter la FDA pour analyser et conserver toute la variabilité de l'hydrogramme. Dans cette partie de la thèse, on intègre les techniques et outils de la FDA pour définir les fonctions objectifs utilisées avec la procédure de calibration définie lors du premier objectif spécifique.

Le troisième objectif traite la même problématique que le deuxième et vise à comparer l'approche de calibration développée et les approches de calibration utilisées dans la littérature. Cet objectif vise aussi à faire une première analyse d'incertitude liée à l'estimation des paramètres par les approches considérées. La procédure définissant l'espace optimal développée au premier objectif est utilisée à cet effet.

Le quatrième objectif spécifique traite les problématiques liées aux lacunes des modèles hydrologiques conceptuels en ce qui a trait à la représentation adéquate des processus physiques et des états internes hydrologiques bien qu'il soit performant à simuler les débits. Basé sur les procédures et techniques développées lors des deux premiers objectifs, cet objectif vise à maximiser le degré de réalisme du modèle, toujours dans les limites imposées par le cadre conceptuel, via la définition et la considération de critères reflétant certains mécanismes hydrologiques du bassin versant en question.

Le cinquième objectif spécifique est de développer une stratégie de calibration multisite pour améliorer la performance du modèle à l'échelle du bassin ainsi que de réduire le biais des paramètres estimés qui peut découler du choix d'un site pour la calibration. On compare aussi la performance de cette calibration à celle d'une calibration locale (calibration sur un seul site) sur les différents sites considérés. À travers cet objectif, on explore aussi la variabilité spatiale des paramètres dans le bassin versant du lac St-Jean.

3. Modèle hydrologique CEQUEAU

Toutes les approches de calibration proposées sont appliquées au modèle hydrologique déterministe, conceptuel et semi-distribué CEQUEAU (Charbonneau et al., 1977; St-Hilaire et al., 2000; Morin et Paquet, 2007). Ce modèle hydrologique a fait l'objet de plusieurs études (ex. Ayadi and Bargaoui, 1998 ; St-Hilaire, 2000 ; St-Hilaire et al. 2000, 2003 ; Eleuch et al., 2010; Diaz Mercado et al., 2015 ; Arsenault et al., 2016 ; Boisvert et al., 2016).

Ce modèle est actuellement utilisé par le partenaire industriel du projet, Rio Tinto, pour des prévisions hydrologiques dans certaines régions du Canada (voir section 4. zones d'étude). CEQUEAU a été reprogrammé sous forme modulaire en C++ par Rio Tinto, ce qui a permis d'intégrer de nouveaux outils au modèle. Dans cette section, on fournit une description de la structure de base du modèle. La description détaillée de la version originale du modèle est disponible dans Morin et Paquet (2007) et la nouvelle version dans St-Hilaire et al. (2015). À noter que la nouvelle version fournie par Rio Tinto est utilisée dans cette thèse.

3.1 Schématisation du bassin versant et données météorologiques

À fin de tenir compte des caractéristiques physiques et leur variation spatiale dans le bassin versant, le modèle CEQUEAU requiert un double découpage du bassin. Le premier découpage consiste à superposer une grille délimitant des surfaces élémentaires appelées carreaux entiers (voir Figure 2-2). Le second consiste à subdiviser chaque carreau entier en un maximum de quatre carreaux partiels selon les lignes de partage. Ce double découpage permet au modèle de simuler l'évolution de l'écoulement de l'eau dans l'espace, simuler le débit en n'importe quel point au bassin versant, et représenter les modifications artificielles de l'écoulement dans les rivières (ex. barrage).

Pour chaque carreau entier, le modèle requiert le pourcentage de forêt, pourcentage de lacs et marais et altitude du coin sud-ouest. Les données requises pour chaque carreau partiel sont sa superficie (sous forme de pourcentage du carreau entier), et le sens de l'écoulement. Le sens de l'écoulement transitant sur chaque carreau partiel est obtenu à partir des données topographiques telles que dénivelée du terrain, ruisseau et rivière.

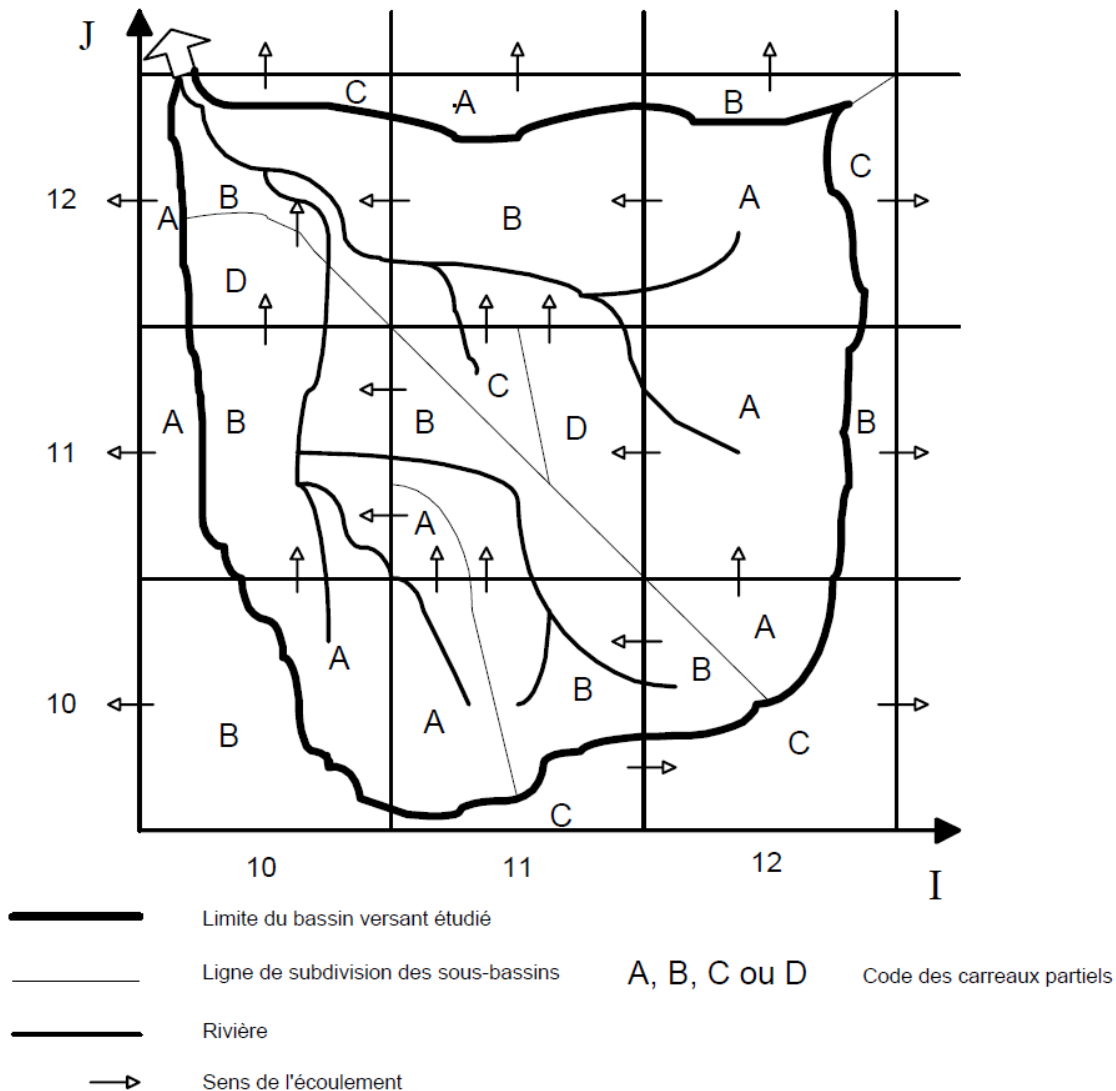


Figure 2-2 Subdivision d'un bassin versant en carreaux entiers et carreaux partiels selon les lignes de partages des eaux (Morin et Paquet, 2007)

Les données physiographiques et la schématisation d'un bassin versant peuvent être obtenues soit à partir de cartes topographiques ou à l'aide d'un système d'information géographique (SIG). Un module de préparation de données physiographiques à l'aide d'un SIG pour la nouvelle version de CEQUEAU est disponible dans Dugdale et al., (2017).

Le modèle hydrologique requiert une distribution spatiale des données météorologiques de précipitations journalières, températures maximales et minimales journalières. Ces données sont obtenues pour chaque carreau entier par interpolation à partir de stations météorologiques. Deux

méthodes sont disponibles pour l'interpolation de ces données : par polygones de Thiessen ou par pondération des trois stations les plus proches.

3.2 Simulation du débit

CEQUEAU simule l'écoulement de l'eau à l'aide de deux fonctions mathématiques appelées fonction de production et fonction de transfert. La première, présentée dans la figure 2-3, simule de manière simple et réaliste le processus que peut subir toute goutte d'eau tombée dans le bassin : formation et fonte du stock de neige, évapotranspiration, écoulement dans la zone saturée et non saturée, stockage d'eau dans les lacs et marécages. Elle représente, alors, l'écoulement vertical de l'eau à partir de la source atmosphérique. La fonction de transfert simule l'écoulement horizontal de l'amont vers l'aval.

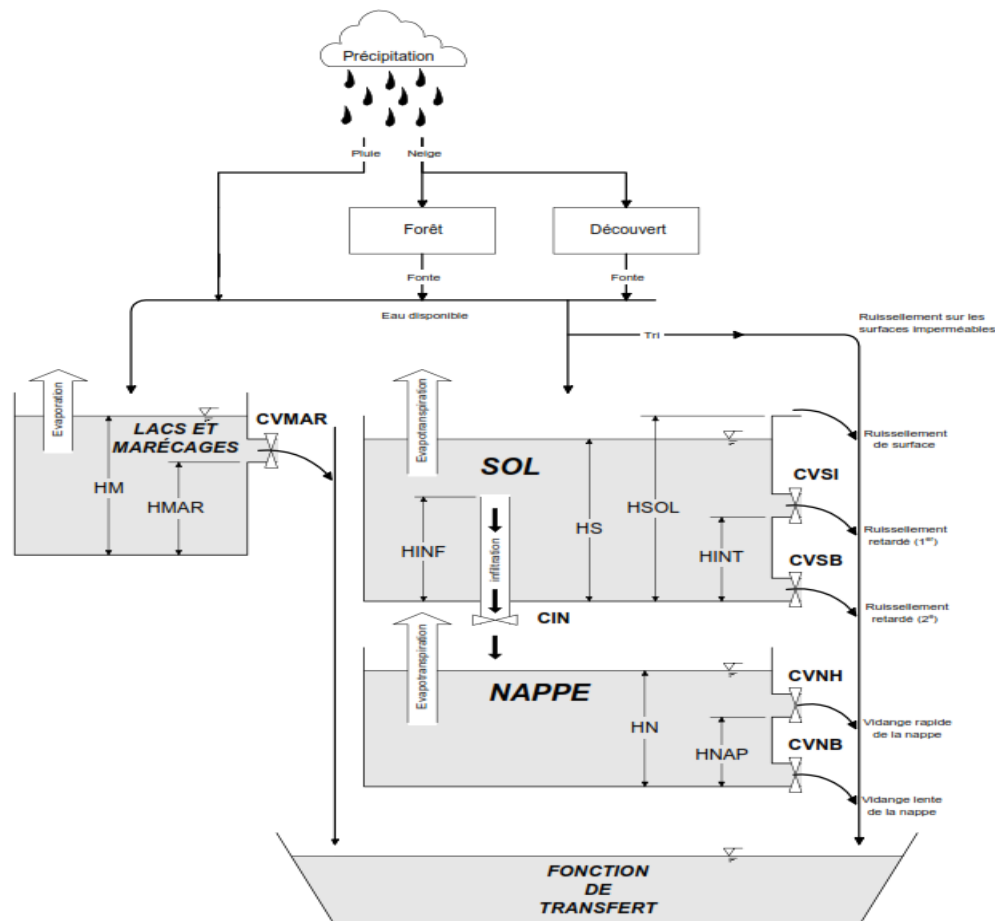


Figure 2-3 Schéma de la fonction de production du modèle CEQUEAU (Morin et Paquet, 2007).

Pour chaque carreau entier et pour chaque pas de temps journalier, la fonction de production calcule un bilan hydrique en considérant le sol comme des réservoirs interconnectés avec la formule suivante :

$$Q_t = P_t - ETP_t + (HN_t - HN_{t-1}) + (HS_t - HS_{t-1}) \quad (2.1)$$

Avec Q (mm) : débit total, P (mm) : précipitation liquide ou solide, ETP (mm) : évapotranspiration, HS (mm) : eau accumulée dans le réservoir SOL, HN (mm) : eau accumulée dans le réservoir NAPPE, t : pas de temps.

L'eau de pluie (précipitations liquides) est prise en compte immédiatement dans l'Eq. 2.1 alors que l'eau des précipitations solides provient de la fonte du manteau nival. Le modèle CEQUEAU effectue un bilan de fonte des stocks de neige sur chaque carreau entier séparé en zone de forêt et zone de clairière. La lame de fonte globale est la somme des lames produites, dans les deux zones, pondérées en fonction des surfaces de chaque zone. Pour le calcul de la fonte réelle, le module de fonte dans CEQUEAU tient compte du mûrissement du stock de neige et de la capacité du manteau nival à absorber les précipitations liquides. L'ETP est calculée à l'aide de la formule de Thornthwaite modifiée pour tenir compte de la durée potentielle d'ensoleillement en fonction de la latitude et du jour de l'année ainsi que de l'eau disponible dans la couche superficielle du sol (Morin et Paquet, 2007).

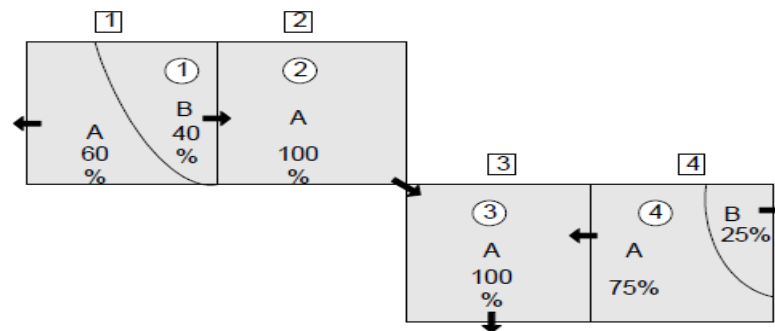
L'infiltration du réservoir Sol vers le réservoir Nappe n'est pas explicitement considérée dans le calcul du volume disponible sur chaque carreau entier à l'aide de l'équation 2.1. En revanche, cette composante est considérée pour déterminer la quantité d'eau accumulée dans les deux réservoirs (voir Figure 2-3).

La lame d'eau disponible pour chaque carreau partiel est calculée en fonction de sa superficie en ajoutant les volumes d'eau entrants des carreaux en amont (voir Figure 2-4). Ce volume d'eau est finalement acheminé d'un carreau partiel à l'autre jusqu'à l'exutoire grâce au coefficient de transfert donné par l'équation :

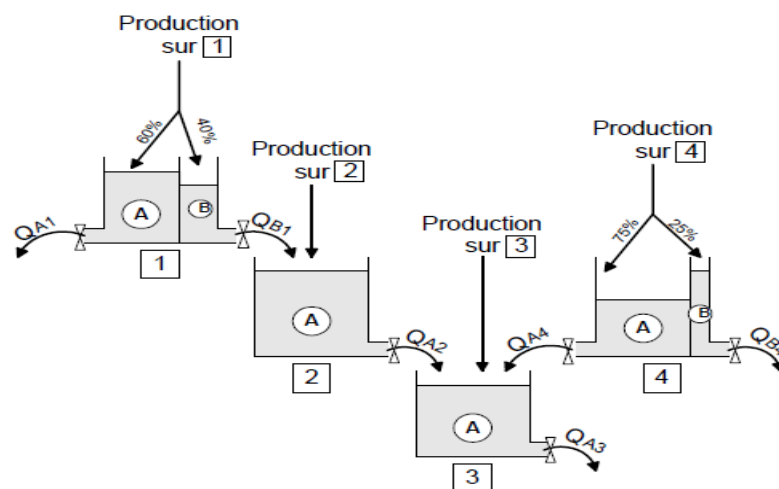
$$XKT_i = 1 - \exp\left(-\frac{EXXKT \cdot RMA3_i}{\max(SI, SI_{ac})} * \frac{100}{CEKM2}\right) \quad (2.2)$$

Avec XKT_i : coefficient de transfert du carreau partiel i , $EXXKT$: paramètre d'ajustement des coefficients de transfert d'un carreau partiel à l'autre, $RMA3_i$ (km²) : cumul des pourcentages de superficies des carreaux partiels en amont du $i^{\text{ème}}$ carreau, SI : superficie d'eau en amont du

carreau partiel, Slac (km²): pourcentage de superficie d'eau sur le carreau partiel, CEKM2 (km²): superficie des carreaux entiers conformes.



A) Vue en plan d'un partie du bassin



B) Schéma du transfert en rivière

Figure 2-4 Schéma de la fonction de transfert du modèle CEQUEAU (Morin et Paquet, 2007).

Les paramètres du modèle qui nécessitent l'ajustement régissent les processus de formation et fonte de neige, l'écoulement dans les zones saturées et non saturées du sol (réservoirs SOL et NAPPE), l'évapotranspiration, et l'acheminement de l'eau jusqu'à l'exutoire. Le tableau 2-2 présente la liste de ces paramètres avec l'intervalle de valeurs possibles fixé pour chaque paramètre.

Table 2-2 Liste des paramètres du modèle CEQUEAU et leurs descriptions.

Paramètre	Description (unité)	Intervalle	
STRNE	Seuil de transformation Pluie/neige (°C)	-2	3
TFC	Taux potentiel de fonte en forêt (mm/°C/jour)	0	10
TFD	Taux potentiel de fonte en clairière (mm/°C/jour)	0	10
TSC	Seuil de température de fonte en forêt (°C)	-2	3
TSD	Seuil de température de fonte en clairière (°C)	-2	3
TTD	Coefficient de déficit calorique	0	3
TTS	température du mûrissement du stock de neige (°C)	-5	2
CIN	Coefficient d'infiltration dans le réservoir Nappe	0	1
CVMAR	Coefficient de vidange du réservoir Lac et Marrais	0	1
CVNB	Coefficient de vidange bas du réservoir Nappe	0	1
CVNH	Coefficient de vidange haute du réservoir Nappe	0	1
CVSB	Coefficient de vidange basse du réservoir Sol	0	1
CVSI	Coefficient de vidange haute du réservoir Sol	0	1
XINFMA	Infiltration maximale par jour (mm/jour)	0	40
HINF	Seuil d'infiltration du réservoir Sol vers le réservoir Nappe (mm)	0	100
HINT	Seuil de vidange intermédiaire du réservoir Sol (mm)	0	100
HMAR	Seuil de vidange du réservoir Lac et Marrais (mm)	200	500
HNAP	Seuil de vidange supérieur du réservoir Nappe (mm)	20	500
HPOT	Seuil de prélèvement d'eau à taux potentiel par évapotranspiration (mm)	0	80
HSOL	Hauteur du réservoir Sol (mm)	100	300
HRIMP	Lame d'eau nécessaire pour que débute le ruissèlement sur les surfaces imperméables (mm)	0	10
EVNAP	Fraction d'évapotranspiration prise du réservoir Nappe	0	1
EXXKT	Paramètre d'ajustement des coefficients de transfert d'un carreau partiel à l'autre	0	5
ZN	Temps de concentration du bassin (jour)	0.1	20
TRI	Fraction de surface imperméable des carreaux entiers	0	1
COEP	Coefficient de correction des précipitations annuelles en fonction de l'altitude (mm/m/an)	0	1
COET	Correction des températures en fonction de l'altitude (°C/1000m)	-5	0
XAA	Exposant de la formule de Thorntwaite	0	2

4. Zones d'étude

Pour tester la robustesse des méthodes développées face à différentes caractéristiques hydrométéorologiques, le modèle hydrologique CEQUEAU (décrit dans la section 3) est calibré sur trois bassins localisés dans différentes régions du Canada. Les deux premiers sont des bassins pour lesquels des prévisions hydrologiques sont requises quotidiennement par le partenaire industriel du projet, Rio Tinto, alors que le troisième est un bassin versant forestier situé au Nouveau-Brunswick. Les caractéristiques de ces bassins sont résumées dans le tableau 2-3. On peut y constater que les trois bassins diffèrent par leur taille, leur climat et leur orographie.

4.1. Bassin versant du Lac St-Jean

Le bassin versant du lac St-Jean est situé dans la région du Saguenay-Lac Saint Jean au nord-Est de la province du Québec et couvre un territoire de 73 800 km². Il draine ses eaux vers le lac St-Jean, considéré comme le troisième plus grand lac dans la province. On considère trois sous-bassins de ce même bassin : le sous-bassin du lac St-Jean avec une superficie de 45 261 km², celui des Passes-Dangereuses au nord avec une superficie de 11 485 km² et celui de la Chute-Du-Diable, ayant une superficie de 9 701 km². La région est caractérisée par une forte couverture forestière, occupant 90% du territoire. La Figure 2-5 présente la localisation des sous-bassins Passes Dangereuses et Chute-Du-Diable dans le grand bassin du Lac St-Jean, considérés pour le premier objectif de la thèse. Les débits observés en ces bassins ainsi que ceux à l'exutoire du bassin du lac St-Jean sont des débits calculés par bilan hydrique. Huit autres sous-bassins, dont les caractéristiques sont présentées dans le tableau 2-3, sont considérés pour le quatrième et cinquième objectif de la thèse. Les débits observés issus de ces sous-bassins sont des mesures direct en rivière.

Les données météorologiques et hydrométriques pour ce bassin proviennent de stations opérées par Rio Tinto (voir Figure 2-5). Les températures moyennes mensuelles varient entre -16.8 °C (janvier) et 18.8°C (juillet). Les précipitations totales annuelles au sein du bassin du Lac St-Jean en entier varient de 837 mm à 1046 mm. Les précipitations en hiver sont majoritairement sous forme de neige. La pointe de crue occurrente au printemps est causée par la fonte de neige.

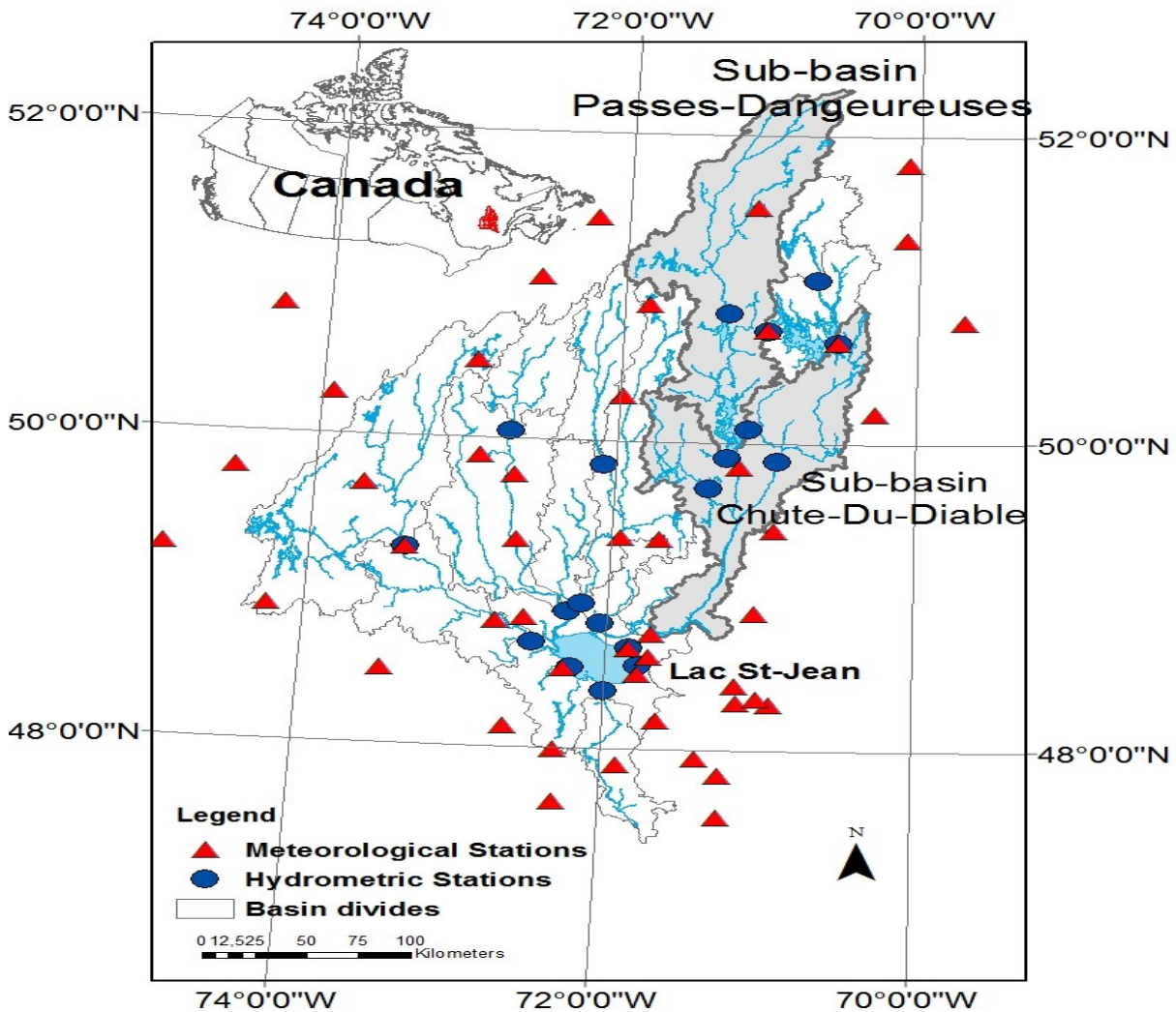


Figure 2-5 Carte du bassin versant du Lac St-Jean et la localisation des sous-bassins Passes-Dangeureuses et Chute-Du-Diable.

4.2. Bassin versant de la rivière Nechako

Le bassin versant de la rivière Nechako est situé dans la province de la Colombie-Britannique avec une superficie de 47 200 km² (Figure 2-6). Ses deux principaux affluents sont les rivières Nautley et Stuart qui doublent son volume d'eau (Macdonald et al., 2011). En 1951, La compagnie d'Aluminium du Canada (maintenant Rio Tinto) a commencé la construction du barrage de Kenney incluant une chaîne de rivières et de lacs pour former le réservoir du Nechako (Macdonald et al., 2007). Les volumes d'eau de ce réservoir se déversent majoritairement dans

une centrale électrique située à Kemano via un tunnel de 792 m de long et avec une chute de 792m d'altitude (voir Figure 2-6).

Les débits observés utilisés dans le premier objectif de la thèse sont des débits calculés par bilan hydrique dans le réservoir du Nechako, tandis que les débits observés dans le deuxième objectif de la thèse sont les débits observés en la station Vanderhoof (voir Figure 2-6 ci-dessous).

Les données météorologiques sont issues de stations opérées par Rio Tinto par interpolation par pondération des stations les plus proches avec corrections en fonction de l'altitude et des précipitations moyennes. La température moyenne maximale dans ce bassin atteint 19.8°C (en août) et la température moyenne minimale atteint -11.8°C (en janvier). Les précipitations moyennes annuelles varient entre 369 et 901mm. Les précipitations hivernales sont sous forme de neige ce qui aboutit à un débit de pointe occurrent au printemps par la fonte de neige dans les affluents non régulés du Nechako.

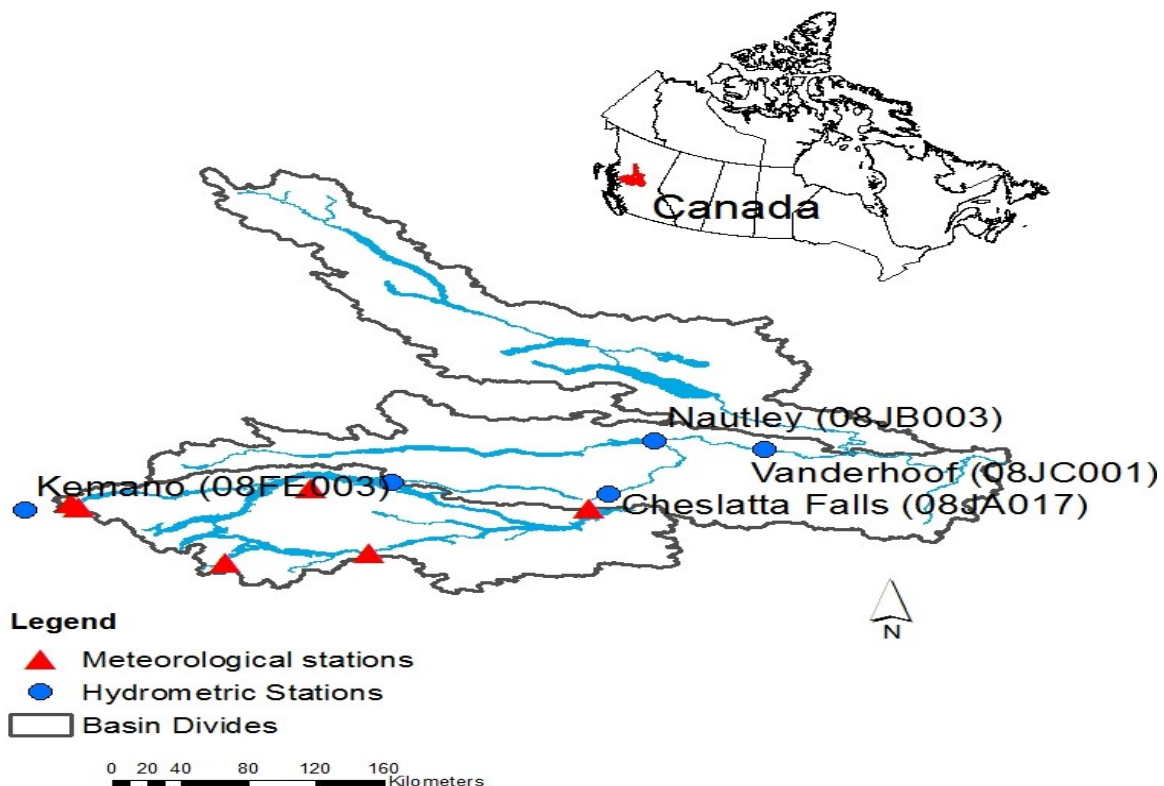


Figure 2-6 Carte du Bassin de la rivière Nechako.

4.3. Bassin versant du ruisseau Catamaran

Le troisième bassin est le bassin du ruisseau Catamaran qui est un petit bassin forestier de 52 km² (Figure 2-7). Ce bassin, situé dans la province du Nouveau-Brunswick draine ses eaux vers la rivière Little Southwest Miramichi dans le centre de la province (St-Hilaire et al., 2000). L'altitude varie entre 120 et 343m. Ce bassin versant est couvert majoritairement par des conifères(65%).

Les données météorologiques et hydrométriques proviennent de la station opérée par Pêches et Océans Canada située au milieu du bassin versant. Les précipitations annuelles observées dans ce bassin varient entre 867mm et 1412mm. Les températures moyennes mensuelles varient entre 11.8°C (en janvier) et 18.8°C (en juillet). Les pointes de débits sont principalement causées par la fonte de neige accumulée en hiver. Toutefois, les précipitations automnales contribuent aussi à l'écoulement en ce bassin.

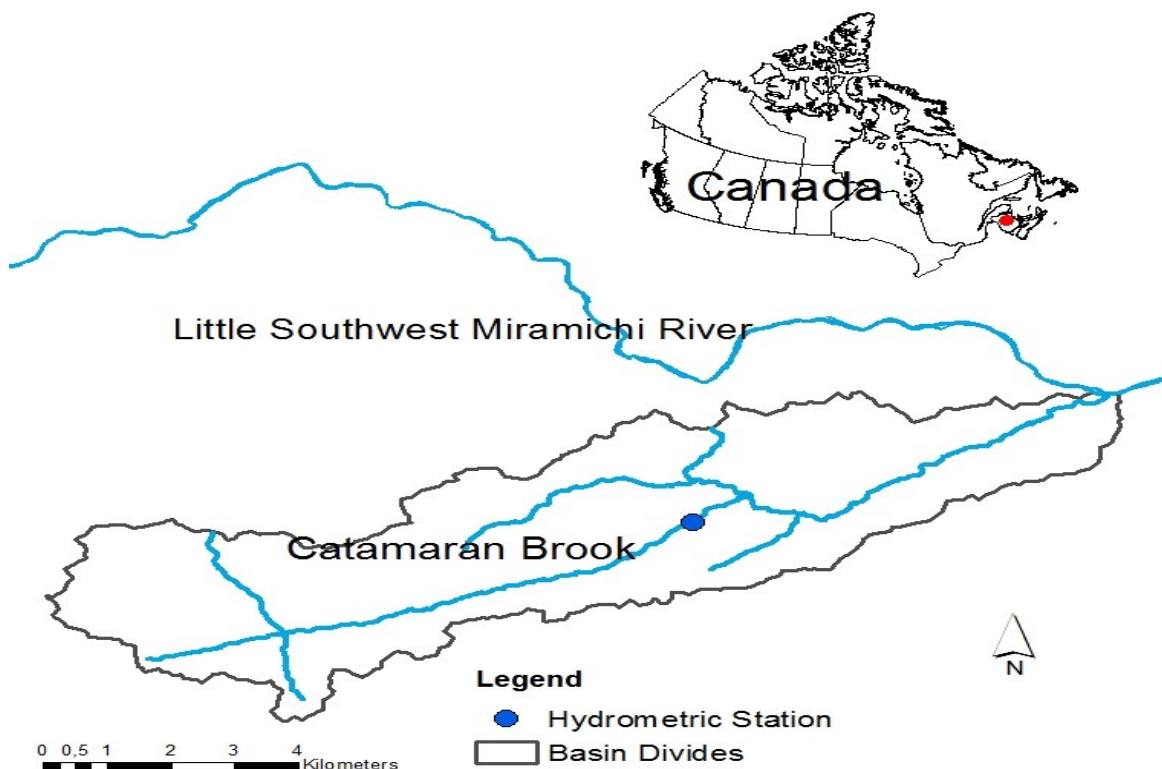


Figure 2-7 Carte du bassin du ruisseau Catamaran (adaptée de Caissie et al., 1998).

Table 2-3 Caractéristiques des bassins versants considérés.

Bassin	Lac St-Jean	Chute-Du-Diable	Passes-Dangereuses	Nechako	Catamaran
Superficie (km ²)	45 261	9 701	11 485	47 200	52
Altitude (m)	88-792	215-615	450-914	986-1863	120-343
Précipitation moyenne annuelle (mm)	914	1000	1082	602	1180
Débit maximum (m ³ /s)	6 733	1 796	2 339	1 338	7.42
(Date)	(12-05-1960)	(12-05-1960)	(21-05-2014)	(05-06-2007)	(30-10-2003)
Débit maximum moyen annuel (m ³ /s)	4187	1072	1181	771	5.21
Débit minimum moyen annuel (m ³ /s)	108	20	10	9	0.034
Débit moyen annuel (m ³ /s)	878	204	250	190	0.54
Volume moyen annuel (m ³)	320727	74387	91447	69572	196
Période de données météo disponible	01-01-1985 au 31-12-2013	01-01-1985 au 31-12-2013	01-01-1985 au 31-12-2013	01-01-1985 au 31-12-2013	01-01-1985 au 31-12-2013

Les statistiques moyennes annuelles sont calculées pour la période de données météo disponible

Table 2-4 Caractéristiques annuelles des sous-bassins du lac St-Jean.

Station	Notation	Superficie (km ²)	Altitude (m)	Précipitation moyenne (mm)	Débit minimum moyen (m ³ /s)	Débit maximum moyen (m ³ /s)	Débit moyen (m ³ /s)	Volume moyen (m ³)
Lac St-Jean	LSJ	45 261	88-792	1 006	46	4 139	914	333 868
Ashapshuam-Aval	ASHA	11 318	108-715	757	85	1 353	307	112 006
Ashapshuam-Amont	ASHAM	15 489	298-671	1 046	118	1 022	304	94 294
Metabetchouane	METAB	2 199	230-780	1 022	14	288	52	19 061
Mistassibi 3	MISBI3	3 065	299-721	988	12	382	80	27 503
Mistassini	MISNI	9 581	125-669	931	58	1 140	222	75 751
Mistassini 2	MISNI2	2 554	319-674	976	6	546	60	21 669
Petite Peribonka	PPER	7 476	439-760	837	5	738	22	7 862
Mistassibi	MISBI	8 697	126-720	960	35	1 000	203	72 835

Les statistiques présentées sont des moyennes annuelles calculées entre 2002 et 2009.

5. Outils statistiques

Cette section expose brièvement le principe des fonctions de profondeur et l'analyse de données fonctionnelles. Ces deux éléments sont rarement utilisés dans des applications hydrologiques notamment dans le cadre de l'estimation des paramètres hydrologiques. Ces deux outils statistiques sont utilisés dans les travaux de cette thèse pour combler certaines problématiques de calage de modèle hydrologique. La fonction de profondeur est utilisée comme métrique de sélection de jeux de paramètres performants avec un algorithme d'optimisation. L'analyse de données fonctionnelles est utilisée pour combler les limites des mesures de performance classiques exposées dans la sous-section 1.2.2. Elle est utilisée alors comme outil d'évaluation d'un modèle hydrologique.

5.1. Fonction de profondeur

La notion de fonction de profondeur est introduite par Tukey (1975) comme mesure statistique pour ordonner un échantillon multivarié. Cette mesure quantifie entre autres la centralité d'un point x par rapport à un échantillon issu d'une loi de probabilité multivariée P . Zuo et Serfling (2000) définissent les quatre propriétés désirables d'une fonction de profondeur d comme :

- Invariance affine : $d(x, P)$ est indépendant du système de coordonnées ;
- Le maximum (point le plus profond) est atteint au centre : $d(x, P)$ est le maximum si et seulement si x est le point de symétrie ;
- Monotonie par rapport au point profond : $d(x, P)$ est décroissante au fur et à mesure qu'on s'éloigne du centre ;
- Nulle à l'infini : $d(x, P)$ tend vers 0 quand $\|x\|$ tends vers l'infini.

Il existe différentes fonctions de profondeur dans la littérature, ex. la profondeur de Tukey ou demi-plan (Tukey, 1975), celle de Mahalanobis (Mahalanobis, 1936), Oja (Oja, 1983), 'simplicial depth' (Liu, 1990). Zuo et Serfling (2000) présentent quatre types de fonctions de profondeur selon leurs structures. Les deux types les plus utilisés sont les fonctions de profondeur géométriques, qui utilisent des demi-plans, simplex, etc., pour calculer la profondeur d'un point (ex. Tukey, simplicial depth), et les fonctions de profondeur basées sur une distance (ex. profondeur de Mahalanobis, 'projection depth').

La profondeur de Tukey est définie comme le minimum de points de P qui se trouve d'un côté d'un demi-plan où la frontière passe par x :

$$d_{HS}(x, P) = \min_{n_h} \{ \min[|(p \in P \text{ et } \langle n_h, p - x \rangle > 0)|], \min[|(p \in P \text{ et } \langle n_h, p - x \rangle < 0)|] \} \quad (2.3)$$

où $\langle x, y \rangle$ est le produit scalaire, n_h est vecteur normal d'un hyperplan.

La profondeur de Mahalanobis, basée sur la distance de Mahalanobis (Mahalanobis, 1936), est définie par :

$$d_{MH}(x, P) = \frac{1}{1 + (x - \mu)' \Sigma^{-1} (x - \mu)} \quad (2.4)$$

où μ est le vecteur moyen de la distribution P , Σ^{-1} inverse de la matrice de covariance de P .

Les fonctions de profondeur de type géométrique tel que la fonction de profondeur de Tukey ont l'avantage d'extraire de l'information plus robuste que les profondeurs basées sur une distance tel que la profondeur de Mahalanobis. Le premier type de profondeur semble être un choix intéressant, mais au détriment d'une charge computationnelle importante parce qu'elles utilisent des projections dans tous les plans unidimensionnels possibles (Zuo et Serfling, 2000). Cette charge computationnelle augmente avec la dimension (le nombre de variables) ou la taille de l'échantillon. Utiliser un algorithme d'optimisation pour calculer la profondeur devient nécessaire. Par exemple, Chakrapoty et Chaudhuri (2006) proposent l'algorithme '*Simulated Annealing*' pour calculer la profondeur de Tukey. La fonction de profondeur de Mahalanobis est un choix alternatif prometteur (Zuo et Serfling, 2000). Cette profondeur est la plus simple à calculer et son calcul est exact pour n'importe quelle dimension (pas besoin d'approximation). Pour ces raisons, cette fonction de profondeur est utilisée dans le cadre des travaux de cette thèse pour estimer les paramètres d'un modèle hydrologique.

Cette notion de fonction de profondeur a récemment été utilisée dans quelques études hydrologiques (ex. Chebana et Ouarda, 2008; Wazneh et al., 2013). Dans le cadre de la calibration de modèle hydrologique, la notion de fonction de profondeur est utilisée par Bardossy et Singh (2008) pour définir la structure géométrique des paramètres performants. Cette étude stipule que les paramètres performants (paramètres définissant un modèle performant) sont structurés dans l'espace. Les mêmes auteurs, dans une autre étude, utilisent la notion de fonction de profondeur pour définir les événements hydrologiquement inhabituels (Singh et Bardossy, 2012).

5.2. Analyse de données fonctionnelles

L'analyse de données fonctionnelles (FDA) c'est tout un cadre d'analyses statistiques constitué d'un ensemble d'outils statistiques permettant d'analyser une série de données discrètes observées sur un intervalle régulier (temps, espace, etc.) en une seule entité sous forme de fonctions continues (courbes). Les objectifs de la FDA sont les mêmes que les analyses statistiques classiques telles l'exploration des données, la comparaison des échantillons, la régression. Ramsay et Silverman (2002) exposent plusieurs applications de la FDA dans quelques domaines telles l'économie et l'archéologie. Récemment, cet outil est introduit en hydrologie par Chebana et al. (2012) comme outil pour l'analyse exploratoire des hydrogrammes et la détection des valeurs (courbes) aberrantes. Ternynck et al. (2015) utilisent la FDA pour la classification des hydrogrammes. Récemment, Masselot et al. (2016) exploitent la FDA pour la prédiction des courbes de débits saisonnières grâce aux modèles de régression fonctionels. Ces études exposent les avantages à utiliser la FDA pour conserver la forme entière et la variabilité de l'hydrogramme contrairement à l'approche multiobjective qui nécessite une sélection préalable des caractéristiques à étudier.

Dans le cadre de cette thèse, la FDA est utilisée pour évaluer la performance du modèle à reproduire les hydrogrammes observés où chaque hydrogramme est considéré comme une seule observation (courbe) au lieu de 365 observations journalières.

La première étape de la FDA est la transformation des observations brutes en fonctions continues ou courbes. Soit N échantillons de données $\mathbf{Y}_i = \{y_i(t_1), \dots, y_i(t_j), \dots, y_i(t_T)\}$ pour $i=1, \dots, N$ observées aux points discrets $\{t_j\}$ où $j=1, \dots, T$. Le principe de la FDA repose sur le fait que chaque observation $y_{i,j}$ est observé selon un processus continu tel (Ramsay et Ramsey, 2002) :

$$y_{i,j} = x_i(t_j) + \varepsilon_{i,j} ; \quad (2.5)$$

où $\varepsilon_{i,j}$ est l'erreur, $y_{i,j}$ est l'observation générée par le processus au point t_j . La fonction $x_i(t)$ peut être construite à partir d'une combinaison linéaire de fonctions indépendantes:

$$x_i(t) = \sum_{k=1}^K c_{i,k} \phi_k(t) = \mathbf{c}_i^T \boldsymbol{\phi}(t), \quad (2.6)$$

où $\mathbf{c}_i$ est le vecteur des coefficients de longueur K et $\boldsymbol{\phi}$ est le vecteur de K fonctions indépendantes.

Le choix d'une base (constante, polynomial, séries de Fourier, splines, ondelettes, etc.) dépend fortement du type de données. Les séries de Fourier sont le choix le plus naturel pour représenter des données périodiques alors que les splines sont généralement utilisées pour des données non périodiques (Levitin et al., 2007).

Pour estimer une fonction $x_i(t)$ générant les données discrètes, on considère une technique de lissage. La méthode la plus utilisée est la méthode de moindres carrée qui consiste à minimiser la somme carrée des erreurs (SSE). Les fonctions $x_i(t)$ sont alors estimées en minimisant:

$$SSE(y_i|c_i) = \sum_{j=1}^T (y_{i,j}(t_j) - x_i(t_j))^2, i=1, \dots, N. \quad (2.7)$$

La méthode des moindres carrée ne tient pas compte du degré de lissage des fonctions estimées. Pour remédier à cette contrainte, la méthode des moindres carrée pénalisée peut être utilisée. Cette procédure se base sur le même principe que la méthode des moindres carrée en ajoutant un terme de pénalisation à l'équation 2.7 :

$$PENSSSE_m(y_i|c_i) = \sum_{j=1}^T (y_{i,j}(t_j) - x_i(t_j))^2 + \lambda PEN_m(x_i) \quad i=1, \dots, N. \quad (2.8)$$

Où PEN_m est le terme de pénalisation définie comme l'intégrale de la dérivée d'ordre m de la fonction $x_i(t)$; λ est le paramètre de lissage à estimer. Ce paramètre contrôle le compromis entre le faible biais par rapport aux observations et la variabilité des fonctions estimées $\hat{x}(t)$ reflétée par le degré de lissage. En effet, quand $\lambda \rightarrow 0$, $\hat{x}_i(t)$ est plus fidèle aux données alors que quand $\lambda \rightarrow \infty$, $\hat{x}_i(t)$ est plus lisse. Ce paramètre est estimé soit avec la validation croisée ou la validation croisée généralisée (GCV). Cette dernière est plus rapide et tend à lisser plus. Elle consiste à choisir pour une gamme de valeurs de λ la valeur qui minimise le critère noté :

$$GCV_i(\lambda) = \frac{\sum (y_{i,j} - x_{i,\lambda}(t_j))^2}{[\text{trace}(I - S(\lambda))]^2} \quad (2.9)$$

où $S(\lambda)$ est la matrice de lissage :

$$S(\lambda) = \Phi[\Phi^T \Phi + \lambda R_m]^{-1} \Phi^T, \quad (2.10)$$

où $R_m = \int [D^m \phi(t)][D^m \phi^T(t)] dt$: matrice de pénalité et D^m est la dérivée d'ordre m .

La trace de $S(\lambda)$, qui est le degré de liberté $df(\lambda)$, correspond au nombre de fonctions de la base dans le cas d'un ajustement sans pénalisation.

Les statistiques descriptives de la FDA sont les mêmes que les statistiques classiques sauf que celle-ci sont sous forme de fonctions. Pour un échantillon de fonctionnels $x_i(t)$ définies sur le domaine continue D :

- La fonction moyenne, comme illustrée dans la Figure 2-8, s'écrit :

$$\bar{x}(t) = 1/N \sum_{i=1}^N x_i(t), \quad t \in T \quad (2.11)$$

- La fonction variance est donnée par la formule :

$$var_x(t) = \frac{1}{N-1} \sum_{i=1}^N (x_i(t) - \bar{x}(t))^2, \quad t \in D. \quad (2.12)$$

- La fonction (surface) covariance décrit la variation conjointe de deux fonctions autour de leur moyenne :

$$\sigma(s, t) = cov(x(s), x(t)) = \frac{1}{(N-1)} \sum_{i=1}^N (x_i(s) - \bar{x}(s))(x_i(t) - \bar{x}(t)). \quad (2.13)$$

Le test de Student pour les fonctionnelles compare les moyennes de deux échantillons de fonctionnelles pour chaque pas de temps t . La statistique du test $T(t)$ est donnée pour deux échantillons x_1 et x_2 définies dans le domaine D par :

$$T(t) = \frac{|\bar{x}_1(t) - \bar{x}_2(t)|}{\sqrt{\frac{1}{T} \{Var[x_1(t)] + Var[x_2(t)]\}}} \quad \forall t \in D. \quad (2.14)$$

Le fait que ces statistiques soient définies comme des fonctionnelles permet de comparer entre les échantillons et représenter la variabilité de cette même statistique en fonction du temps contrairement aux statistiques classiques données par une seule valeur fixe. Conséquemment, ces statistiques fonctionnelles sont adéquates pour analyser et conserver la forme des hydrogrammes, qui sont une variation du débit en fonction du temps. Par exemple, le test de Student fonctionnel permet de détecter les périodes où les échantillons sont différents. La Figure 2-8 présente un exemple de test de Student appliqué à deux échantillons de fonctionnelles dont les fonctions moyennes sont présentées dans la Figure 2-8 à gauche.

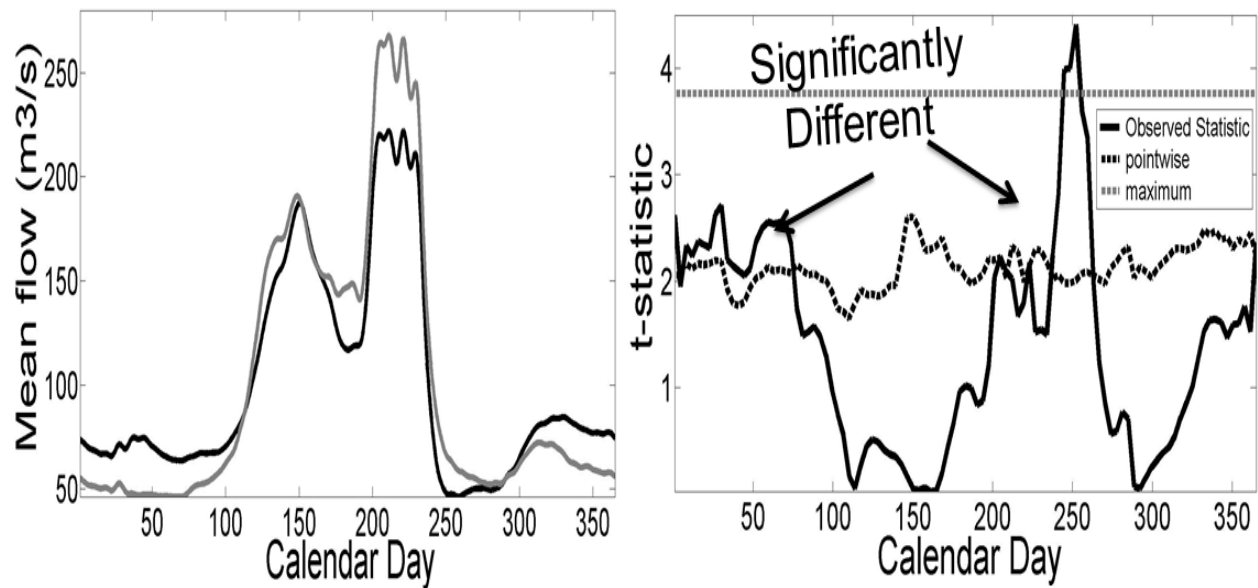


Figure 2-8 Fonctions moyennes (à gauche) et test de Student associé (à droite).

Le test rejette l'hypothèse que ces deux échantillons ont la même moyenne si la statistique de Student représentée par la courbe noire continue dépasse la p-value représentée par la courbe noire en pointillé. On constate alors dans cet exemple que ces deux échantillons sont différents en hiver (jours de l'année entre 0 et 75) et au début de l'été (jours 230 à 270).

6. Synthèse des travaux de recherche

Le présent chapitre offre une synthèse des méthodes et stratégies développées dans le cadre de cette thèse ainsi que les résultats obtenus. L'objectif global de ces méthodes est de calibrer un modèle hydrologique conceptuel performant, fiable et réaliste. Chaque partie traite une problématique reliée à un élément particulier du processus de la calibration automatique présenté dans la Figure 2-9 ci-dessous. On s'intéresse d'abord à définir une méthodologie de calibration efficace capable de fournir des paramètres performants et réalistes, ce qui est l'objectif 1. Les algorithmes proposés dans cette partie ayant pour objectif de mieux délimiter l'espace optimal des paramètres et trouver un optimum global sont utilisés dans la suite de la thèse. Ensuite, on s'intéresse à définir une fonction objectif permettant d'obtenir un modèle calibré capable de reproduire non seulement la pointe de la crue, mais aussi l'hydrogramme en entier (ex. les débits d'étiages et de crues). Cette approche est comparée dans la partie 3 aux autres approches conventionnellement utilisées dans la littérature. Ensuite, cette approche est exploitée dans la partie 4 proposant une stratégie de calibration multicritères par processus dans le but de définir des critères d'ajustement. L'objectif de cette partie est de définir un modèle qui est non seulement performant, mais reproduit aussi de manière réaliste les mécanismes hydrologiques dominants, conceptualisés dans le modèle. La dernière partie s'intéresse à une calibration globale multisite pour améliorer la performance du modèle à l'échelle du bassin versant du Lac St-Jean en utilisant les algorithmes développés dans la première partie.

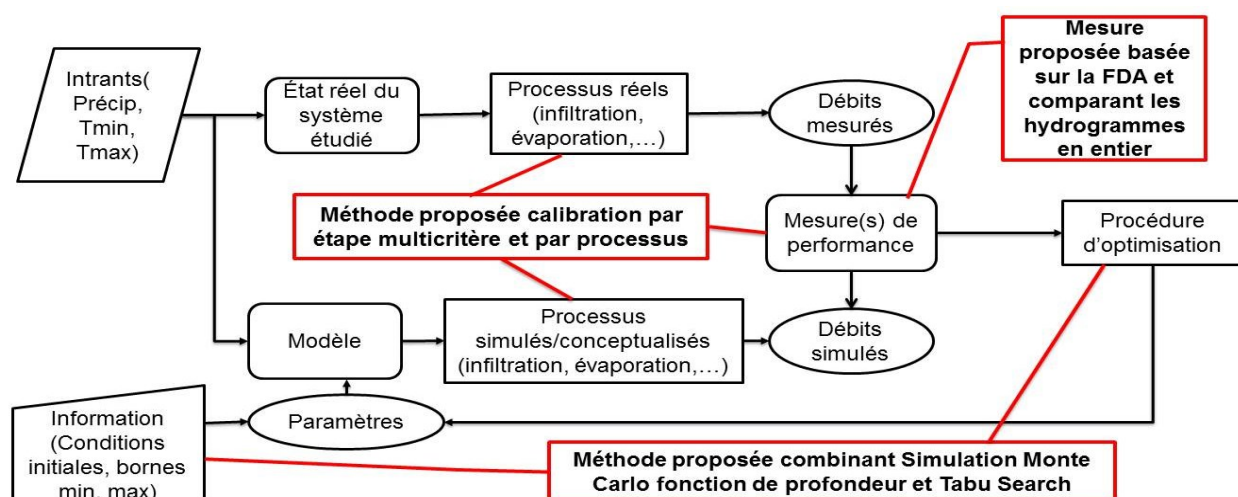


Figure 2-9 Schéma de la calibration automatique d'un modèle hydrologique et les méthodes proposées (en rouge) pour définir chaque élément.

6.1. Objectif 1 : Simulation Monte Carlo couplé au Tabu Search pour estimer les paramètres d'un modèle hydrologique

Cette partie est consacrée à la définition d'une méthode efficace et performante pour réduire le risque d'équifinalité d'une part et obtenir un jeu de paramètres ayant des valeurs réalistes. La méthode proposée s'échelonne sur deux étapes :

1. Définition d'un espace optimal de paramètres via une simulation Monte Carlo couplée avec une fonction de profondeur.
2. Ajustement des paramètres du modèle hydrologique dans l'espace défini en utilisant une métaheuristique : le 'Tabu Search'.

La première étape vise à réduire l'espace des paramètres selon les caractéristiques du bassin en question. Cet espace est exploré dans la seconde étape par le 'Tabu Search' pour définir un seul jeu de paramètres optimal.

6.1.1 Méthodologie

a) Étape 1 : Définition de l'espace optimal des paramètres

La méthode proposée pour définir l'espace optimal des paramètres est inspirée de l'algorithme 'Robust Parameter Estimation' ROPE proposé par Bardossy et Singh (2008). La méthode consiste en une simulation Monte Carlo couplée avec la notion de fonction de profondeur. À chaque itération un ensemble de simulations hydrologiques pour la période de calibration est effectué en utilisant des paramètres tirés aléatoirement de l'espace des valeurs possibles des paramètres à partir d'une loi uniforme. La fonction de profondeur est exploitée pour restreindre l'échantillonnage des paramètres dans la zone contenant des paramètres performants selon la mesure de performance considérée. Dans cette étude, on utilise la profondeur de Mahalanobis décrite dans la partie 5.1. Ce choix est plus adapté dans notre cas puisque son calcul est facile et exact pour n'importe quelle dimension, même si cette dernière est élevée (ici la dimension est 22). En plus, cette profondeur tient compte des corrélations éventuelles entre les paramètres via la matrice de covariance calculée dans la distance de Mahalanobis. Pour diminuer la taille de l'espace et définir la région de l'espace contenant des paramètres performant selon les caractéristiques du bassin d'intérêt, les bornes de valeurs possibles des paramètres sont mises à jour à chaque itération. L'algorithme est comme suit :

1- Fixer les bornes initiales des paramètres selon les valeurs possibles que peuvent prendre les paramètres (voir tableau 2-2).

2- Échantillonner uniformément un ensemble initial X de N jeux de paramètres.

3- Calculer la performance du modèle avec une mesure de performance (ex. NSE) et avec chaque vecteur N de paramètres de l'ensemble X .

4- Sélectionner l'ensemble X^* des meilleurs vecteurs de l'ensemble de X selon la mesure de performance (par exemple un pourcentage P de meilleurs vecteurs ; dans notre cas, on prend $P=10\%$).

5- Substituer les bornes de l'ensemble X^* aux bornes de paramètres.

6- Échantillonner uniformément un nouvel ensemble Y entre les nouvelles bornes et sélectionner uniquement les vecteurs Y^* profonds parmi X^* selon un seuil fixé (on prend comme seuil le premier quartile des profondeurs).

7- $X \leftarrow Y^*$; Répéter à partir de l'étape 3 jusqu'à non amélioration de la performance de Y^* . L'espace défini par l'ensemble final est l'espace optimal.

Pour quantifier le pourcentage de réduction de l'espace des paramètres et comparer le nouvel espace avec l'espace fixé au préalable, l'indice suivant est calculé pour chaque paramètre i :

$$G_i = \frac{optborneMax_i - OptborneMin_i}{borneMax_i - borneMin_i} \quad (2.15)$$

Où $optborneMax_i$ et $OptborneMin_i$ sont les limites de l'intervalle optimal, $borneMax_i$ et $borneMin_i$ sont les limites de l'intervalle initial du paramètre i défini dans le tableau 2-2.

b) Étape 2 : Estimation des paramètres du modèle

L'espace optimal défini est exploré davantage avec l'algorithme d'optimisation global 'Tabu Search', algorithme rarement utilisé en hydrologie, dans le but de trouver un optimum global. Le 'Tabu Search' (TS) ou la recherche taboue est une méta-heuristique basée sur la notion du voisinage et qui a été introduite par Fred Glover en 1986. Un voisinage d'une configuration s appartenant à l'espace E , noté $V(s)$, est défini comme un sous ensemble de l'espace E obtenu à partir d'une légère modification de celle-ci. Généralement, le voisinage est défini comme un

ensemble de mouvements v qui mènent d'une solution s aux nouvelles solutions du sous-ensemble $V(s)$. Le principe de l'algorithme TS consiste à évaluer à chaque itération le voisinage de la configuration courante s , ensuite la remplacer par la meilleure configuration de $V(s)$ jusqu'à satisfaction d'un critère d'arrêt. La particularité de cette recherche est l'utilisation d'une liste de type FIFO (First In First Out) nommée liste Taboue afin de mémoriser les mouvements effectués, qui sont considérés comme mouvements tabous, lors des K dernières itérations. Le but d'utiliser cette liste est d'éviter les cycles de taille K et aussi d'interdire à l'algorithme de ré-effectuer ces mouvements c.-à-d. de réévaluer les configurations déjà explorées. Toutefois cette liste peut engendrer des interdictions radicales dans le sens où elle peut interdire un mouvement qui peut mener à un optimum global. Pour cet effet, on utilise un critère d'aspiration qui permet d'accepter un mouvement bien que tabou sans le risque d'introduire un cycle. Pour une meilleure exploration de l'espace des paramètres, des techniques d'intensification et de diversification peuvent être intégrées. La première incite l'algorithme à explorer davantage les régions prometteuses alors que la seconde l'incite à explorer les régions non encore visitées.

On montre une version de base de l'algorithme TS qu'on a appliquée pour l'ajustement des paramètres du modèle hydrologique et présentée dans la figure 2-10 ci-dessous.

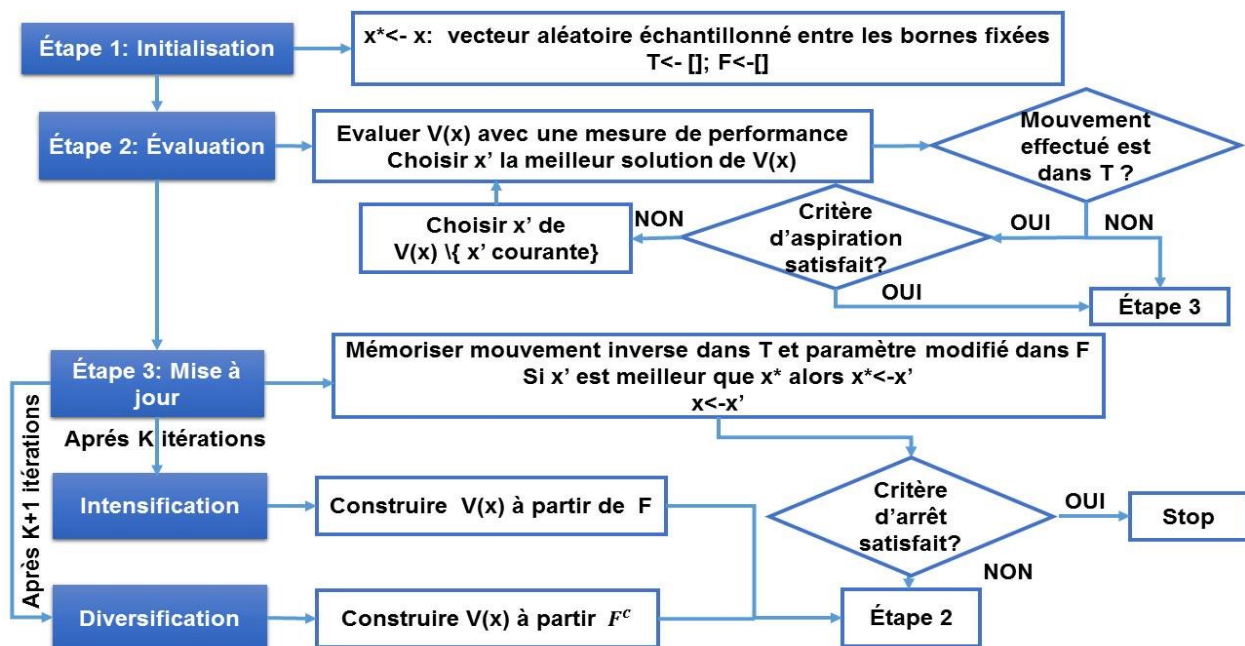


Figure 2-10 Organigramme du Tabu Search.

Les éléments du Tabu Search sont : x : combinaison de paramètres qui est la solution courante; x^* : optimum ; $V(x)$: voisinage de x ; Mouvement: perturbation d'un paramètre, défini comme un couple (indice du paramètre, signe) ; T : contient l'inverse des mouvements effectués ; F : couple (indice du paramètre, nombre de fois modifié) ; F^c : Complémentaire de F ; Critère aspiration: accepter un mouvement tabou si il conduit vers x' meilleur que x^* ; Critère d'arrêt: atteint nombre maximal d'itérations ou évaluation et non amélioration de l'optimum après i itérations.

L'algorithme TS évalue un jeu de paramètre x initial et évalue son voisinage $V(x)$ obtenu en appliquant l'ensemble des mouvements possibles et remplace x par le meilleur jeu de paramètres de $V(x)$. Un mouvement consiste à la modification (augmentation ou diminution) d'un paramètre i dans x par une fraction prédéfinie. La fraction est définie selon l'intervalle des valeurs possibles du paramètre en question. La modification d'un paramètre à la fois permet de maintenir le gain de performance dans les itérations précédentes. On caractérise chaque mouvement par l'indice du paramètre i et le signe de modification selon le fait que le paramètre est augmenté ou diminué. Dans la liste Taboue T , on mémorise l'inverse des mouvements effectués qui sont interdits dans les $|T|$ itérations suivantes. Toutefois, l'algorithme accepte un mouvement tabou si et seulement si il mène à un jeu de paramètres plus performant que l'optimum trouvé.

L'intensification incite l'algorithme à modifier uniquement les paramètres de la liste F . Quant à la diversification, elle incite l'algorithme à modifier les paramètres qui non pas été modifiés et donc ceux qui ne sont pas mémorisés dans la liste F . L'algorithme alterne alors entre phase d'apprentissage (collecter l'information sur les paramètres améliorant la performance du modèle), intensification (mieux définir les paramètres influents) et diversification (explorer les paramètres non influents).

c) Structure de calibration/validation

Deux structures de calibration/validation sont utilisées pour tester la robustesse des méthodes et la fiabilité du modèle calibré. La première structure est la structure de base, couramment utilisée dans la pratique: la validation croisée avec échantillon fractionné ('split-sample'). Celle-ci consiste à diviser l'échantillon en deux périodes égales. Ensuite l'une est utilisée pour la calibration et la seconde pour la validation. Le modèle est jugé acceptable s'il a des performances similaires lors des deux périodes. C'est une condition nécessaire et non suffisante pour accepter la performance et la fiabilité d'un modèle (Klemeš, 1986). La structure de

validation croisée différentielle ('Differential split-sample'), qui est la seconde structure utilisée, est plus puissante et permet d'évaluer la performance du modèle face à des conditions climatiques ou géographiques différentes que celles utilisées lors de la calibration (Klemeš, 1986; Kirchner, 2006). Le principe de cette structure consiste à utiliser deux périodes de données avec des conditions hydrologiques indépendantes et différentes. Dans notre cas, le modèle est calibré sur une période humide ensuite il est validé sur une période sèche et vice-versa. Cette structure permettra d'évaluer la performance du modèle à prédire des débits très élevés non observés en calibration et inversement.

d) Mesures de performance

Dans les deux étapes de la calibration, le critère Nash-Sutcliffe est utilisé comme mesure de performance. Ce critère est défini comme suit :

$$NSE = [1 - \frac{\sum_{i=1}^n (q_{ci} - q_{oi})^2}{\sum_{i=1}^n (q_{oi} - \bar{q}_o)^2}] \quad (2.16)$$

Avec q_{ci} et q_{oi} les débits simulés et observés au pas de temps i .

Pour la validation et l'évaluation de la performance du modèle calibré, en plus du NSE, le critère Kling-Gupta (KGE, Gupta et al., 2009) et le biais relatif sont utilisés :

$$Biais = \left| \frac{\sum_{i=1}^n (q_{ci} - q_{oi})}{\sum_{i=1}^n q_{oi}} \right| \quad (2.17)$$

q_{ci} et q_{oi} les débits simulés et observés au pas de temps i .

$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (2.18)$$

Avec $\alpha = \sigma_s / \sigma_o$ ratio entre l'écart type des simulations et observations. $\beta = \mu_s / \mu_o$: ratio entre la moyenne des simulations et des observations, et r : la corrélation linéaire entre simulations et observations.

Le KGE sert à comparer les deux moments (moyenne et écart type) de la distribution des simulations et observations ainsi que leur occurrence.

6.1.2 Résultats et discussion

La procédure exposée est appliquée aux bassins de Chute-du-Diable, Passes-Dangereuses, la rivière Nechako et le ruisseau Catamaran.

Pour illustrer la problématique concernant la convergence des algorithmes vers un jeu de paramètres irréalistes, les paramètres du modèle sont ajustés avec la structure de la validation croisée avec échantillon fractionné dans l'espace défini dans le tableau 2-2. L'algorithme TS converge vers des solutions ayant des valeurs NSE et KGE supérieures à 0.8 en période de calibration et validation pour les bassins Chute-Du-Diable, Passes-Dangereuses et le ruisseau Catamaran. Pour le bassin de la rivière Nechako, le NSE et le KGE en calibration et validation sont entre 0.6 et 0.8. Bien que ces modèles calibrés soient performants et acceptables, les valeurs fixées par l'algorithme pour certains paramètres sont irréalistes dans les zones d'étude. Par exemple, EVNAP (fraction d'évaporation prise du réservoir) fixé à 1 est traduit par une source d'évaporation provenant uniquement du réservoir nappe ce qui est irréaliste selon les conditions des zones d'étude.

L'optimisation de cet espace de paramètres a permis de le réduire de 50 % à 80% dans les quatre bassins. L'espace des paramètres les plus sensibles du modèle à savoir CVSB (Coefficient de vidange basse du réservoir Sol) et TFC (taux potentiel de fonte en forêt) est réduit de plus de 90% pour les quatre bassins. Les bornes minimales et maximales obtenues sont conceptuellement réalistes et correspondent aux conditions physiographiques de chaque bassin. En effet, les bornes obtenues sont différentes d'un bassin à l'autre. Par exemple, les bornes de TFC obtenues (voir Table 2-5) sont comprises entre 0.9 et 1.54 mm/°C/jour pour la rivière Nechako alors que pour le ruisseau Catamaran elles sont entre 2.6 et 3 mm/°C/jour. Ceci est corroboré par les observations qui démontrent que la fonte est plus rapide dans le bassin versant du ruisseau Catamaran que dans celui de la rivière Nechako.

Table 2-5 Bornes optimales obtenues après optimisation de l'espace initial avec la procédure proposée.

Paramètres	Bornes initiales		Bornes optimales							
			Chute-Du-Diable		Passes-Dangereuses		Rivière Nechako		Ruisseau Catamaran	
STRNE	-2	3	-1.11	-0.11	-1.6	-1.1	2.14	2.79	-0.31	0.42
TFC	0	10	2.14	3.04	2.0	2.7	0.90	1.54	2.63	3.02
TFD	0	10	3.31	7.11	5.8	8.5	2.79	6.70	6.37	7.61
TSC	-2	3	0.35	1.47	0.4	1.2	1.43	2.58	1.10	1.47
TSD	-2	3	-0.67	1.10	-1.4	0.0	-1.14	1.05	-0.87	0.06
TTD	0	3	1.23	2.25	1.2	2.3	0.73	2.41	1.21	1.96
TTS	-5	2	-3.51	-1.95	-3.7	-1.8	-1.67	0.88	-2.42	-1.22
CIN	0	1	0.31	0.69	0.2	0.7	0.27	0.85	0.46	0.64
CVMAR	0	1	0.29	0.64	0.2	0.6	0.22	0.79	0.37	0.60
CVNB	0	1	0.07	0.09	0.1	0.6	0.09	0.62	0.07	0.08
CVNH	0	1	0.28	0.68	0.3	0.7	0.25	0.80	0.28	0.57
CVSB	0	1	0.04	0.05	0.03	0.05	0.08	0.16	0.03	0.04
CVSI	0	1	0.20	0.58	0.2	0.5	0.16	0.71	0.46	0.59
XINFMA	0	40	12.89	26.51	11.2	26.9	11.58	30.95	14.90	20.41
HINF	0	100	57.03	80.74	67.0	83.6	19.86	73.25	56.29	63.32
HINT	0	100	63.12	84.18	67.6	88.5	38.83	86.28	53.53	60.41
HMAR	200	500	291.0	420.31	293.9	442.5	280.22	451.37	300.89	377.25
HNAP	20	200	89.97	152.26	73.7	154.6	53.55	157.18	115.11	145.92
HPOT	0	80	23.37	41.35	32.5	50.7	50.28	66.41	15.77	26.12
HSOL	100	300	142.4	213.00	156.0	235.5	137.50	241.25	165.19	209.94
HRIMP	0	10	3.27	7.03	2.8	6.8	2.36	7.76	3.02	5.87
EVNAP	0	1	0.34	0.70	0.3	0.7	0.22	0.80	0.38	0.66
Iteration			37		32		23		53	

L'algorithme a convergé vers un espace de paramètres où le NSE moyen est autour de 0.8 sauf pour le bassin de la rivière Nechako (voir Table 2-6). Pour ce bassin l'algorithme converge vers un espace dont le NSE moyen est de 0.31. Cette première étape de la calibration définit un espace optimal dont la taille dépend étroitement de la fonction de profondeur choisie. En effet selon la définition de la profondeur, un point profond pour une fonction de profondeur ne l'est pas nécessairement pour une autre.

Table 2-6 Performance des jeux de paramètres de l'espace optimal avec les quantiles 0.05 et 0.95 et l'écart type aux quatre bassins Chute-Du-Diable (CD), Passes-Dangereuses(PD), rivière Nechako (Nech) et ruisseau Catamaran (Cat) .

Bassin	NSE moyen	Quantile 0.05	Quantile 0.95	Écart Type
CD	0.86	0.83	0.88	0.015
PD	0.79	0.77	0.80	0.011
Nech	0.31	0.21	0.39	0.058
Cat	0.83	0.81	0.84	0.008

L'ajustement des paramètres avec le TS dans cet espace optimal, fournit des modèles calibrés performants avec un NSE et KGE supérieurs à 0.7 pour le bassin Passes-Dangereuses et le ruisseau Catamaran en période de calibration et validation. Pour le bassin Chute-Du-Diable les NSE et KGE atteignent 0.9 en période de calibration et validation. Pour le bassin de la rivière Nechako, le NSE et KGE sont supérieurs à 0.6. Les modèles calibrés sont performants sur les débits élevés, l'estimation de la crue et sont moins performants à simuler les débits d'étiages. Ceci est dû à la fonction objectif utilisée lors de la calibration, le NSE, qui est inadéquat pour représenter les débits d'étiages.

L'application de la procédure de calibration avec la structure de la validation croisée différentielle (calibration sur période sèche et validation sur période humide et vice versa) montre la robustesse du modèle dans les bassins Chute-Du-Diable et Passes-Dangereuses. Les valeurs du NSE obtenues (voir Table 2-7) pour le bassin Chute-Du-Diable sont entre 0.87 et 0.93, quelle que soit la nature de la période de calibration et validation (période sèche ou humide). Pour le bassin Passes-Dangereuses, le NSE du modèle calibré selon une période humide est supérieur 0.85 pendant les deux périodes de calibration et validation. Le KGE est de 0.9 pour les deux périodes de calibration et validation. Alors que pour le modèle calibré selon une période sèche, le NSE est de 0.75 et le KGE est de 0.72 en calibration et validation.

Table 2-7 Performance du modèle calibré avec la validation croisée différentielle aux trois bassins (Chute-Du-Diable (CD), Passes-Dangereuses(PD), rivière Nechako (Nech).

Bassin	Calibration sur période humide			Évaluation sur période sèche		
	NSE	Bias	KGE	NSE	Bias	KGE
CD	0.92	0.03	0.94	0.88	0.09	0.82
PD	0.85	0.01	0.92	0.87	0.03	0.91
Nech	0.60	0.09	0.78	0.44	0.21	0.63
Bassin	Calibration sur période sèche			Évaluation sur période humide		
	NSE	Bias	KGE	NSE	Bias	KGE
CD	0.93	0.04	0.93	0.87	0.08	0.88
PD	0.75	0.04	0.72	0.74	0.03	0.72
Nech	0.50	0.25	0.65	0.52	0.21	0.70

Globalement, le modèle a des performances similaires en calibration et validation. Autrement dit, le modèle est capable de reproduire des débits élevés (ou faibles) non observés durant la période de calibration. Le modèle calibré au bassin de la rivière du Nechako est moins performant, quelle que soit la nature des périodes de calibration avec un NSE entre 0.44 et 0.6 et KGE entre 0.63 et 0.78. Cette modeste performance du modèle sur ce bassin peut s'expliquer entre autres par les fortes incertitudes sur les intrants (précipitations, températures), ou les observations de débits en ce bassin. En effet, les coefficients de correction des précipitations annuelles et des températures en fonction de l'altitude n'ont pas été ajustés. Étant donné que ce bassin se caractérise par de hautes altitudes, l'ajustement de ces paramètres pourrait améliorer la performance du modèle en ce bassin. La structure de calibration basée sur la validation croisée différentielle n'est pas appliquée au bassin du ruisseau Catamaran, car on ne dispose pas d'assez de données pour le fractionner en deux périodes différentes.

6.1.3 Conclusion

La procédure de calibration proposée permet de réduire l'espace des paramètres du modèle en particulier les paramètres sensibles. L'espace optimal obtenu dépend fortement du bassin où toutes les valeurs inadéquates sont rejetées. Il a été démontré que le risque d'obtenir des paramètres conceptuellement irréalistes est considérablement réduit quand l'ajustement des paramètres est fait selon l'espace optimal défini avec la simulation Monte Carlo et la fonction de

profondeur. Globalement, le modèle calibré a des performances similaires en période de calibration et validation avec les deux structures de calibration. Toutefois, le modèle calibré est moins performant à simuler les débits d'étiage. Ceci s'explique d'une part par les incertitudes liées aux intrants et la structure du modèle. D'autre part, l'utilisation d'une mono-fonction objectif, qui est le NSE, impacte fortement la performance globale du modèle. Cette problématique est traitée dans la partie suivante avec l'objectif de mieux représenter l'hydrogramme en entier et remédier aux problématiques liées aux pertes d'informations.

6.2. Objectif 2 : Analyse de données fonctionnelles pour calibrer et évaluer un modèle hydrologique

Cette partie s'intéresse à l'introduction d'un nouveau concept pour la calibration des modèles hydrologiques basé sur l'analyse de données fonctionnelles. L'hydrogramme est une variation du débit en fonction du temps. L'utilisation des fonctions objectives classiques évalue la variabilité de manière réductrice en une seule valeur, ce qui aboutit très souvent à une performance du modèle centrée sur la capacité à représenter uniquement une partie de l'hydrogramme (ex. la crue). Grâce à la FDA, une nouvelle fonction objectif est définie comparant les hydrogrammes simulés et observés comme des fonctions continues. Ainsi, les débits de crue et d'étiages sont comparés en une seule étape.

À noter que dans la présente thèse le package FDA en MATLAB a été exploité pour utiliser les techniques de la FDA. Ce package est disponible au lien suivant : <http://www.psych.mcgill.ca/misc/fda/software.html>.

6.2.1 Méthodologie

La première étape de la FDA est la transformation des données brutes en fonctions continues $x_i(t)$. Comme les données représentent une certaine périodicité, les observations de séries annuelles de débits journaliers $\{y_{i,1}, y_{i,2}, \dots, y_{i,365}\}$ sont transformées en fonctions continues lissées via des séries de Fourier :

$$x_i(t) = c_1 + c_2 \sin \omega t + c_3 \cos \omega t + c_4 \sin 2\omega t + \dots c_K \cos K\omega t, \quad (2.19)$$

Où $\omega = \frac{2\pi}{T}$ avec T la période(dans notre cas $T=365$), k =nombre de fonctions de la base.

La méthode des moindres carrés pénalisés (présentée dans la partie 5.2) est utilisée pour estimer les coefficients c_k . Pour chaque bassin, le nombre de fonctions K et le paramètre de lissage λ sont fixés selon les observations de débits journaliers. Ensuite, les mêmes valeurs sont utilisées pour transformer les débits simulés par le modèle à chaque itération.

La Figure 2-11 ci-dessous présente un exemple de lissage des hydrogrammes observés au Lac St-Jean entre 2002 et 2009 pour $K=255$ et $\lambda =0.01$. La Figure 2-11a présente les débits journaliers observés qui montrent une certaine périodicité. Dans la Figure 2-11b, on remarque une bonne corrélation entre les débits observés et leurs valeurs lissées à travers le scatter plot. Les Figures 2-11c à 2-11f montrent que les hydrogrammes lissés s'ajustent adéquatement aux hydrogrammes observés et que les débits de pointe sont atteints.

D'autres valeurs de K ont été testées de 50 à 365. Les valeurs faibles de K n'ont pas permis d'atteindre les pointes de crues avec des courbes plus lisses, tandis que les valeurs proches de 365 permettent d'obtenir des courbes plus fidèles aux hydrogrammes observés. Toutefois, les valeurs élevées de K augmentent la complexité des courbes à ajuster vu le nombre de coefficients à estimer.

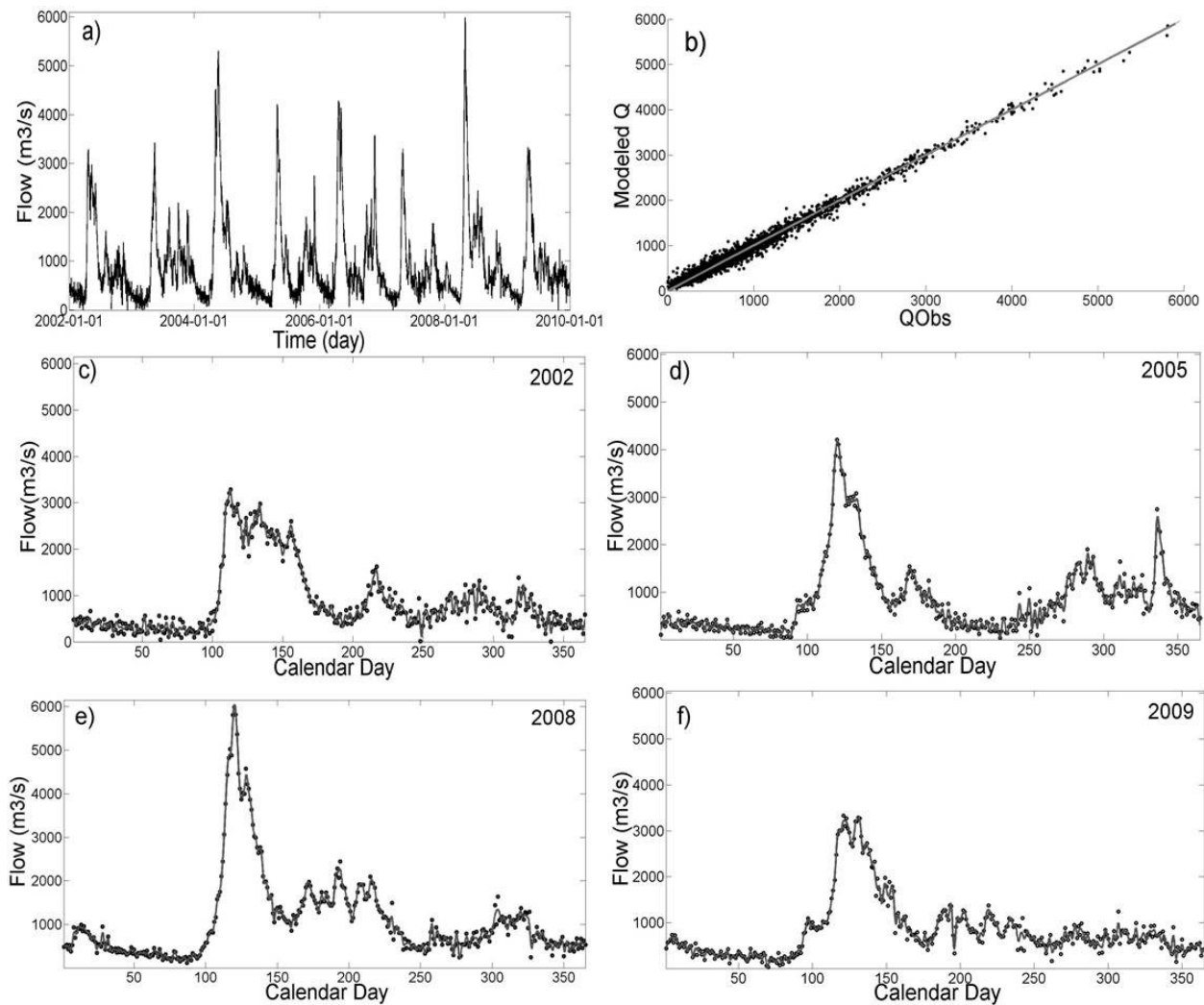


Figure 2-11 Exemple de lissage des hydrogrammes au lac St-Jean pour les débits observés entre 2002 et 2009. a) débits journaliers observés. b) Scatter plot entre les débits observés et les débits lissés. c)-f) hydrogrammes observés et lissés pour les années sélectionnées.

On compare à chaque itération les nouvelles données que sont des hydrogrammes lissés observés et simulés. On définit la fonction objectif comme la distance entre les hydrogrammes simulés et observés. On exploite aussi la dérivée de ces hydrogrammes pour inclure dans les barèmes de calibration, l'occurrence des évènements hydrologiques (période d'étiage, crue, récession) à travers le changement de pente, comme illustré dans la Figure 2-12 ci-dessous.

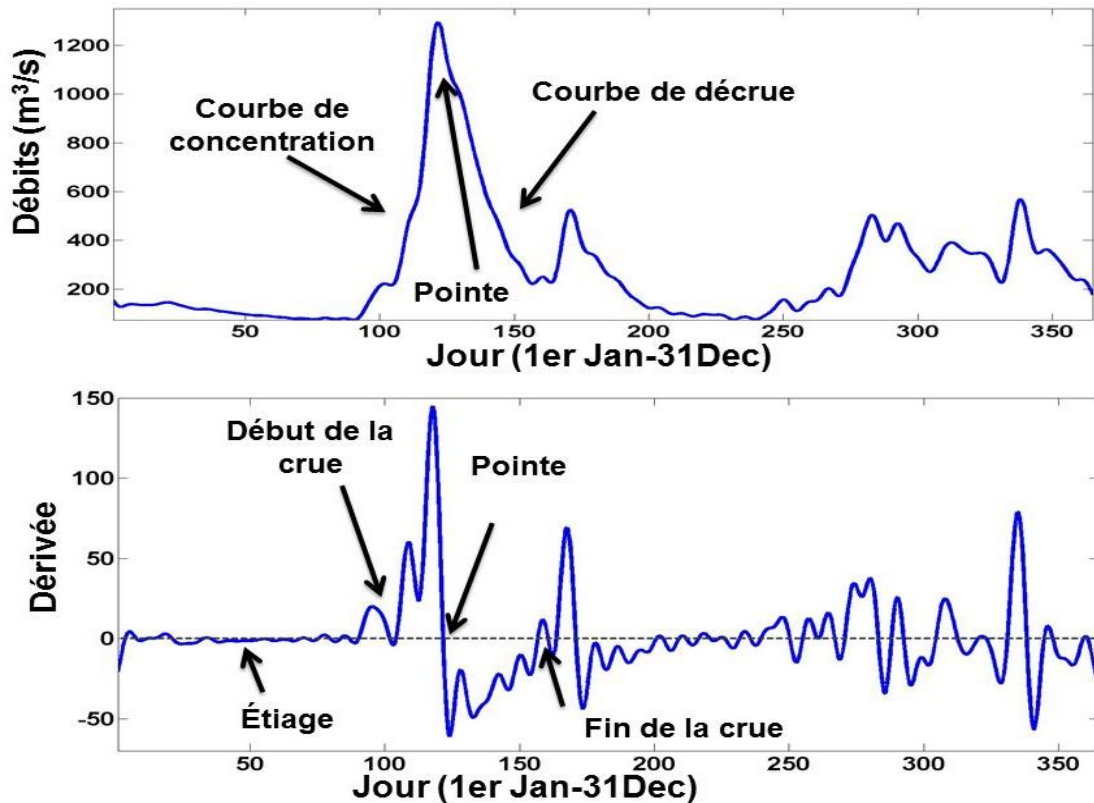


Figure 2-12 Illustration de l'extraction des événements hydrologiques à partir de la dérivée de l'hydrogramme.

Pour minimiser l'effet de la crue, la fonction objectif est pondérée par la moyenne de la statistique de Student fonctionnelle (voir sous-section 5.2). En effet, les jeux de paramètres ne reproduisant pas adéquatement les débits observés sur une période de l'année, leurs statistiques seront élevées. Par conséquent, ces jeux de paramètres sont rejetés puisque la pondération par ce critère amplifie la fonction objectif qui est à minimiser.

Comme la procédure de calibration développée lors de la première partie de la thèse est utilisée, deux fonctions objectives sont définies. La première fonction objectif, utilisée lors de l'étape 1 de la procédure de calibration (réduction de l'espace), minimise la distance entre l'hydrogramme moyen simulé et observé et leur dérivée. On cherche alors un sous-espace optimal où les jeux de paramètres sont capables au moins de reproduire l'hydrogramme moyen. La fonction objectif s'écrit :

$$M_{OF} = S * \int_{t=1}^{T=365} \{ |\bar{x}_{sim}(t) - \bar{x}_{obs}(t)|^2 + |\bar{x}'_{sim}(t) - \bar{x}'_{obs}(t)|^2 \} dt \quad (2.20)$$

avec S est la moyenne de la statistique du test de Student fonctionnel; $\bar{x}_{sim}$ et $\bar{x}_{obs}$ sont les hydrogrammes moyens simulé et observé, $\bar{x}'_{sim}$ et $\bar{x}'_{obs}$ sont leurs dérivées ; T est le nombre d'observations par année (ici $T=365$).

La deuxième fonction objectif utilisée avec l'algorithme du Tabu Search minimise la distance entre les hydrogrammes annuels simulés et observés à chaque année i . La fonction objectif est comme suit :

$$OF = S * \sqrt{\sum_{i=1}^N \{ \int_{t=1}^{T=365} |x_{i,sim}(t) - x_{i,obs}(t)|^2 + \int_{t=1}^{T=365} |x'_{i,sim}(t) - x'_{i,obs}(t)|^2 \} } \quad (2.21)$$

Avec: N est le nombre d'années; T est le nombre d'observations par année (ici $T=365$); $x_{i,sim}$ et $x_{i,obs}$ sont les hydrogrammes simulés et observés à la $i^{ème}$ année ; $x'_{i,sim}$ et $x'_{i,obs}$ sont les dérivées des hydrogrammes simulés et observés à la $i^{ème}$ année ; S est la moyenne de la statistique du test de Student fonctionnel.

Pour évaluer le modèle calibré, différents critères, notamment des signatures hydrologiques, sont calculés en plus du NSE. Ces signatures hydrologiques visent à évaluer différents aspects du régime hydrologique d'un bassin. On procède aussi à une analyse des courbes de débits classés simulés et observés. La courbe des débits classés permet d'évaluer la performance du modèle à reproduire des débits de différentes magnitudes. Pour cela, elle est décomposée en trois segments comme suggérés par Yilmaz et al. (2008). Le premier segment, avec des probabilités de dépassement entre $[0-0.2]$, représente la réponse du bassin aux précipitations et événements extrêmes. Le deuxième, avec des probabilités de dépassement entre $]0.2-0.8 [$, représente la réponse du bassin aux précipitations modérées. Le troisième segment caractérise la période d'étiage du bassin, avec des probabilités de dépassement entre $[0.8-1]$.

6.2.2 Résultats et discussion

Le modèle hydrologique CEQUEAU est calibré sur les trois bassins : Le lac St-Jean, la rivière Nechako et le ruisseau Catamaran. On compare la performance du modèle calibré avec la procédure développée en réponse au premier objectif en utilisant : 1) les deux fonctions objectives définies avec les équations 2.20 et 2.21 (noté OptFDA) ; et 2) calibration en utilisant comme seul critère le NSE (OptNSE). Le but est d'évaluer la capacité des fonctions objectifs définies à partir de la FDA à fournir la même qualité de simulation de la crue qu'avec le critère NSE tout en améliorant l'estimation d'autres composantes de l'hydrogramme.

La Figure 2-13 présente les résultats du test de Student fonctionnel appliqué aux hydrogrammes observés et simulés sur les trois bassins avec les deux méthodes de calibration. La statistique de Student est très faible durant la période de crue printanière et durant l'été (jours entre 100 et 200) pour OptFDA et OptNSE sur les trois bassins, ce qui signifie que le modèle calibré avec OptFDA ou OptNSE est capable de reproduire cette composante de l'hydrogramme. En revanche, les débits simulés avec OptNSE en période étiage hivernale (jour entre 1 et 90) sont significativement différents des débits observés sur le bassin du Lac St-Jean et la rivière Nechako. Sur ce dernier, on remarque une amélioration significative avec OptFDA et malgré l'amélioration apportée sur le bassin Lac St-Jean le test rejette l'hypothèse que les débits simulés et observés sont similaires. Sur le bassin du ruisseau Catamaran, OptFDA améliore l'estimation des débits en novembre/décembre.

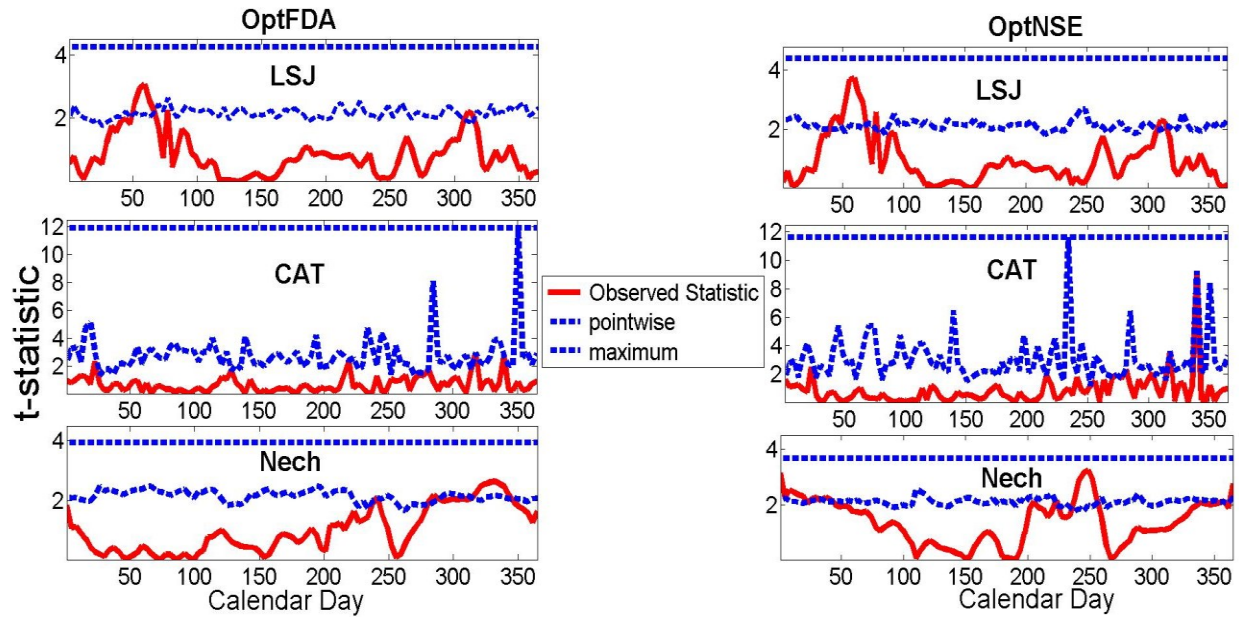


Figure 2-13 Test de Student fonctionnel utilisé pour comparer les hydrogrammes observés et simulés.

Selon les valeurs du NSE obtenues, les deux structures de calibration du modèle (OptNSE et OptFDA) ont la même performance. En effet, le NSE au lac St-Jean est autour de 0.9 en période de calibration et validation avec les deux méthodes, et au bassin de la rivière du Nechako est supérieur à 0.8 en période de calibration et autour de 0.66 en période de validation. Sur le ruisseau Catamaran OptFDA est légèrement moins performant selon le critère NSE (0.73 avec OptFDA contre 0.8 avec OptNSE). Toutefois, comme illustré et cité auparavant, le test de Student accepte l'hypothèse que les hydrogrammes simulés et observés sont similaires. En plus, les biais calculés pour les signatures hydrologiques avec OptFDA sont faibles par rapport à ceux obtenus avec OptNSE. Sur ce bassin, la calibration avec la FDA améliore la simulation de la variabilité des débits (biais du coefficient de variation est de -0.7% pour OptFDA contre -20% pour OptNSE, biais de la variabilité des débits élevés est de -8% pour OptFDA contre -11% pour OptNSE) et de la fréquence des débits élevés excédant 3 fois le débit médian (biais de -8% en période de calibration et validation avec OptFDA contre -12% en période de calibration et -29% en période de validation). Sur le bassin du lac St-Jean, la calibration avec FDA améliore la simulation de signatures décrivant la distribution des débits. Par exemple, le biais du débit moyen et médian est de 7 et 4 m³/s resp. avec OptFDA contre 17 et 29 m³/s avec OptNSE en

période de calibration. Sur le bassin de la rivière du Nechako, les simulations de signatures hydrologiques avec OptNSE et OptFDA sont équivalentes.

L'analyse des courbes de débits classés simulés et observés (Figure 2-14 à 2-16) montre qu'aucune méthode n'est capable de simuler entièrement cette courbe avec la même précision sur chaque portion (débits faibles-moyens-élevés). La calibration basée sur la FDA simule adéquatement : les débits avec probabilité de dépassement inférieur à 2% au bassin du lac St-Jean, le débit maximum au ruisseau Catamaran, les débits avec probabilité de dépassement entre 20 % et 80 % au bassin du lac St-Jean et ruisseau Catamaran. Quant à la calibration en utilisant le NSE, elle simule adéquatement : le débit maximum au bassin du lac St-Jean et la rivière du Nechako, débits avec probabilité de dépassement entre 0.5% et 2% au ruisseau Catamaran, débits avec probabilité de dépassement entre 20% et 40% à la rivière du Nechako. Les deux configurations du modèle avec OptNSE et OptFDA sont incapables de reproduire adéquatement les débits faibles avec probabilité de dépassement supérieur à 80%.

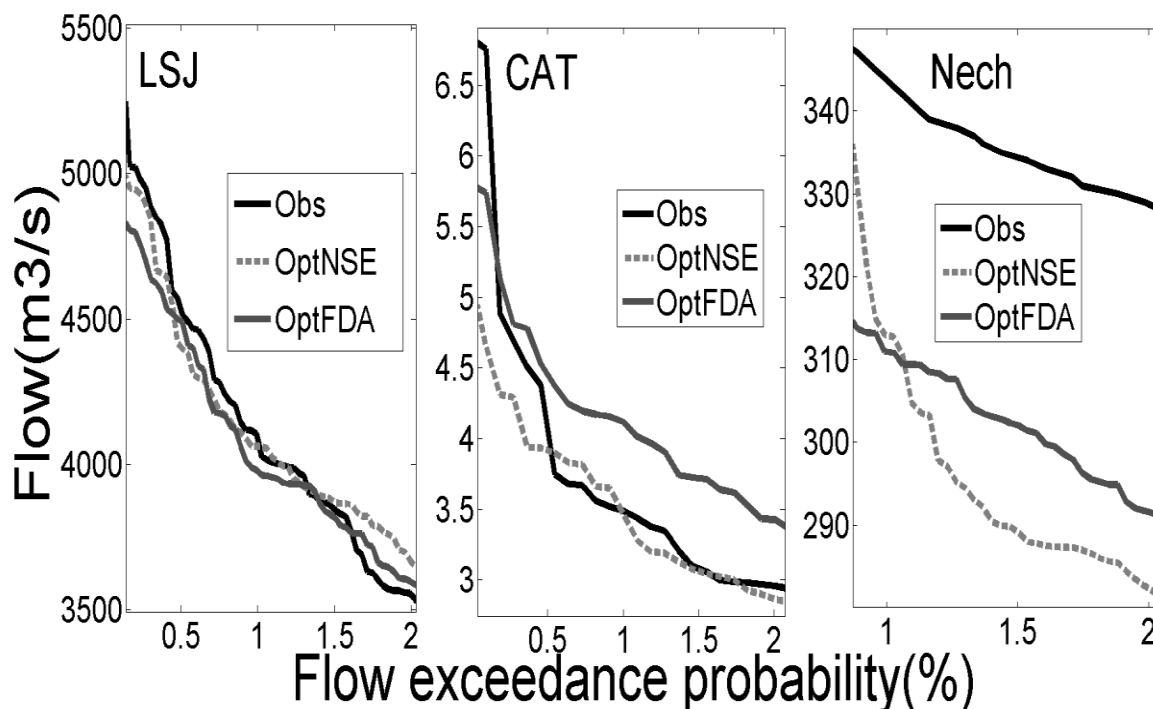


Figure 2-14 Analyse des débits maximums extrêmes de courbes de débits classés observés et simulés avec le modèle paramétré avec OptFDA (modèle calibré avec FDA) et OptNSE (modèle calibré avec NSE) aux trois bassins : Lac St-Jean (LSJ), ruisseau Catamaran (CAT) et à la rivière Nechako (Nech).

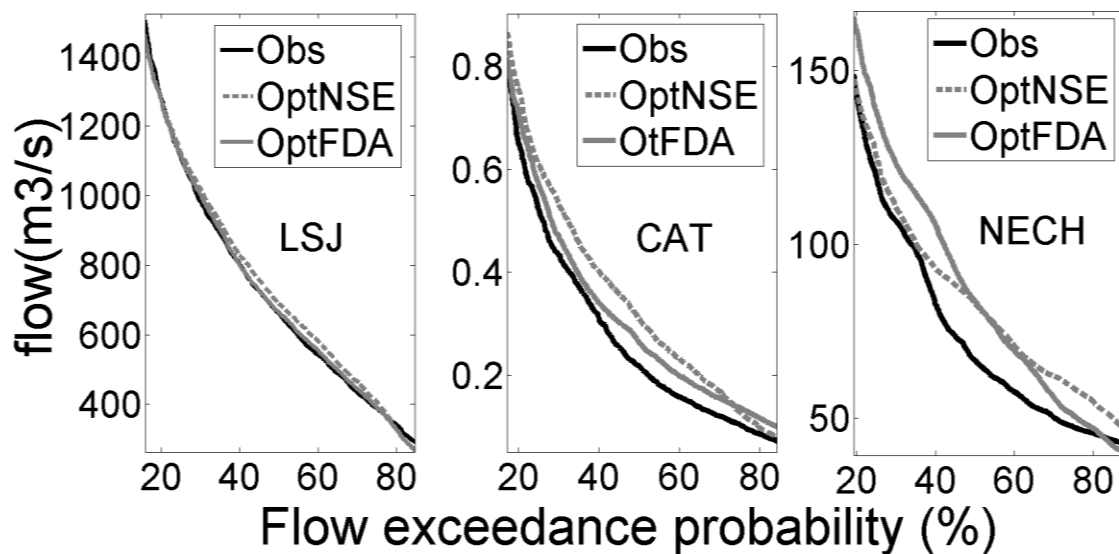


Figure 2-15 Analyse des débits moyens de la courbe de débits classés observés et simulés avec OptFDA (modèle calibré avec FDA) et OptNSE (modèle calibré avec NSE) aux trois bassins : lac St-Jean (LSJ) , Ruisseau Catamaran (CAT) et la rivière Nechako (Nech).

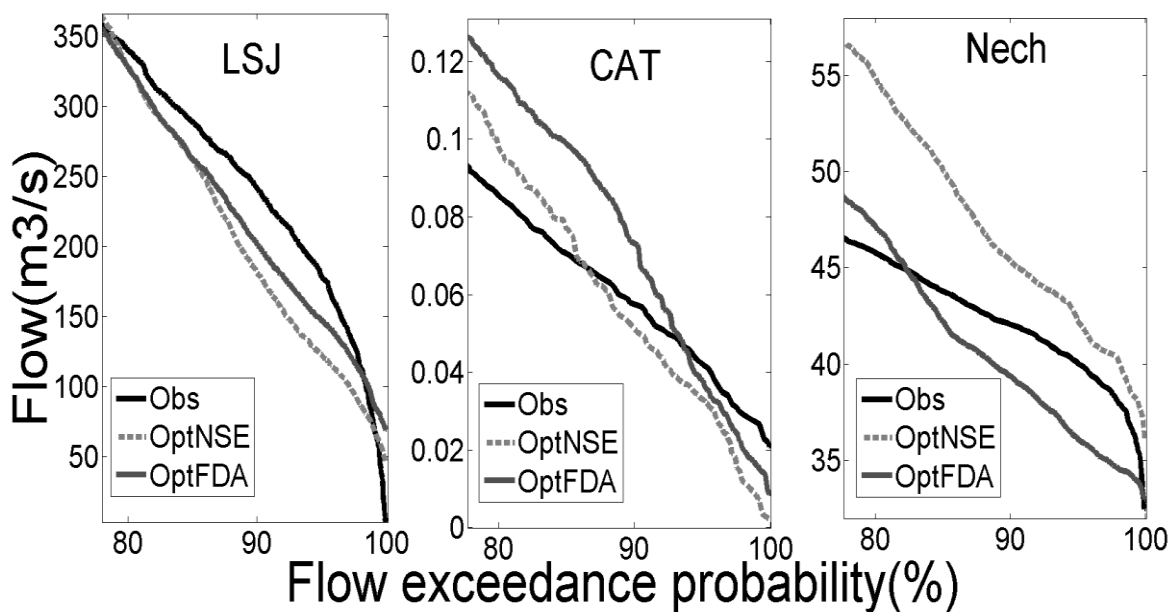


Figure 2-16 Analyse des débits faibles de la courbe de débits classés observés et simulés avec OptFDA (modèle calibré avec FDA) et OptNSE (modèle calibré avec NSE) aux trois bassins : lac St-Jean (LSJ) , Ruisseau Catamaran (CAT) et la rivière Nechako (Nech).

Le tableau 2-8 présente le biais saisonnier des deux configurations OptNSE et OptFDA sur les trois bassins. Au bassin Lac St-Jean les biais saisonniers sont presque similaires pour les deux configurations. On note une amélioration avec OptFDA de 2% pour les deux périodes de janvier à mars et octobre à décembre et une dégradation de 3% des étiages estivaux par rapport à OptNSE. Au bassin du ruisseau Catamaran, on note une réduction significative avec OptFDA du biais calculé pour les deux périodes d'été et d'automne. Au bassin de la rivière du Nechako, le biais des étiages hivernaux et estivaux est significativement amélioré avec OptFDA alors que le biais calculé entre octobre-décembre se dégrade de 20% comparativement à OptNSE. Le biais saisonnier calculé en période de crue est similaire sur les trois bassins avec les deux configurations.

Table 2-8 Biais saisonnier relatif absolu des deux configurations OptFDA et OptNSE.

	Méthode/Bassin	LSJ	CAT	NECH
JFM	OptNSE	0.14	0.01	0.31
	OptFDA	0.12	0.05	0.03
AMJ	OptNSE	0.02	0.04	0.02
	OptFDA	0.02	0.03	0.02
JOS	OptNSE	0.07	0.51	0.14
	OptFDA	0.11	0.04	0.04
OND	OptNSE	0.16	0.40	0.30
	OptFDA	0.14	0.20	0.51

Selon les résultats obtenus, les deux méthodes démontrent des lacunes en ce qui a trait à leur capacité à simuler entièrement et de manière réaliste le mécanisme hydrologique du bassin. En effet, les deux calibrations mènent à des jeux de paramètres différents, ce qui reflète des stratégies différentes selon la fonction objectif. On note des différences significatives des valeurs fixées pour les paramètres de la fonction de production du modèle CEQUEAU qui contrôle les seuils d'infiltration entre les deux réservoirs Sol et Nappe et un seuil de vidange du réservoir Sol (HINF et HINT), les hauteurs des réservoirs (HSOL et HNAP) et le paramètre EVNAP (fraction d'évaporation prise de réservoir Nappe).

Table 2-9 Paramètres ajustés avec la calibration basée sur la FDA (OptFDA) et avec le NSE (OptNSE) au bassin du Lac St-Jean (LSJ), au ruisseau Catamaran (CAT) et à la rivière du Nechako (Nech).

Parameter	LSJ		CAT		Nech	
	OptNSE	OptFDA	OptNSE	OptFDA	OptNSE	OptFDA
CIN	0.53	0.53	0.58	0.17	0.19	0.40
CVMAR	0.18	0.40	0.59	0.79	0.65	0.67
CVNB	0.05	0.08	0.07	0.05	0.19	0.31
CVNH	0.42	0.53	0.32	0.53	0.36	0.47
CVSB	0.02	0.02	0.03	0.02	0.01	0.03
CVSI	0.22	0.39	0.58	0.15	0.26	0.32
XINFMA (mm/day)	23.63	19.54	17.88	30.73	19.45	20.12
HINF (mm)	71.18	81.62	58.30	67.26	83.03	81.67
HINT (mm)	84.33	82.44	55.76	67.90	82.58	80.09
HMAR (mm)	341.48	355.99	358.26	453.80	403.85	334.71
HNAP (mm)	144.36	116.61	116.80	80.95	65.60	87.71
HPOT (mm)	41.49	41.59	15.83	17.24	13.60	2.56
HSOL (mm)	209.58	185.44	196.47	190.04	241.86	226.12
HRIMP	6.19	3.94	5.55	7.10	4.62	4.73
EVNAP	0.43	0.51	0.65	0.19	0.67	0.56
STRNE (°C)	-1.45	-1.55	0.04	2.02	1.99	1.93
TFC (mm/°C/day)	2.01	3.56	2.67	5.63	0.38	0.33
TFD (mm/°C/day)	8.10	8.00	6.67	7.74	4.16	3.95
TSC (°C)	0.12	1.13	1.23	2.44	0.08	1.16
TSD (°C)	-1.34	-1.11	-0.12	-0.90	0.37	1.07
TTD	1.65	0.08	0.25	0.73	0.85	0.62
TTS (°C)	-3.40	-3.11	-1.73	-3.36	-3.05	-1.36

Au ruisseau Catamaran, OptNSE fixe des seuils d'infiltration et percolation (58 mm et 55 mm resp.) plus bas qu'OptFDA (67 mm et 69.9 mm resp.) et une hauteur du réservoir Nappe supérieur à celle d'OptFDA de 36 mm. Autrement dit, OptNSE infiltre plus d'eau vers le réservoir Nappe qu'OptFDA. Toutefois, 65% de l'eau évaporée provient du réservoir Nappe avec OptNSE alors que seulement 19% provient de ce même réservoir avec OptFDA. La réduction d'eau stockée dans le réservoir Nappe est causée principalement par la forte contribution à l'évapotranspiration alors qu'avec OptFDA elle est due aussi à la contribution au débit de base avec le Coefficient de vidange haut du réservoir Nappe CVNH égale à 0.53 contre une valeur de 0.32 pour OptNSE. Conséquemment, le mécanisme simulé par la configuration du modèle avec OptFDA semble plus réaliste. La faible performance du modèle peut s'expliquer par différents facteurs. Au bassin de la rivière Nechako, la faible performance peut être due à une structure de modèle inadaptée pour ce bassin et/ou un biais important sur les intrants (précipitations et températures) puisque la qualité de toutes les calibrations en ce bassin est faible. En effet, les altitudes sur ce bassin varient entre 853m et 1863m ce qui entraîne un biais important sur la météo en haute altitude qui n'est pas disponible, mais interpolée à partir de stations en basse altitude. Le modèle est moins performant sur les débits d'étiages, en particulier les débits hivernaux dans les trois bassins. Cela se traduit par une définition peu réaliste de la composante du modèle qui contrôle l'infiltration vers le réservoir Nappe. La nature conceptuelle du modèle et la non-disponibilité d'information sur les variables d'états limitent les contraintes qu'on peut imposer à ces paramètres.

L'avantage de la FDA est de conserver la variabilité temporelle du débit et ainsi calibrer le modèle en tenant compte de toutes les composantes de l'hydrogramme en une seule étape. Les stratégies de calibration, qu'elles soient mono-objective ou multiobjective, basées sur les mesures de performance traditionnelles (NSE, RMSE, R^2) ou signatures hydrologiques (ex. FDC) nécessitent une sélection préalable de caractéristiques spécifiques du régime hydrologique à considérer lors de la calibration (Yapo et al, 1998 ; Gupta et al. 2008; Yilmaz et al. 2008; Hingray et al. ; 2010; Euser et al, 2013; Shafii et Tolson, 2015). En plus, l'information sur la variabilité temporelle de l'écoulement est perdue puisque ces mesures de performance sont données par une seule valeur. Les statistiques fonctionnelles sont formulées pour inclure tous les pas de temps, ce qui les rend plus adaptées pour analyser les hydrogrammes, comme illustré avec le test de Student fonctionnel.

6.2.3 Conclusion

Cette partie de la thèse propose une fonction objectif basée sur la FDA qui permet de conserver la forme, la variabilité de l'hydrogramme ainsi qu'une analyse en fonction du temps, ce qui est mieux adapté que les approches classiques pour évaluer la performance d'un modèle hydrologique. Cette méthode n'est pas suffisante pour assurer une représentation adéquate de l'ensemble des processus hydrologiques conceptualisés dans le modèle à cause des différentes sources d'incertitudes (observations, structure du modèle), compensation de paramètres et le manque d'information sur les processus hydrologiques pour contraindre les paramètres. L'approche proposée dans cette partie est la première application de la FDA dans un contexte de calibration de modèle hydrologique. Une différente formulation de la fonction objectif basée sur la FDA peut fournir différents résultats. Ceci est l'objectif de la partie suivante qui vise à explorer une fonction objectif basée sur la FDA différente et comparer l'efficacité de celle-ci avec les différentes approches de calibration.

6.3. Objectif 3 : Comparaison d'approches de calibration avec analyse d'incertitude

Cette partie compare quatre approches de calibration à savoir: 1) calibration traditionnelle avec le NSE, 2) calibration multiobjective, 3) calibration basée sur le principe du diagnostic (i.e. basée sur les signatures hydrologiques), et 4) calibration basée sur l'analyse de données fonctionnelles proposée dans la partie précédente. Dans cette partie, on explore une formulation de fonction objectif basée sur la FDA différente de celle définie dans la partie précédente. La comparaison entre les approches porte sur la comparaison de la performance de l'ensemble des jeux de paramètres optimaux résultants de l'application de chaque approche avec des Simulations Monte Carlo couplée avec la notion de fonction de profondeur telle que développée dans la première partie. On analyse aussi des bandes d'incertitude liée à l'estimation des paramètres par chaque approche. Le bassin versant Chute-Du-Diable (sous-bassin du Lac St-Jean) est utilisé comme cas d'étude.

6.3.1 Méthodologie

a) Calibration basée sur la FDA

La première étape de la calibration avec la FDA est la même que celle décrite dans la sous-section 6.2.1. Cette étape consiste à convertir les données de débits journaliers en courbes continues représentant les hydrogrammes lissés. Comme le principe de la calibration avec la FDA consiste à comparer les courbes observées et simulées par le modèle, le critère évaluant la performance du modèle défini compare les surfaces sous chaque courbe simulée et observée. Cette surface représente dans notre contexte un volume d'eau annuel intégré. Pour chaque année i , on calcule alors le critère :

$$of_i = \left| 1 - \frac{\int_{t=1}^{T=365} Q_{i,sim}(t)}{\int_{t=1}^{T=365} Q_{i,obs}(t)} \right| \quad (2.22)$$

avec : $Q_{i,sim}(t)$ et $Q_{i,obs}(t)$ sont l'hydrogramme lissé simulé et observé pour l'année i avec l'équation 2.19.

Le critère global de calibration est la somme de of_i pour la période de calibration pondérée par la statistique du test de Student comme pour l'équation 2.20 :

$$FDA_OF = S \times \sum_{i=1}^N of_i \quad (2.23)$$

avec: of_i comme défini dans l'équation 2.22, S est la moyenne de la statistique de Student et N est le nombre d'années d'observations.

b) Calibration avec une seule fonction objectif

Pour cette approche de calibration, le NSE (décrit par l'équation 2.16) est utilisé comme fonction objectif. Ce critère est le critère le plus utilisé en hydrologie. En plus, il est reconnu pour sa performance à capter l'occurrence et l'amplitude de la crue, l'événement dominant dans le bassin versant étudié.

c) Calibration multiobjective

Pour l'approche multiobjective, une agrégation de trois critères : corrélation, mesure de biais et mesure de variance est utilisée. Cette agrégation est le critère KGE (décrit par l'équation 2.18). Ce critère est une amélioration du NSE pour capter la variabilité du débit, la pointe de crue avec une forte corrélation entre les débits simulés et observés (Gupta et al., 2009).

d) Calibration basée sur le principe du diagnostic ou signatures hydrologiques

Le principe du diagnostic consiste à évaluer la performance du modèle hydrologique à reproduire un nombre de signatures hydrologiques décrivant le processus dominant. La calibration basée sur ce principe consiste à une calibration multiobjective du modèle selon les signatures hydrologiques préalablement sélectionnées. Dans cette étude, on opte pour la stratégie développée par Shafii et Tolson (2015). Cette stratégie consiste à maximiser le nombre de signatures satisfaites selon un seuil d'acceptation fixé. En effet, Shafii et Tolson (2015) ont démontré que considérer explicitement le nombre de signatures hydrologiques satisfaites permet d'explorer efficacement l'espace des paramètres. Douze signatures hydrologiques sont utilisées dans cette étude. Ces signatures décrivent la distribution du débit (ex. débit moyen, débit médian et variance), la variabilité du débit (ex. coefficient de variation), les débits de crue et débits d'étiages selon des indices extraits de la courbe de débits classés et des indices de débits annuels. La fonction objectif est une métrique de consistance définie comme suit :

$$SIG_OF = \sum_{i=1}^{n^*+1} C_i, \quad (2.24)$$

avec: n^* : nombre de signatures satisfaites et C_i est une fonction de score continue tel que $C_i = 1$ si le biais relatif entre la signature observée et simulée est inférieur au seuil d'acceptation, $C_i = 0$ si le biais entre la signature observée et simulée est supérieur à un seuil maximum (ici 100%), sinon $C_i = \frac{100 - |D_i|}{100 - |D_i^*|}$ avec D_i est le biais relatif entre la signature observée et simulée et D_i^* et le seuil d'acceptation. Ce critère, à maximiser, atteint sa valeur idéale à 12 puisque douze signatures hydrologiques sont considérées ; ce qui signifie que toutes les signatures sont satisfaites selon le seuil d'acceptation considéré. La valeur minimale est de 0 avec aucune signature satisfaite. Deux seuils d'acceptation sont fixés pour la calibration avec cette stratégie : le premier à 10% et le deuxième plus restrictif à 5%.

6.3.2 Résultats et discussion

On applique les quatre approches pour définir l'espace optimal des paramètres de CEQUEAU sur le bassin Chute-Du-Diable. Le modèle est calibré sur la période comprise entre 1991 et 1999 et validé sur la période entre 2000 et 2013. On note l'ensemble des jeux de paramètres de la calibration avec NSE OptNSE, avec le KGE OptKGE, avec les signatures hydrologiques avec le seuil de 10 % OptSIG10 avec le seuil 5% OptSIG5 et avec le critère basé sur la FDA OptFDA.

La procédure d'identification des paramètres optimaux converge rapidement avec le critère KGE et le critère basé sur la FDA comparativement à l'approche utilisant le NSE ou les signatures hydrologiques (voir Figure 2-17). L'étape de transformation des débits simulés en fonctions temporelles, requise par l'approche FDA, impacte légèrement le temps de calcul qui reste similaire au temps de calcul requis par la calibration avec le NSE. Quant à l'approche basée sur les signatures hydrologiques, le temps de calcul requis est important à cause du nombre d'itérations (plus de 40 itérations). Les gammes de valeurs des jeux de paramètres du critère basé sur la FDA et du NSE OptNSE et OptFDA sont très étroites comparativement aux autres approches.

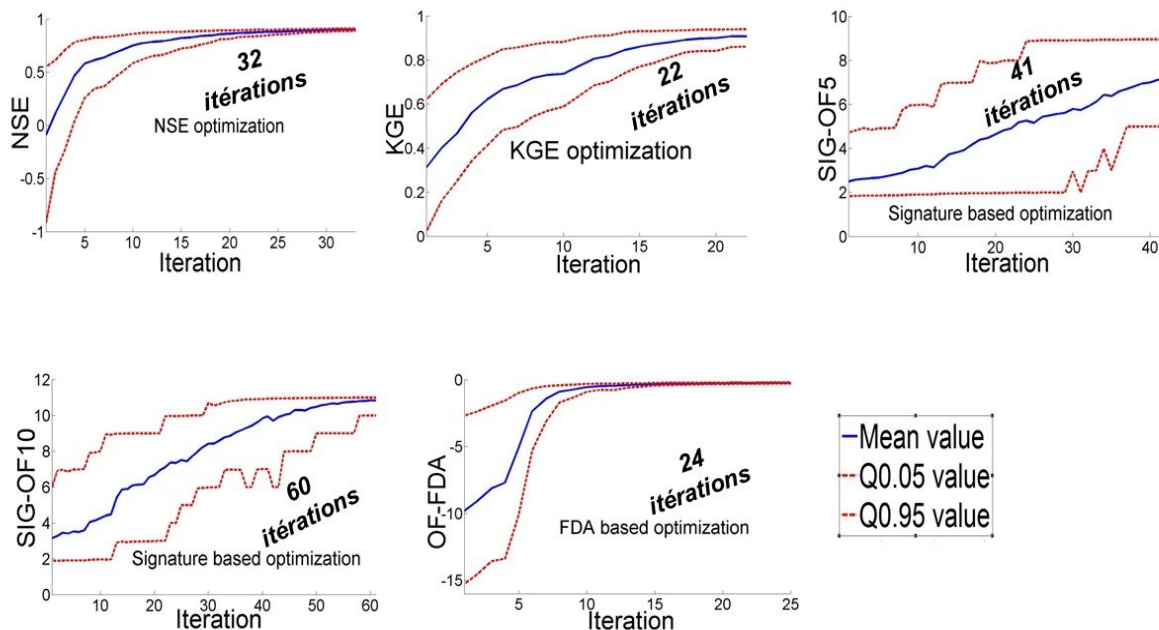


Figure 2-17 Convergence de l'algorithme selon l'approche de calibration et analyse de la distribution de la fonction objectif considérée.

Le tableau 2-10 présente la distribution des jeux de paramètres optimaux obtenus avec les approches de calibration considérées. Les calibrations avec le NSE et KGE convergent vers des valeurs de paramètres similaires pour presque tous les paramètres avec une dispersion plus importante avec le KGE. Globalement, toutes les approches mènent à des valeurs similaires pour certains paramètres régissant l'écoulement dans le réservoir Sol (valeurs en bleu dans le tableau 2-10). Ces paramètres sont les seuils d'infiltration maximale par jour et du réservoir Sol vers le réservoir Nappe (HINT et XINFMA), de vidange (HINF), hauteur du réservoir Sol (HSOL) et

coefficient de vidange basse du réservoir Sol (CVSB). On note aussi des valeurs similaires du pourcentage d'évapotranspiration pris du réservoir Nappe (EVNAP) et le coefficient de déficit calorique (TTD). Des différences significatives sont remarquées pour les paramètres de fonte et la hauteur du réservoir Nappe (HNAP).

Table 2-10 Distribution des jeux de paramètres optimaux obtenus avec les approches de calibration avec NSE (OptNSE) avec KGE (OptKGE), basée sur les signatures hydrologiques au seuil d'acceptation de 10% (OptSIG10) et de 5% (OptSIG5) et l'approche proposée basée sur la FDA (OptFDA).

Paramètre	OptNSE		OptKGE		OptSIG10		OptSIG5		OptFDA	
	Ecart		Ecart		Ecart		Ecart		Ecart	
	Moyenne	type	Moyenne	type	Moyenne	type	Moyenne	type	Moyenne	type
STRNE (°C)	-0.50	0.24	-0.65	0.42	-0.76	0.08	0.29	0.52	-0.89	0.31
TFC (mm/°C/day)	3.27	0.30	3.31	0.62	1.54	0.09	1.54	0.22	2.56	0.53
TFD (mm/°C/day)	6.88	0.92	7.00	1.08	6.86	0.21	6.04	0.96	6.21	0.97
TSC (°C)	1.48	0.31	1.44	0.44	1.58	0.10	1.37	0.36	1.30	0.58
TSD (°C)	-0.60	0.43	-0.55	0.52	-0.89	0.08	-0.48	0.35	-0.55	0.52
TTD	0.38	0.11	0.40	0.14	0.33	0.02	0.46	0.08	0.36	0.10
TTS (°C)	-2.08	0.59	-2.37	0.87	-2.45	0.28	-2.16	0.73	-1.71	0.99
CIN	0.41	0.10	0.51	0.17	0.59	0.04	0.48	0.10	0.51	0.14
CVNH	0.16	0.02	0.40	0.17	0.52	0.05	0.51	0.12	0.47	0.17
CVSB	0.03	0.00	0.04	0.01	0.04	0.00	0.05	0.01	0.01	0.00
CVSI	0.40	0.11	0.37	0.15	0.38	0.07	0.42	0.08	0.47	0.16
XINFMA (mm/day)	20.86	4.04	20.09	6.32	19.65	2.27	16.52	3.13	21.74	6.23
HINF (mm)	80.03	3.28	71.63	7.22	58.19	1.85	63.45	5.59	82.40	4.40
HINT (mm)	79.69	4.41	80.04	7.74	61.29	2.55	61.26	6.77	79.44	3.39
HNAP (mm)	105.19	16.75	123.75	27.11	104.89	7.40	105.79	19.44	80.85	18.21
HPOT (mm)	23.90	6.12	21.63	7.87	18.28	0.60	15.00	3.42	42.32	4.74
HSOL (mm)	203.24	24.08	198.31	30.64	199.01	10.56	202.58	21.19	207.07	31.66
EVNAP	0.17	0.02	0.16	0.03	0.16	0.01	0.14	0.02	0.14	0.03

La performance et la consistance des jeux de paramètres obtenus avec les quatre approches sont évaluées avec les critères NSE, KGE et le nombre de signatures hydrologiques satisfaites au seuil d'acceptation de 10% et de 5%.

Toutes les approches sont sujettes à un risque d'équifinalité. Toutefois, l'équifinalité avec le critère NSE et KGE est plus prononcée, comparativement aux autres approches (voir Figure 2-18). La majorité des jeux de paramètres de OptNSE (environ 95 %) ont une valeur de NSE entre 0.88 et 0.90 en période de calibration et entre 0.86 et 0.9 en validation.

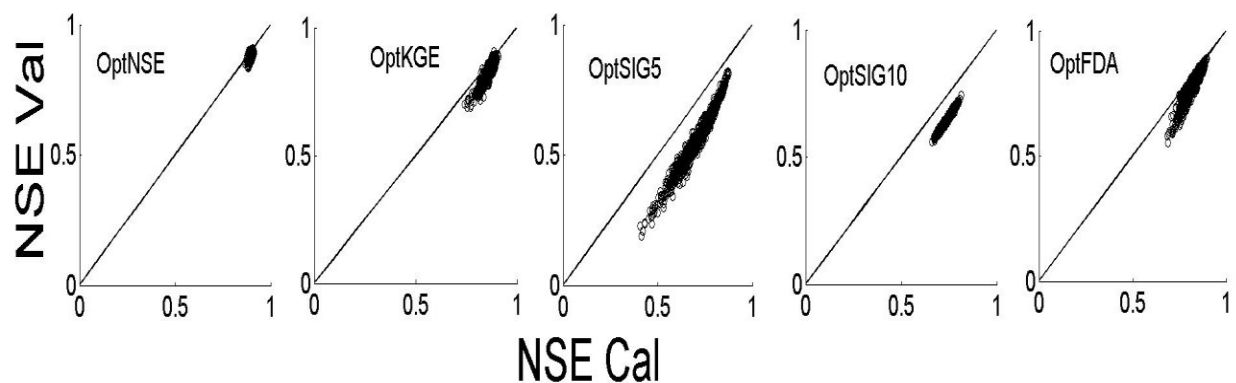


Figure 2-18 Valeurs de NSE en calibration vs en validation pour les jeux de paramètres optimaux obtenus avec la calibration avec NSE (OptNSE), avec KGE (OptKGE), basée sur les signatures hydrologiques au seuil d'acceptation de 10% (OptSIG10) et de 5% (OptSIG5) et l'approche proposée basée sur la FDA (OptFDA).

Environ 95% des jeux de paramètres de OptKGE ont une valeur du KGE entre 0.86 et 0.94 en calibration. La performance de cet espace se dégrade en validation avec des valeurs entre 0.76 et 0.91. En augmentant le seuil d'acceptation, l'approche basée sur les signatures hydrologiques devient plus pénalisante et distinctive en termes de valeurs de similarité dans les valeurs de paramètres. Environ 88% des paramètres de l'espace défini avec les signatures hydrologiques en utilisant un seuil d'acceptation de 10% satisfont 10 sur 12 signatures alors que seulement 74% de paramètres de l'espace défini en utilisant le seuil de 5 % satisfont de 5 à 7 des 12 signatures.

Les Figures 2-19 et 2-20 présentent le nombre de signatures hydrologiques satisfaites au seuil d'acceptation de 10% et 5% en calibration et en validation. L'ensemble OptNSE est très consistant comparativement à OptKGE. Environ 73 % de paramètres de cet espace satisfont plus de 10 sur 12 signatures avec un seuil d'acceptation de 10 % alors que seulement 37 % des jeux de paramètres de OptKGE satisfont le même nombre de signatures. Ces jeux de paramètres ont une faible consistance au seuil de 5% avec seulement 10 % de OptNSE et 2% OptKGE sont capables de reproduire 10 à 11 signatures hydrologiques.

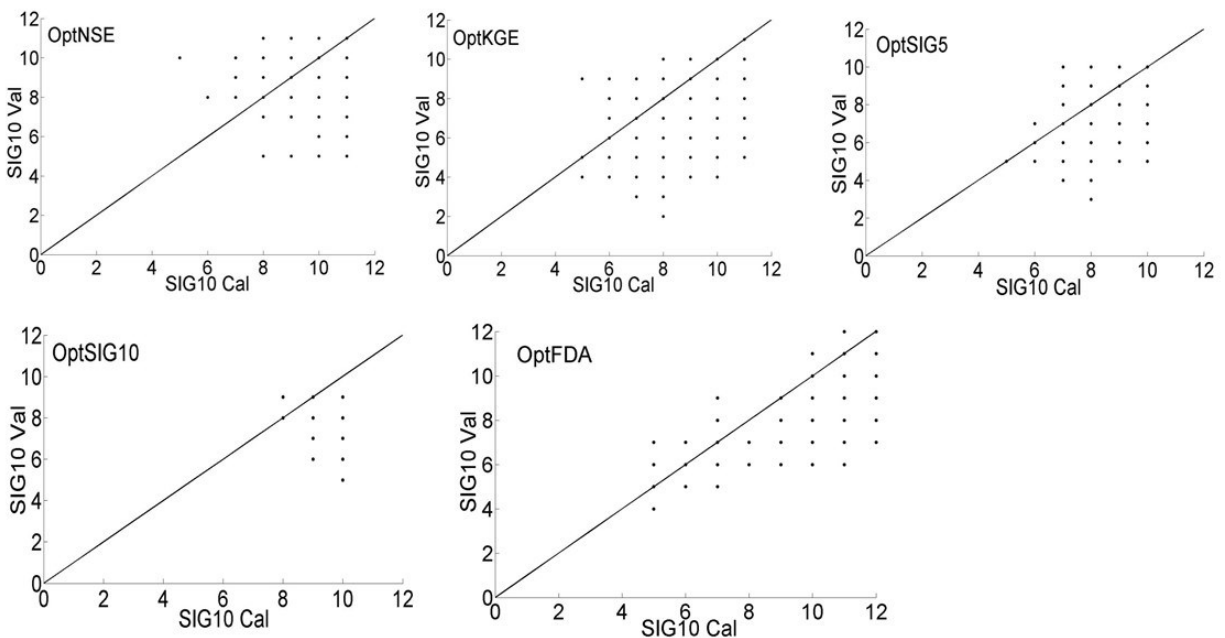


Figure 2-19 Nombre de signatures satisfaites au seuil d'acceptation 10% en calibration vs en validation avec les jeux de paramètres optimaux obtenus avec la calibration avec NSE (OptNSE), avec KGE (OptKGE), basée sur les signatures hydrologiques au seuil d'acceptation de 10% (OptSIG10) et de 5% (OptSIG5) et l'approche proposée basée sur la FDA (OptFDA).

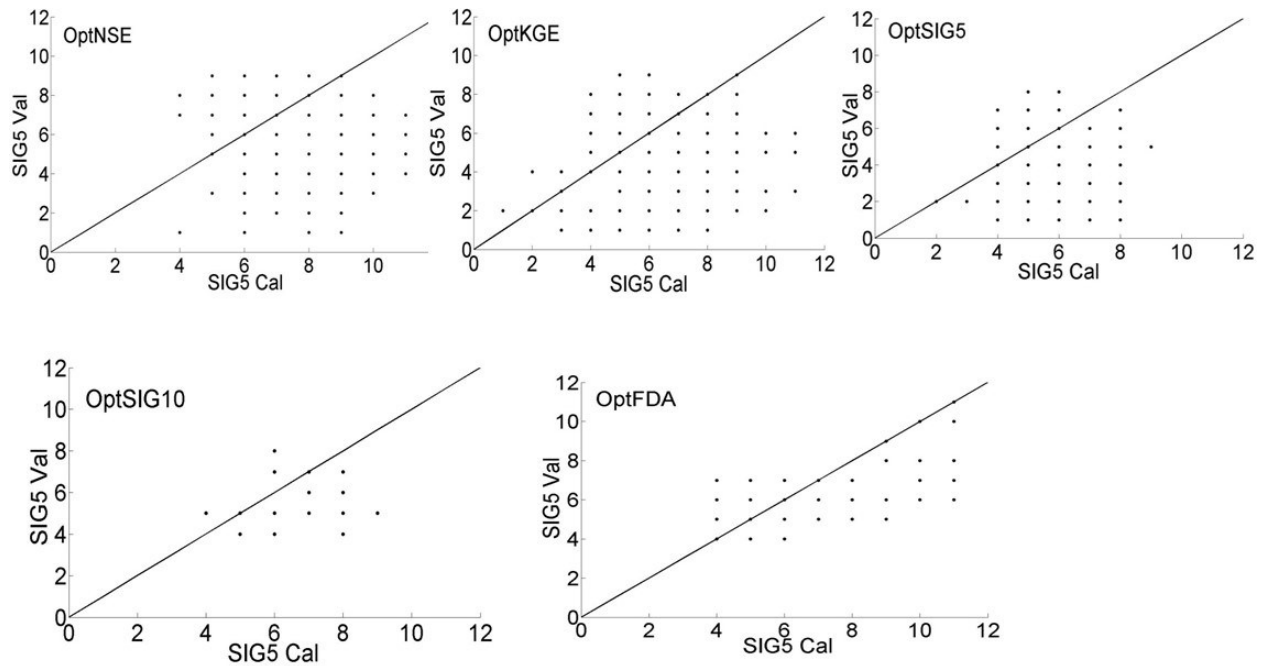


Figure 2-20 Nombre de signatures hydrologiques satisfaites au seuil d'acceptation 5% en calibration vs en validation avec les jeux de paramètres optimaux obtenus avec la calibration avec NSE (OptNSE), avec KGE (OptKGE), basée sur les signatures hydrologiques au seuil d'acceptation de 10% (OptSIG10) et de 5% (OptSIG5) et l'approche proposée basée sur la FDA (OptFDA).

La performance des espaces définis par la calibration avec les signatures hydrologiques aux deux seuils d'acceptation est acceptable avec des valeurs de NSE variant entre 0.56 et 0.84 et KGE entre 0.76 et 0.89 en période de calibration. En période de validation, ces valeurs varient entre 0.36 et 0.74 pour le NSE et entre 0.66 et 0.83 pour le KGE. L'ensemble OptFDA est performant avec 95 % de paramètres ayant des valeurs de NSE entre 0.75 et 0.88 et de KGE entre 0.77 et 0.92 en période de calibration. En validation les valeurs du NSE varient entre 0.67 et 0.86 et le KGE entre 0.71 et 0.87. Cet ensemble de paramètres est moins consistant que les autres ensembles avec seulement 17% des paramètres satisfaisant plus de 10 signatures. En revanche, presque 3% des jeux de paramètres satisfont toutes les signatures hydrologiques en période de calibration et 1.7% en période de validation au seuil d'acceptation de 10%. Les autres approches n'ont pas atteint ce degré de consistance (i.e. aucune autre des approches n'a défini un jeu de paramètres qui satisfait toutes les 12 signatures considérées). La consistance peut être augmentée en augmentant la taille de l'échantillon qui implique une augmentation du temps de calcul. Dans

cette étude la taille de l'échantillon est fixée à 1000 jeux de paramètres pour des temps de calcul raisonnables à cause de la complexité du modèle.

Pour l'analyse d'incertitude, on compare les bandes d'incertitudes (valeurs minimales et maximales pour chaque pas de temps) de l'hydrogramme moyen résultant de chaque approche (voir Figure 2-21). Les approches conventionnelles sont incapables de reproduire les débits d'étiages en hiver avec des bandes en dessous de la valeur observée en période de calibration et validation. Les bandes d'incertitude de l'approche proposée basée sur la FDA sont très étroites autour de la valeur observée en hiver, que ce soit en période de calibration ou validation. Les bandes sont larges en période de crue printanière, quelle que soit l'approche de calage. Une faible performance à capter l'occurrence de la crue et sa pointe est notée pour l'approche basée sur les signatures hydrologiques. Pour cette approche, le coefficient de corrélation entre simulations et observations n'est pas utilisé, ce qui peut expliquer cette faible performance. Toutes les approches surestiment les débits en période d'automne avec seule la valeur minimale qui représente adéquatement ces débits. Avec l'approche FDA et celle basée sur les signatures au seuil d'acceptation de 10%, le biais entre la valeur de débit observé et la borne maximale de l'ensemble des débits simulés est moins élevé.

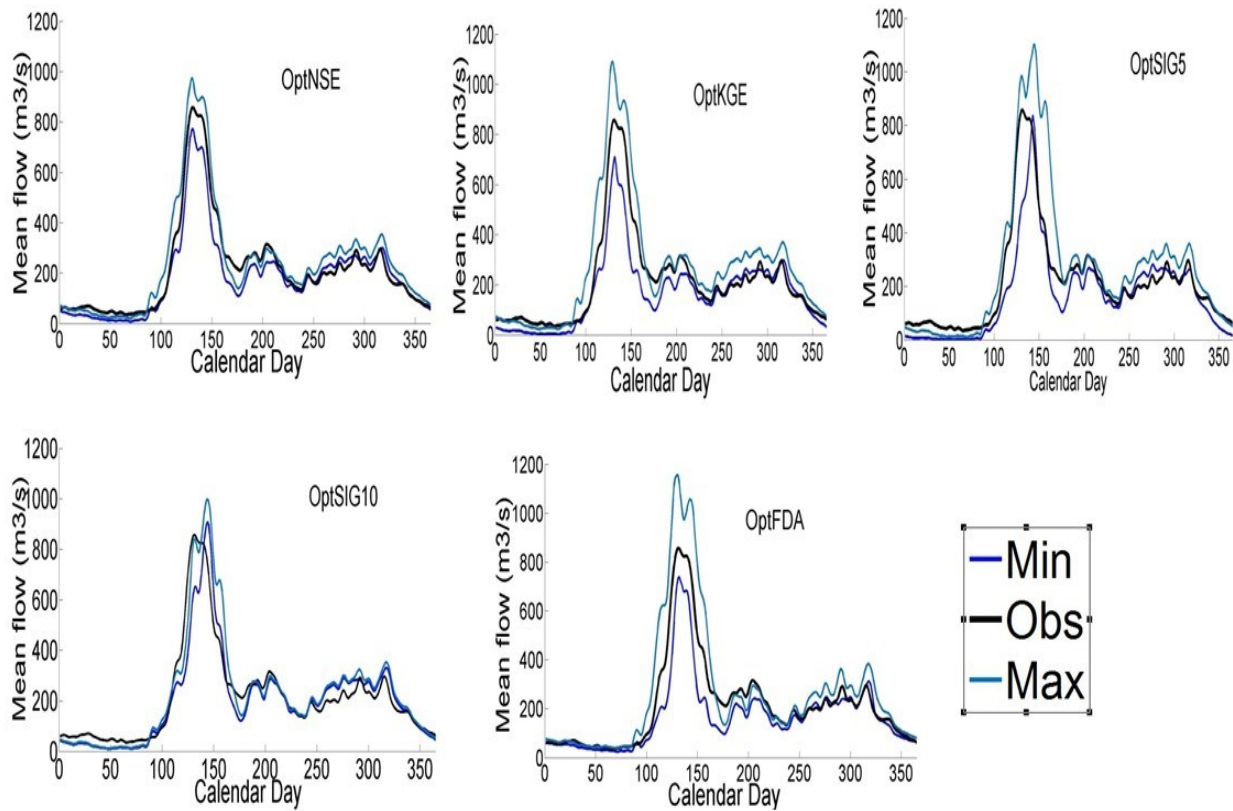


Figure 2-21 Hydrogramme moyen observé avec les hydrogrammes moyens minimaux et maximaux simulés en période de calibration avec le modèle calibré avec NSE (OptNSE), avec KGE (OptKGE), avec les signatures hydrologiques au seuil d'acceptation de 10% (OptSIG10) et de 5% (OptSIG5) et avec l'approche proposée basée sur la FDA (OptFDA).

L'ensemble OptSIG5 obtenu avec l'approche basée sur les signatures hydrologiques au seuil d'acceptation de 5% est le moins performant. Dans cette étude l'algorithme utilisé vise à optimiser une seule fonction objectif. L'application de cette approche avec un algorithme qui vise à explorer le front de Pareto est plus efficace et peut atteindre un degré de consistance très élevé en particulier à des seuils très restrictifs tels qu'à 5 % (Shafii et Tolson, 2015). L'approche basée sur la FDA peut aussi être utilisée avec un tel algorithme en optimisant simultanément le vecteur des critères annuels ($\{of_1, \dots, of_N\}$) avec of_i tel que défini avec l'Eq. 2.22 au lieu d'utiliser une agrégation.

6.3.3 Conclusion

Dans ce chapitre de la thèse, quatre approches de calage sont comparées sur un bassin à forte dominance hydrologique de la crue printanière causée par la fonte de neige. Toutes les approches ont démontré des difficultés pour capter toute la forme de l'hydrogramme. Toutefois, l'approche proposée basée avec la FDA montre une supériorité en termes de meilleure représentation de l'hydrogramme en entier. Cette approche a montré un potentiel à atteindre un degré de consistance très élevé (reproduire toutes les signatures hydrologiques considérées). L'incertitude sur les débits de la crue est plus significative que sur les débits d'étiage ou de la crue automnale. L'approche classique la plus performante pour une meilleure estimation de la crue est le calage avec le critère NSE.

6.4. Objectif 4 : Calibration par processus multicritère pour un modèle hydrologique conceptuel

Dans cette partie, on s'intéresse à une calibration par processus hydrologiques dans le but d'obtenir un modèle qui, dans les limites de sa conceptualisation, reproduit de manière réaliste les processus hydrologiques et les états internes. Les méthodes développées dans la deuxième partie de la thèse sont exploitées et les critères de calibration sont définis basés sur le même principe. Toutes les observations sont transformées en fonctionnelles comme exposé dans la partie 6.2.1.

6.4.1 Méthodologie

Pour minimiser l'effet de la compensation des paramètres, chaque composante du modèle est calibrée séparément selon un critère qui reflète l'adéquation du processus qu'elle régit. Les paramètres contrôlant l'accumulation et la fonte de neige sont ajustés en premier selon la courbe fonctionnelle de crue. Ensuite, on calibre les paramètres de la fonction de production du modèle CEQUEAU, notés paramètres physiques, qui distribuent verticalement les volumes d'eau dans les différentes zones du sol conceptualisées. Ces paramètres sont ajustés sur la base d'une séparation de l'hydrogramme en débit de base et ruissellement de surface. Finalement, les paramètres de la fonction de transfert sont ajustés selon la qualité d'ajustement et d'occurrence des hydrogrammes sur au moins deux stations : une en amont et une en aval du bassin. Le

processus est répété jusqu'à non amélioration de la performance globale du modèle. La performance globale du modèle est évaluée avec le critère défini par l'équation 2.21 évaluant la performance du modèle à reproduire adéquatement les hydrogrammes annuels observés.

a) Ajustement des paramètres de fonte

Les paramètres de fonte régissent le processus d'accumulation du stock de neige en hiver et la fonte de celui-ci au printemps. Ils affectent alors l'occurrence, la pointe et le volume de la crue printanière. Ces paramètres sont ajustés uniquement selon la courbe de la crue printanière dans un cadre d'analyse fonctionnelle. Pour faciliter l'extraction automatique de cette courbe lors de la calibration, les hydrogrammes lissés sont alignés selon l'hydrogramme moyen. L'alignement des courbes consiste en une translation dans le temps selon un intervalle de référence en appliquant une fonction appelée 'warping function' (Ramsay et Silverman, 2002). Cet alignement permettra de fixer l'occurrence de la crue à un intervalle fixé au préalable (entre $t_1=81$ correspondant au 21 mars et $t_2=181$ correspondant au 30 juin) ce qui permettra l'extraction automatique du volume de la crue. Pour extraire l'occurrence réelle de la crue et de la pointe, les dérivées des hydrogrammes observés et simulés non alignés sont comparées. La fonction objectif est définie, basée sur le même principe que celui présenté dans la partie 6.2.1, en utilisant la distance entre les courbes de crue alignées et les dérivées des hydrogrammes non alignés :

$$Snow_OF = S * \sqrt{\sum_{i=1}^N \int_{t=81}^{T=181} |x_{i,sim}(t) - x_{i,obs}(t)|^2 + \int_{t=1}^{T=365} |x'_{i,sim}(t) - x'_{i,obs}(t)|^2} \quad (2.25)$$

Avec : $x_{i,sim}$ et $x_{i,obs}$: hydrogrammes simulé et observé à l'année i ; $x'_{i,sim}$ et $x'_{i,obs}$: dérivée des hydrogrammes simulé et observé à l'année i ; S : la moyenne de la statistique du test de Student fonctionnel.

b) Ajustement des paramètres physiques

Les paramètres physiques de la fonction de production du modèle CEQUEAU (voir figure 2-3) régissent les processus d'infiltration, d'évapotranspiration et de l'écoulement dans les réservoirs conceptualisés par le modèle. Ces paramètres sont ajustés selon l'écoulement de base et le ruissellement de surface. Comme des observations de débits de base et ruissellement sont indisponibles, on procède à une séparation de l'hydrogramme grâce au filtre paramétrique

d'Eckhardt (2005). Ce filtre à deux paramètres généralise les filtres récursifs à un seul paramètre. De plus, cet outil permet la séparation automatique de l'hydrogramme. Le débit de base s'écrit :

$$b_{t+1} = \frac{(1-BFI_{max})\alpha b_t + (1-\alpha)BFI_{max}y_{t+1}}{1-\alpha BFI_{max}} \quad (2.26)$$

Avec : b est le débit de base; y est le débit; BFI_{max} est la proportion du débit de nappe par rapport au débit total annuel; α est la constante de récession; si $b_{t+1} < y_{t+1}$; sinon : $b_{t+1} = y_{t+1}$.

Les paramètres du filtre BFI_{max} et α sont fixés selon les valeurs recommandés par Eckhardt (2005) parce que les observations de débits de base et de ruissellement de surface sont indisponibles. Il est fortement recommandé d'ajuster les paramètres BFI_{max} et α avant de procéder à la calibration du modèle dans le cas où les mesures de ruissellement de surface et débit de base sont disponibles (Lim et al. 2010). Pour évaluer la performance du modèle à simuler adéquatement l'évapotranspiration, le ratio entre débits et précipitations totales est inclus dans la fonction objectif.

Les paramètres physiques sont regroupés en deux catégories où chaque catégorie est ajustée selon un critère différent. La première concerne les paramètres qui régissent l'écoulement dans le réservoir SOL qui représente la zone saturée. Ces paramètres sont ajustés selon le ruissellement de surface. La fonction objectif est comme suit :

$$runoff_OF = \sqrt{\sum_{i=1}^N \int_{t=1}^{T=365} |f_{sim,i}(t) - f_{obs,i}(t)|^2 dt} + \frac{\sum_{t=1}^{T=365*N} |Q_{sim,t} - Q_{obs,t}|}{\sum_{t=1}^{T=365*N} P_t} \quad (2.27)$$

où : $f_{sim,i}$ et $f_{obs,i}$ sont les courbes du ruissellement de surface simulé et observé à l'année i ; N est le nombre d'années d'observations; $Q_{sim,t}$ et $Q_{obs,t}$ sont le débit simulé et observé au jour t ; P_t : précipitations au jour t .

La deuxième catégorie est reliée à l'écoulement dans le réservoir nappe représentant l'écoulement la zone non saturée reliée directement à l'écoulement de base. En conséquence, ces paramètres sont ajustés selon l'adéquation des simulations du débit de base. La fonction objectif est définie comme suit :

$$baseflow_OF = \sqrt{\sum_{i=1}^N \int_{t=1}^{T=365} |b_{sim,i}(t) - b_{obs,i}(t)|^2 dt} + \frac{\sum_{t=1}^{T=365*N} |Q_{sim,t} - Q_{obs,t}|}{\sum_{t=1}^{T=365*N} P_t} \quad (2.28)$$

où : $b_{sim,i}$ et $b_{obs,i}$ sont les courbes de débits de base simulés et observés à l'année i ; N est le nombre d'années d'observations; $Q_{sim,t}$ et $Q_{obs,t}$ sont le débit simulé et observé au jour t ; P_t : précipitation totale au jour t .

c) Ajustement des paramètres de transfert

Les paramètres de fonte et les paramètres physiques sont ajustés selon les observations de débits à l'exutoire. Supposons que ces paramètres soient parfaitement ajustés, toute désynchronisation entre les simulations et les observations dans le bassin versant est due aux paramètres de transfert. Autrement dit, si les paramètres de fonte et physiques sont parfaitement ajustés le modèle est capable de fournir la même qualité de simulations sur n'importe quelle région du bassin versant ; sinon les paramètres de transfert sont incapables d'acheminer adéquatement l'eau dans le bassin. Sous cette hypothèse, les paramètres de transfert sont ajustés en évaluant la qualité de l'ajustement sur deux stations, tout autre paramètre ayant une valeur constante. La fonction objectif est définie comme l'erreur entre la performance du modèle dans la station en amont et celle à l'exutoire :

$$Routing_OF = \sum_{i=1}^N \varepsilon_{i,downstream} - \sum_{i=1}^N \varepsilon_{i,upstream} \quad (2.29)$$

avec : N est le nombre d'années ; $\varepsilon_{i,downstream}$ et $\varepsilon_{i,upstream}$: sont la somme des ratios entre les volumes d'eau simulés et observés en aval et en amont avec $\varepsilon_i = \sum_{i=1}^N \frac{\int_{t=1}^{T=365} x_{i,sim}}{\int_{t=1}^{T=365} x_{i,obs}}$ où $x_{i,sim}$ et $x_{i,obs}$ sont les hydrogrammes lissés simulé et observé à l'année i .

6.4.2 Résultats et discussion

Cette stratégie de calibration est appliquée seulement au bassin du lac St-Jean où l'on dispose de plusieurs stations hydrométriques. Les paramètres de fonte et les paramètres physiques sont ajustés sur la station à l'exutoire et la station ASHA est utilisée pour l'ajustement des paramètres de transfert. On compare les résultats issus de la calibration par processus (notée OptProcess) avec ceux issus d'une calibration globale (notée OptGlob) avec le seul critère défini avec l'équation 2.21. Dans les deux cas, le coefficient de vidange bas du réservoir Nappe (CVNB) est fixé à 0 parce que le modèle tend à vider le réservoir NAPPE en période d'étiage quand ce paramètre est libre. La borne supérieure du paramètre EVNAP (fraction d'évaporation prise du réservoir Nappe) est fixée à 0.25 afin de limiter la compensation du modèle avec

l'évapotranspiration pour réduire le niveau d'eau dans le réservoir Nappe, conséquemment avantager la contribution de ce réservoir au débit de base.

La calibration par processus mène à une estimation meilleure de la courbe de crue printanière (Figure 2-22 représentant la moyenne interannuelle des débits) ainsi que la pointe de crue sur la station à l'exutoire. En effet, le NSE est de 0.86 et le biais relatif du débit maximum annuel est de 11%. Le modèle paramétré avec OptGlob est moins performant à estimer la crue avec un NSE de 0.71 et un biais relatif du débit maximum annuel de 43%. Cette configuration du modèle surestime la crue comme montrée dans la Figure 2-22, ainsi que les débits de probabilité de dépassement entre 1% et 6% comme le démontre la courbe de débits classés dans la Figure 2-23. OptProcess surestime aussi ces débits, mais avec un biais plus faible et simule adéquatement les débits de probabilité de dépassement entre 20% et 80%. Les deux configurations du modèle sous-estiment les débits d'étiage. Le biais relatif des débits d'étiages estivaux est le même (7%) pour les deux approches. Le biais des débits d'étiages hivernaux est amélioré de 5% avec OptProcess, comparativement à OptGlob, mais reste élevé (valeur autour de 16%).

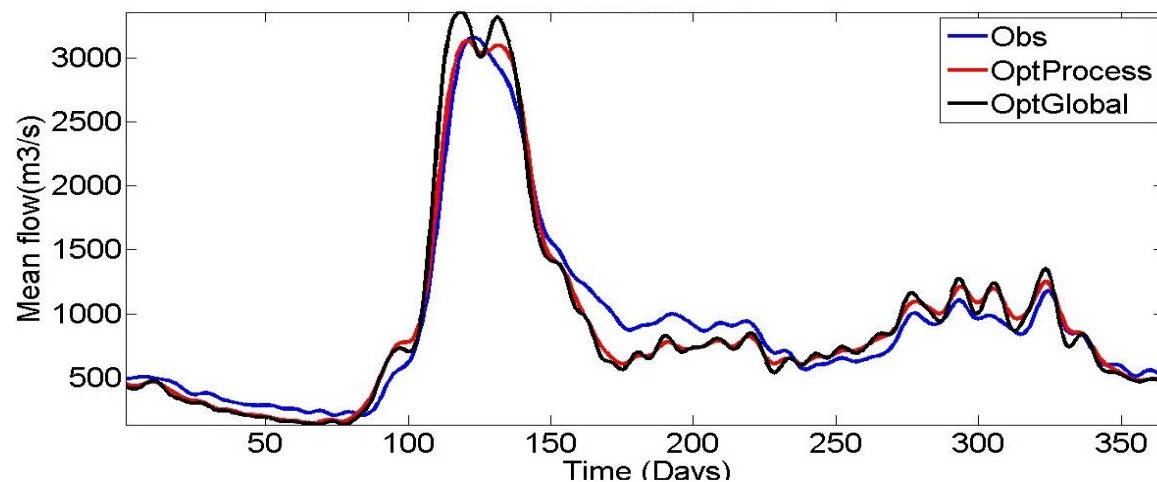


Figure 2-22 Moyenne interannuelle des hydrogrammes lissés observés et simulés à la station LSJ.

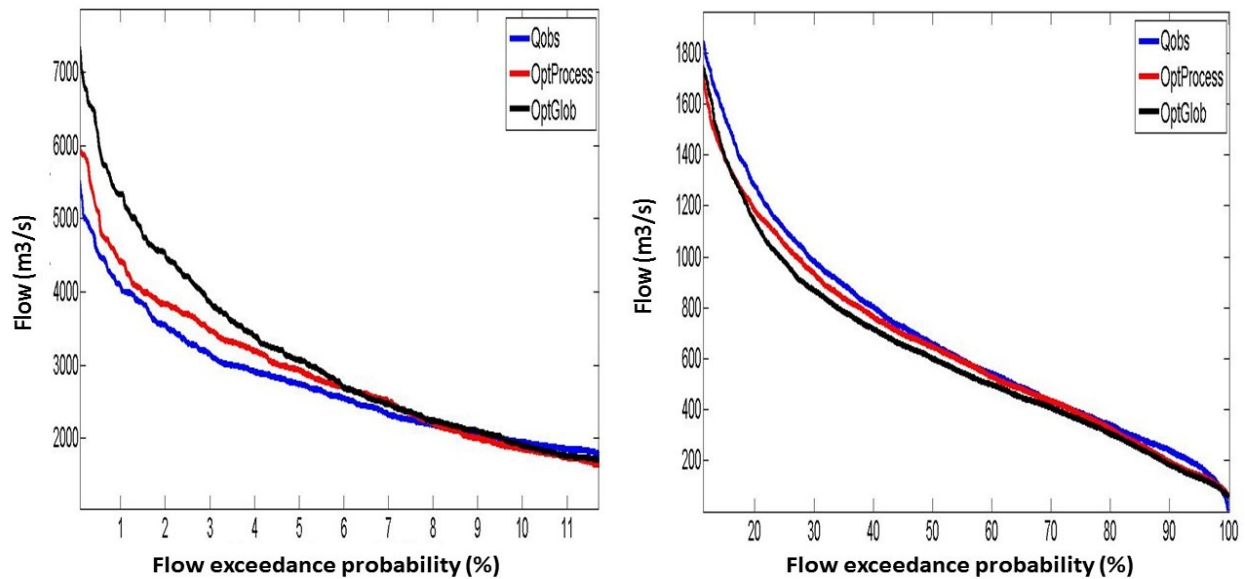


Figure 2-23 Courbe de débits classés observés et simulés par OptProcess et OptGlob.

Le modèle calibré par processus a une performance similaire sur la station ASHA utilisée pour l'ajustement des paramètres de transfert. La valeur du NSE est autour de 0.81, le biais du maximum annuel est de 18% et le biais des débits d'étiages hivernaux est de 33%. La performance du modèle calibré avec le critère global se dégrade sur cette station (NSE=0.44, biais du maximum annuel= 73%, biais des débits hivernaux=38%) et aussi sur toutes les autres stations hydrométriques du bassin (NSE moyen =0.23). La performance de la configuration OptProcess est acceptable sur toutes les stations avec NSE supérieur à 0.60 sauf sur les deux stations ASHA et METAB. Ces sous-bassins se caractérisent par une physiographie différente, avec une pente plus élevée que les autres sous-bassins, ce qui explique en partie la faible performance du modèle. À noter que les coefficients de correction des précipitations annuelles (COEP) et des températures (COET) en fonction de l'altitude n'ont pas été ajustés, ce qui pourrait expliquer la faible performance du modèle sur ces bassins.

Le modèle simule adéquatement la variabilité saisonnière (accumulation en hiver et fonte en printemps) du stock de neige avec une fonte rapide en clairière en comparaison en milieu forestier avec les deux jeux de paramètres (voir Figure 2-24). Les volumes d'eau simulés dans les deux réservoirs Sol et Nappe suivent aussi une certaine périodicité (augmentation en printemps et diminution en période d'étiage) avec les deux calibrations comme démontré dans la Figure 2-25.

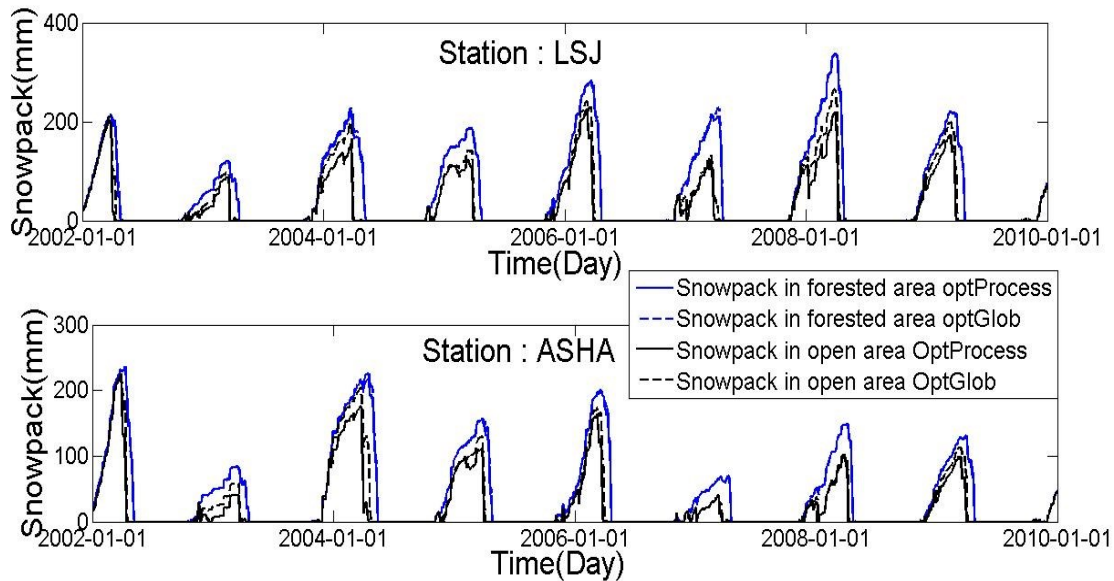


Figure 2-24 Stock de neige simulé avec le modèle calibré contre les débits (optGlob) et avec l'approche par processus (OptProcess) aux deux stations au bassin du Lac St-Jean: à l'exutoire (LSJ) et Ashapshuam (ASHA).

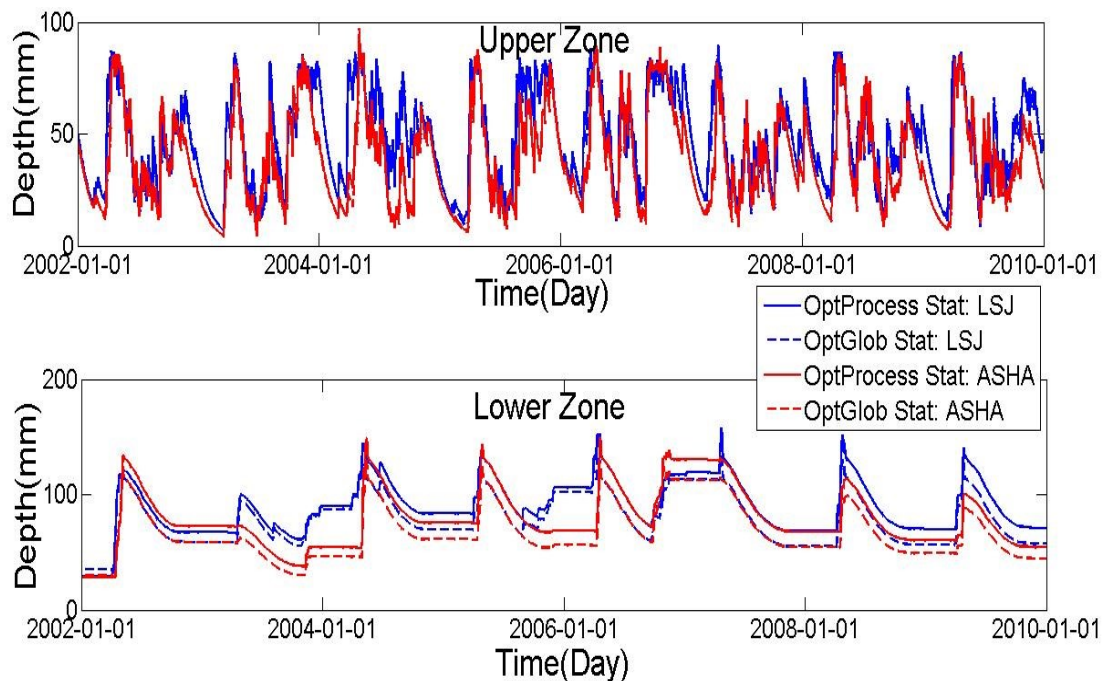


Figure 2-25 Niveau d'eau simulé avec le modèle calibré contre les débits (optGlob) et avec l'approche par processus (OptProcess) dans les deux réservoirs conceptualisé dans CEQUEAU aux deux stations au bassin Lac St-Jean: à l'exutoire (LSJ) et Ashapshuam (ASHA).

Les jeux de paramètres obtenus sont significativement différents (voir tableau 2-11 ci-dessous), ce qui entraîne deux comportements hydrologiques différents. OptGlob infiltre moins d'eau vers le réservoir Nappe avec un seuil d'infiltration maximal par jour XINFMA de 14mm/jour, tandis qu'OptProcess a une valeur de XINFMA de 24mm/jour. Les hauteurs des seuils de vidange des réservoirs conceptualisés sont plus petites avec OptGlob (HNAP=113mm et HSOL= 173mm) en comparaison avec OptProcess (HNAP=130 mm et HSOL= 206mm). Ceci signifie qu'OptProcess emmagasine plus d'eau dans les deux réservoirs. Toutefois, le temps de concentration du bassin (ZN) est de l'ordre de 2 jours ce qui signifie un écoulement rapide. En effet, le transfert dans CEQUEAU est relié à la notion du temps de concentration du bassin et le chemin le plus long entre le point le plus amont et l'exutoire du bassin (Morin et Paquet, 2007). Étant donné que le nombre de carreaux du chemin le plus long au lac St-Jean est de 49 carreaux, une partie de l'eau simulée avec OptProcess en amont sera 24 carreaux plus en aval à la fin de la journée. En revanche, avec ZN fixé à 5 jours avec OptGlob, le volume d'eau simulé en aval parcourt uniquement 9 carreaux en une journée ce qui peut expliquer la surestimation de la crue. En plus, OptGlob ne peut compenser avec l'évapotranspiration (i.e. maximiser EVNAP dont la borne supérieure fixée à 0.25) pour réduire le volume d'eau dans le réservoir Nappe. Ceci explique sa faible performance à l'échelle du bassin ainsi que la surestimation de la crue printanière. Cette performance indique que la calibration basée sur une seule variable (débit à l'exutoire), sous certaines conditions pour limiter la compensation de paramètres, n'est pas capable de reproduire de manière réaliste les mécanismes hydrologiques sous-jacents.

Table 2-11 Valeurs des paramètres du modèle CEQUEAU ajustées contre les débits (OptGLob) et avec l'approche proposée par processus (OptProcess).

Paramètre	OptGLob	OptProcess
STRNE	-1.41	-1.57
TFC	3.21	3.24
TFD	6.95	8.40
TSC	1.23	1.42
TSD	-0.77	-1.25
TTD	1.84	1.43
TTS	-2.57	-3.51
CIN	0.44	0.52
CVMAR	0.56	0.30
CVNH	0.52	0.41
CVSB	0.02	0.02
CVSI	0.51	0.37
XINFMA	14	24
HINF	82	82
HINT	82	79
HMAR	354	335
HNAP	113	130
HPOT	41	38
HSOL	173	206
HRIMP	5.03	3.37
EVNAP	0.23	0.23
Zn	5	2
EXXKT	0.80	0.02

La calibration par processus améliore la performance du modèle sur la station hydrométrique utilisée pour la calibration et aussi à l'échelle du bassin. Malgré les améliorations apportées au modèle, sa performance demeure très faible pour estimer les débits hivernaux, ce qui peut être dû à une déficience de sa structure et/ou au biais sur les mesures de débits et précipitations durant cette saison.

6.4.3 Conclusion

Ce chapitre de la thèse propose une approche de calibration multicritère par processus pour améliorer le réalisme d'un modèle hydrologique conceptuel. L'utilisation de la FDA a permis d'une part de conserver la forme des observations et d'extraire automatiquement certaines parties de l'hydrogramme (dans notre cas la crue) pour définir les critères de calibration. Le modèle calibré résultant de cette procédure simule adéquatement le mécanisme hydrologique dominant dans le bassin versant considéré avec de faibles performances sur certains sous-bassins. D'une part, le bassin lac St-Jean est un large bassin couvrant une superficie de 45 261 km². Pour cette raison, les conditions et le régime hydrologique peuvent différer d'un sous-bassin à l'autre ce qui peut expliquer en partie la faible performance sur ces sous-bassins. En effet, le modèle n'a pas été calibré contre les débits observés en ces bassins, mais plutôt avec des débits reconstitués par bilan hydrique, en particulier pour la station à l'exutoire (LSJ). D'autre part, le modèle est une représentation très simplifiée des processus hydrologiques régissant un bassin versant, ce qui peut entraîner des défaillances de structure. Autrement dit, la simplification des processus peut entraîner un modèle incapable de reproduire avec la même précision tous les processus hydrologiques et leurs variabilités spatiales dans un large bassin versant. Pour considérer cette diversité de régimes hydrologiques dans un bassin versant qui entraîne une variabilité spatiale des paramètres du modèle, une calibration multisite peut améliorer la performance du modèle à l'échelle du bassin. Ceci fait l'objet de la section suivante.

6.5. Objectif 5 : Calibration multisite du modèle CEQUEAU au bassin versant du lac St-Jean

Cette partie présente une procédure de calibration multisite du modèle hydrologique CEQUEAU au bassin du lac St-Jean. Les autres bassins sont omis de cette partie car on n'y dispose pas de plusieurs stations hydrométriques. Cette procédure utilise les deux algorithmes développés au premier objectif de la thèse et est basée sur le critère NSE à cause des observations manquantes à certaines stations hydrométriques, ce qui limite l'utilisation de la FDA. La méthodologie et les résultats en détail sont présentés ici puisque cette partie de la thèse n'a pas fait l'objet d'un article scientifique.

6.5.1 Méthodologie

Dans la présente partie on explore deux stratégies de calibration multisite en utilisant la procédure de calibration développée au chapitre 3 (voir sous-section 6.1.1-a et 6.1.1-b). On suppose que les sous-bassins sont similaires d'un point de vue hydrologique puisque les caractéristiques physiographiques et l'utilisation du territoire sont similaires pour chaque sous-bassin qui est en amont ou en aval de l'autre (i.e chaque sous-bassin déverse dans l'autre).

La première stratégie consiste à ajuster les paramètres du modèle sur l'ensemble des stations simultanément avec les deux étapes de la procédure. On définit la fonction objectif comme la somme du NSE sur toutes les stations :

$$OF = \sum_{i=1}^n NSE_{stat_i}(x) \quad (2.30)$$

Avec : n est le nombre de stations et $NSE_{stat_i}(x)$ est la valeur du NSE du vecteur de paramètres x à la station i .

La deuxième stratégie consiste en une calibration locale par station puis une calibration globale à l'échelle du bassin tel que présenté dans la Figure 2-26. La calibration locale est consacrée essentiellement à l'optimisation de l'espace des paramètres par station avec la procédure basée sur la simulation Monte Carlo et fonction de profondeur (présentée dans la sous-section 6.1.1-a). Pour chaque station, on obtient un ensemble de solutions optimales qu'on note X_{stat_i} . Ensuite, on détermine le vecteur moyen $\bar{X}_{stat_i}$ de cet ensemble pour chaque station i . Ceci nous permet d'analyser la variabilité spatiale des paramètres et calculer ainsi un coefficient de variation. Les paramètres j avec coefficient de variation inférieur à 10 % sont fixés à la valeur moyenne des

$\bar{X}_{stat,i,j}$ pour $i=1, \dots, n$ où n est le nombre de stations. Quant aux paramètres avec coefficient de variation supérieur à 10%, ils sont ajustés avec le Tabu Search (voir la sous-section 6.1.1-b) à l'échelle du bassin avec la fonction objectif définie par l'équation 2.30. Cette étape met l'emphasis sur les paramètres spatialement variables et vise à ajuster ces paramètres de manière à avoir une performance acceptable sur toutes les stations en réduisant le biais des paramètres qui peut résulter d'une simple moyenne. L'espace de paramètres considéré est borné par les vecteurs moyens obtenus pour chaque station tel que pour chaque paramètre j , la borne supérieure est le $\max_{i \in 1:n} \bar{X}_{stat,i,j}$, et la borne inférieure est le $\min_{i \in 1:n} \bar{X}_{stat,i,j}$. Cette deuxième stratégie consiste à calibrer à l'échelle du bassin que les paramètres spatialement variables, sauvant ainsi du temps de calcul en excluant de la calibration les paramètres variant peu d'une station à l'autre. L'hypothèse de travail est que l'optimisation des paramètres variant dans l'espace donnera un meilleur résultat que l'utilisation de la simple moyenne. Les deux stratégies sont appliquées avec une validation croisée, ce qui permet aussi une première analyse de base de la variabilité temporelle des paramètres.

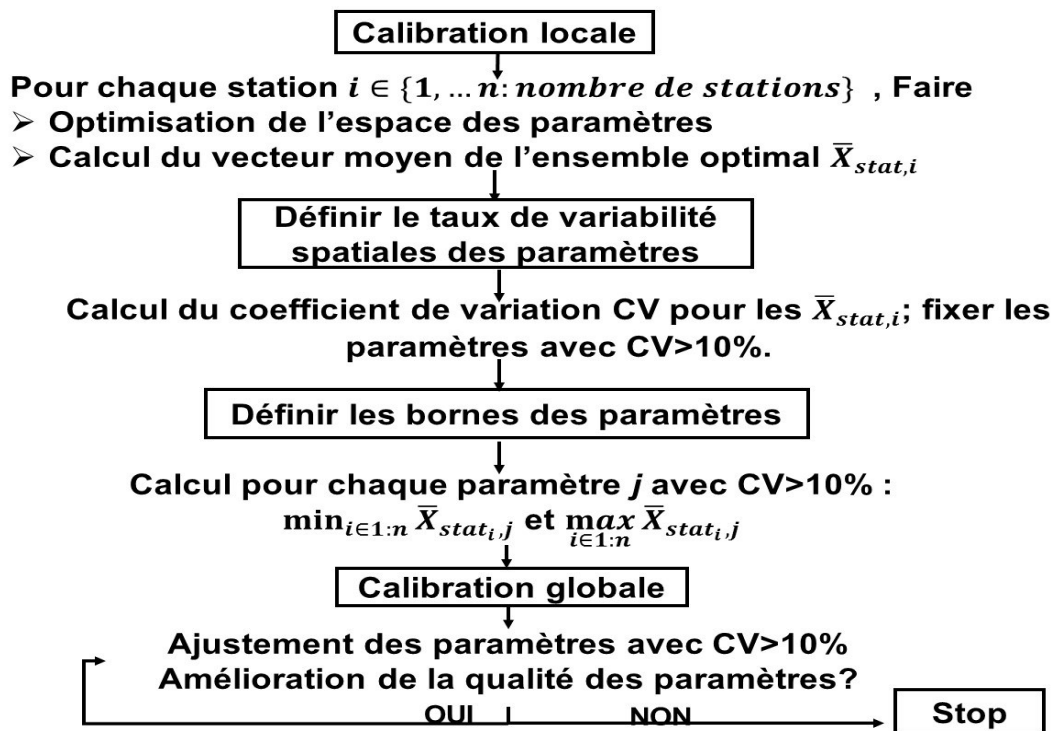


Figure 2-26 Organigramme de la calibration multisite locale suivie par une calibration globale.

6.5.2 Résultats et discussion

La première période de calibration est comprise entre 1998 et 2006 et la deuxième période entre 2003 et 2011. On considère trois stations hydrométriques : Mistassini2 (MISNI2, station la plus en amont du bassin), Ashapshuam aval (ASHA, station en aval du bassin) et la station LSJ (exutoire du bassin). Les autres stations du lac St-Jean sont utilisées pour la validation du modèle. Tous les paramètres sont fixés entre les bornes définies dans le tableau 2-2 (voir section 3) sauf le paramètre EVNAP dont la borne supérieure est fixée à 0.25. Le paramètre CVNB est fixé à 0 afin d'éviter le tarissement du réservoir NAPPE. Les Figures 2-27, 2-28 et 2-29 présentent les débits observés durant les deux périodes de calibration aux trois stations. On peut remarquer que les deux périodes sont différentes avec une crue printanière très importante durant l'année 1999 aux stations MISNI2 et ASHA en comparaison aux autres années.

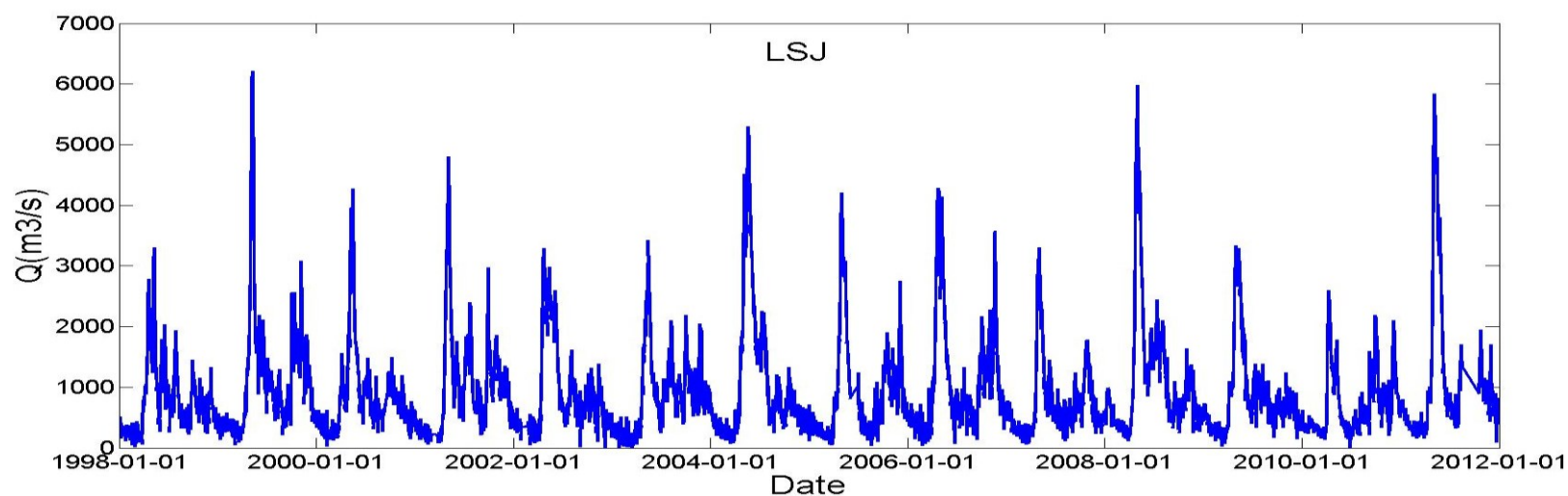


Figure 2-27 Série chronologique des débits observés à l'exutoire du bassin versant du Lac St-Jean.

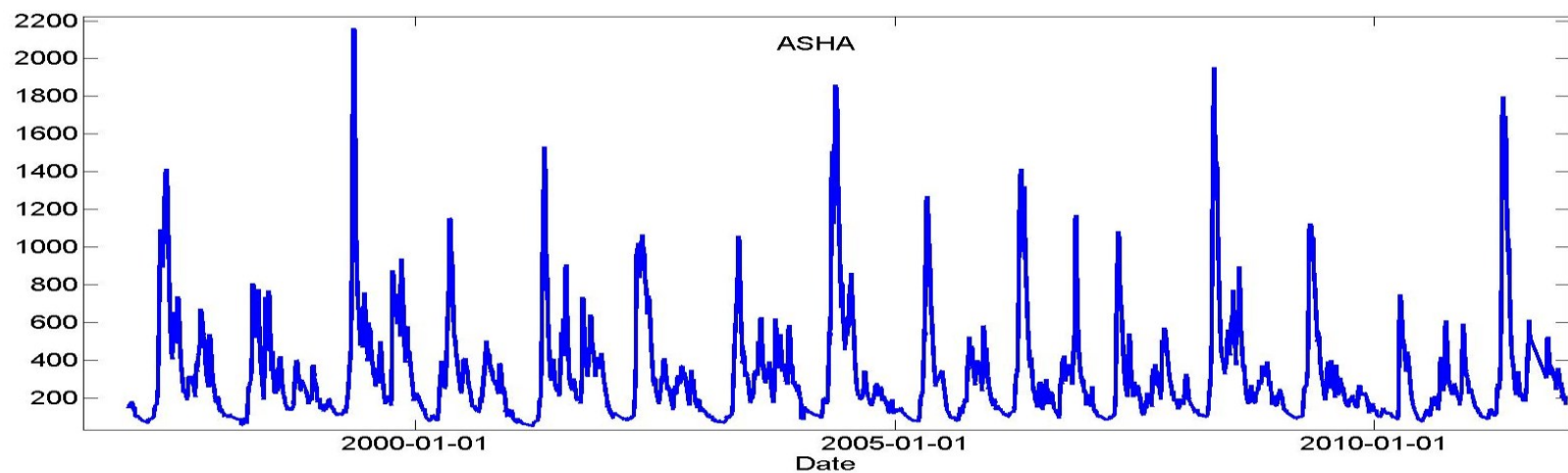


Figure 2-28 Série chronologique des débits observés à la station ASHA.

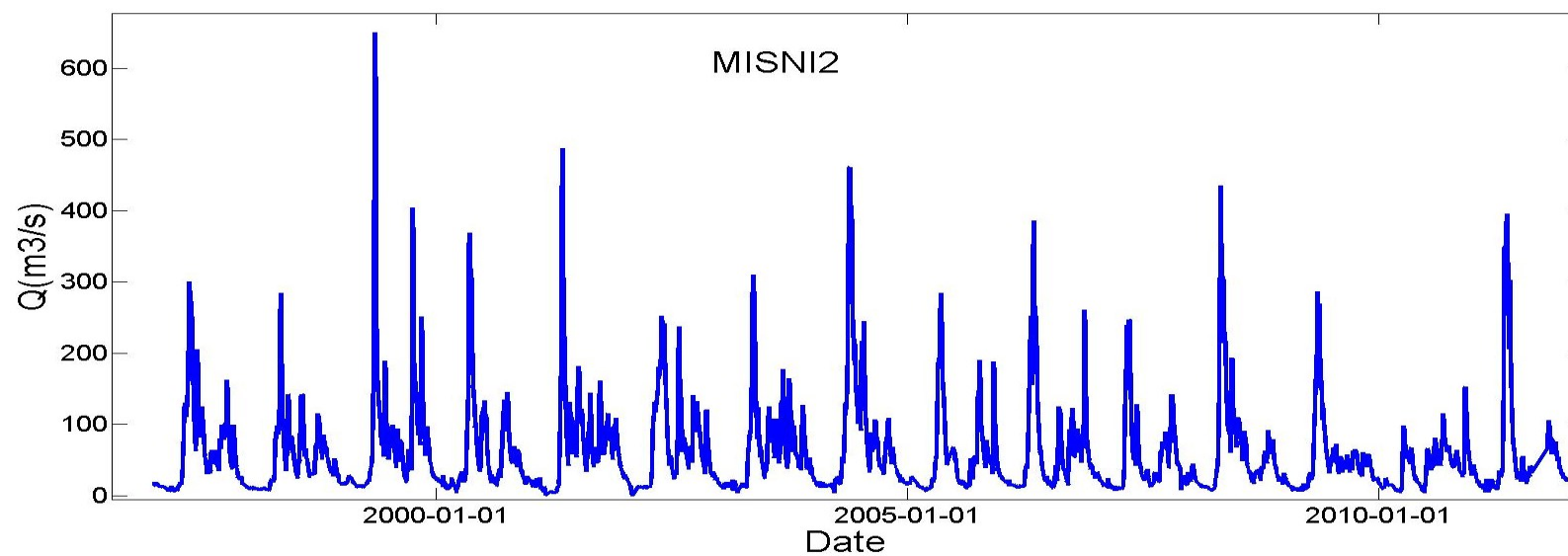


Figure 2-29 Série chronologique des débits observés à la station MISNI2.

Les deux tableaux 2-12 et 2-13 ci-dessous résument les valeurs des paramètres obtenus pour les deux périodes de calibration avec les deux stratégies de calibration. On note le vecteur de paramètres obtenus avec la première stratégie X_Glob1 et avec la deuxième stratégie X_Glob2 . Le vecteur $\bar{X}_{stat_i}$ est le vecteur moyen de l'ensemble optimal de la calibration locale sur la station i . Le vecteur $\bar{X}$ est la moyenne des vecteurs $\bar{X}_{stat_i}$. En rouge sont les paramètres ajustés à l'échelle du bassin lors de la calibration avec la deuxième stratégie.

Table 2-12 Résultats de la calibration multisite sur la période 1998-2006.

Paramètre	$\bar{X}_{stat_i}$			Coefficient de variation(%)	$\bar{X}$	X_Glob2	X_Glob1
	LSJ	ASHAV	MISNI2				
STRNE	-1.22	-1.25	-1.34	-5	-1.27	-1.27	-1.48
TFC	3.44	3.83	3.75	6	3.67	3.67	3.05
TFD	7.85	7.22	7.29	5	7.46	7.46	7.15
TSC	1.08	1.43	0.60	40	1.04	0.98	0.63
TSD	-0.70	-0.78	-0.62	-12	-0.70	-0.64	-0.94
TTD	1.59	1.48	1.38	7	1.48	1.48	1.26
TTS	-2.26	-2.64	-2.83	-11	-2.58	-2.76	-2.34
CIN	0.45	0.39	0.43	7	0.43	0.43	0.44
CVMAR	0.53	0.50	0.52	3	0.52	0.52	0.47
CVNH	0.11	0.06	0.08	29	0.08	0.11	0.10
CVSB	0.03	0.03	0.04	23	0.03	0.04	0.03
CVSI	0.39	0.43	0.26	23	0.36	0.28	0.29
XINFMA	20	24	22	9	22	22	16
HINF	78	74	72	4	74	74	76
HINT	76	79	77	2	78	78	79
HMAR	357	342	347	2	348	348	332
HNAP	121	127	114	5	121	121	118
HPOT	29	32	30	4	30	30	22
HSOL	212	195	210	5	206	206	222
HRIMP	5	5	5	7	5	5	4.86
EVNAP	0.15	0.15	0.15	2	0.15	0.15	0.17
NSE	0.85	0.87	0.78				

Lors de la première période de calibration qui est comprise entre 1998 et 2006, trois paramètres de fonte et trois coefficients de vidange présentent relativement une grande variabilité à l'échelle du bassin. Certains de ces paramètres (TSC, CVSB et CVSI) sont des paramètres sensibles selon l'analyse de sensibilité faite au premier objectif de la thèse (voir résultats dans le tableau 3-5). Malgré cette sensibilité, les deux stratégies de calibrations fixent les trois coefficients de vidanges à des valeurs similaires. On remarque une différence significative entre les valeurs fixées au paramètre de fonte TSC.

Table 2-13 Résultats de la calibration multisite sur la période 2003-2011.

Paramètre	$\bar{X}_{stat_i}$			Coefficient de variation(%)	$\bar{X}$	X_Glob2	X_Glob1
	LSJ	ASHAV	MISNI2				
STRNE	-1.29	-1.39	-1.05	-14	-1.24	-1.39	-1.43
TFC	2.83	3.33	3.74	14	3.30	2.85	3.72
TFD	7.24	7.46	7.49	2	7.40	7.40	8.85
TSC	1.04	1.29	0.73	28	1.02	0.73	0.78
TSD	-0.61	-0.90	-0.72	-19	-0.74	-0.66	-1.20
TTD	1.52	1.72	1.46	8	1.57	1.57	1.22
TTS	-1.98	-2.37	-2.33	-10	-2.23	-2.10	-2.99
CIN	0.39	0.32	0.43	16	0.38	0.35	0.67
CVMAR	0.37	0.37	0.48	16	0.41	0.48	0.35
CVNH	0.37	0.39	0.35	5	0.37	0.37	0.45
CVSB	0.03	0.04	0.06	33	0.04	0.03	0.04
CVSI	0.46	0.48	0.42	6	0.45	0.45	0.73
XINFMA	16	22	20	15	19	19	18
HINF	84	80	70	9	78	78	56
HINT	77	75	72	4	75	75	69
HMAR	357	346	341	2	348	348	375
HNAP	135	123	118	7	125	125	114
HPOT	27	32	29	8	29	29	25
HSOL	199	194	195	2	196	196	245
HRIMP	4	5	5	13	5	4.60	5.37
EVNAP	0.16	0.18	0.15	9	0.16	0.16	0.14
NSE	0.89	0.82	0.61				

La calibration locale sur la deuxième période de calibration entre 2003 et 2011 identifie 5 paramètres de fonte, 3 coefficients de vidange, le seuil d'infiltration maximale par jour (XINFMA) et HRIMP (Lame d'eau nécessaire pour que débute le ruissellement sur les surfaces imperméables (mm)) qui présentent une variabilité entre les trois sous-bassins. Ces paramètres sont des paramètres sensibles en particulier TFC, TSC, CIN, CVMAR, et CVSB (voir résultats analyse de sensibilité tableau 3-5). Les valeurs du paramètre HRIMP convergent autour de la même valeur. On remarque des valeurs similaires pour les paramètres CSVB (autour de 0.04) et XINFMA (autour de 18mm) avec les deux stratégies de calibrations.

Les calibrations locales sur les deux périodes fixent les paramètres régissant les hauteurs des réservoirs (HNAP, HMAR) et les seuils de vidanges (HINF, HINT, HPOT) à des valeurs avec des différences négligeables (de l'ordre de 1 à 4mm), que ce soit à l'échelle locale ou globale. Par exemple le paramètre HMAR est fixé à 348mm avec les deux calibrations à l'échelle du bassin et fixé à l'exutoire à 357mm. HPOT est fixé à 30mm sur la première période et à 29 mm sur la deuxième période. Les différences entre les valeurs du paramètre EVNAP sont aussi négligeables (0.15 pour la première période et 0.16 sur la deuxième période). On remarque aussi des différences négligeables pour les paramètres de fonte TSD (valeur de -0.64 sur la première période et -0.66 sur la deuxième période) et TFD (valeur autour de 7mm/jour/°C pour les deux périodes). La calibration globale sur la première période mène aussi à des valeurs similaires pour certains de ces paramètres (HINF, HINT, HPOT, HNAP, HPOT, HRIMP). On remarque aussi que, quel que soit la période de calibration ou la méthode, le paramètre CVSB converge vers une valeur de 0.03 ou 0.04 et HRIMP vers une valeur autour de 5. Ces paramètres ont donc une variabilité spatiale et temporelle négligeable et peuvent être désormais fixés à ces valeurs sur le bassin du lac St-Jean.

Le tableau 2-14 ci-dessous présente la performance globale des deux stratégies de calibration sur les deux périodes sur les différentes stations du bassin versant. Le NSE et le biais sont calculés sur la période entre 1998 et 2011. On note X_Glob1 et X_Glob2 les calibrations sur la première période (1998-2003) et X_Glob1* et X_Glob2* sur la deuxième période (2003-2011). Les trois premières stations sont les stations utilisées lors de la calibration.

La calibration globale sur la deuxième période est la calibration la moins performante. Le NSE à la station LSJ est de 0.69, et de 0.64 à la station ASHA et de 0.56 à la station MSINI2. La

performance de la calibration locale suivie par la calibration globale sur les deux périodes (X_Glob2 et X_Glob2*) est similaire sur les stations LSJ (NSE égale à 0.86) et ASHAV (NSE autour de 0.83). La paramétrisation du modèle avec X_Glob2* est moins performante sur les autres stations qu'avec X_Glob2 pour la simulation de la crue (faible valeur du NSE). Globalement les deux paramétrisations sont similaires avec des valeurs de biais équivalentes. La calibration globale sur la première période X_Glob1 mène à des résultats similaires à ceux de la calibration avec la deuxième stratégie sur la même période sur les stations utilisées pour la calibration (LSJ, ASHA, MISNI2). Toutefois X_Glob1 est meilleur sur l'ensemble des autres stations.

Table 2-14 Performance des deux stratégies de calibration multisite.

Station	NSE				BIAIS			
	X_Glob1	X_Glob2	X_Glob1*	X_Glob2*	X_Glob1	X_Glob2	X_Glob1*	X_Glob2*
LSJ	0.87	0.86	0.69	0.86	0.01	0.01	0.04	0.001
ASHA	0.84	0.83	0.64	0.82	0.02	0.003	0.03	0.01
MISNI2	0.78	0.77	0.56	0.69	0.04	0.05	0.09	0.05
ASHAM	0.63	0.61	0.41	0.59	0.18	0.17	0.13	0.17
METAB	0.49	0.49	0.25	0.45	0.20	0.18	0.16	0.19
MISBI3	0.81	0.79	0.51	0.69	0.01	0.01	0.04	0.002
MISNI	0.72	0.72	0.60	0.72	0.11	0.09	0.05	0.10
PPER	0.74	0.71	0.39	0.62	0.10	0.13	0.16	0.12
MISBI	0.82	0.80	0.53	0.74	0.01	0.03	0.06	0.02

Le passage d'une calibration locale à une calibration globale maintient globalement la même performance du modèle. On note une légère dégradation du NSE sur la première période sur la station ASHA (NSE de $\bar{X}_{ASHA}$ égale à 0.87 contre une valeur de X_Glob2 de 0.84 sur la même période) et sur la deuxième période sur la station LSJ (NSE de $\bar{X}_{LSJ}$ égale à 0.89 contre 0.86 valeur de X_Glob2*). Cette dégradation est peut-être due à un surajustement lors de la calibration locale sur ces stations, bien que les différences entre les valeurs de NSE soient minimales. Globalement on note des performances similaires avec les deux stratégies de calibration, mais la stratégie Glob1 est moins gourmande en temps de calcul.

Les trois Figures 2-30, 2-31 et 2-32 présentent les courbes de débits classés observés et simulés sur les trois stations. Ces courbes ont été divisées en deux segments : le premier segment contient les débits avec probabilité de dépassement entre 0 et 20% (figure à gauche) et le deuxième

(figure à droite) les débits avec probabilité de dépassement entre 20% et 100%. On remarque que sur les trois stations, le premier segment est adéquatement simulé avec le modèle configuré avec X_Glob1 et X_Glob2 pour la première période. Le deuxième segment de la courbe de débits classés est mieux simulé avec X_Glob1* et X_Glob2*, résultants de la calibration sur la deuxième période. En effet, les débits du premier segment proviennent principalement de la crue printanière de l'année 1999 dans toutes les stations. Cela explique que le modèle calibré sur la première période entre 1998-2003 est meilleur sur ce segment. Les deux dernières configurations X_Glob1* et X_Glob2*, en particulier X_Glob1*, surestiment le premier segment représentant les débits très élevés ce qui explique la faible valeur du NSE. Une légère différence est notée entre la calibration globale et la calibration locale suivie par une calibration globale.

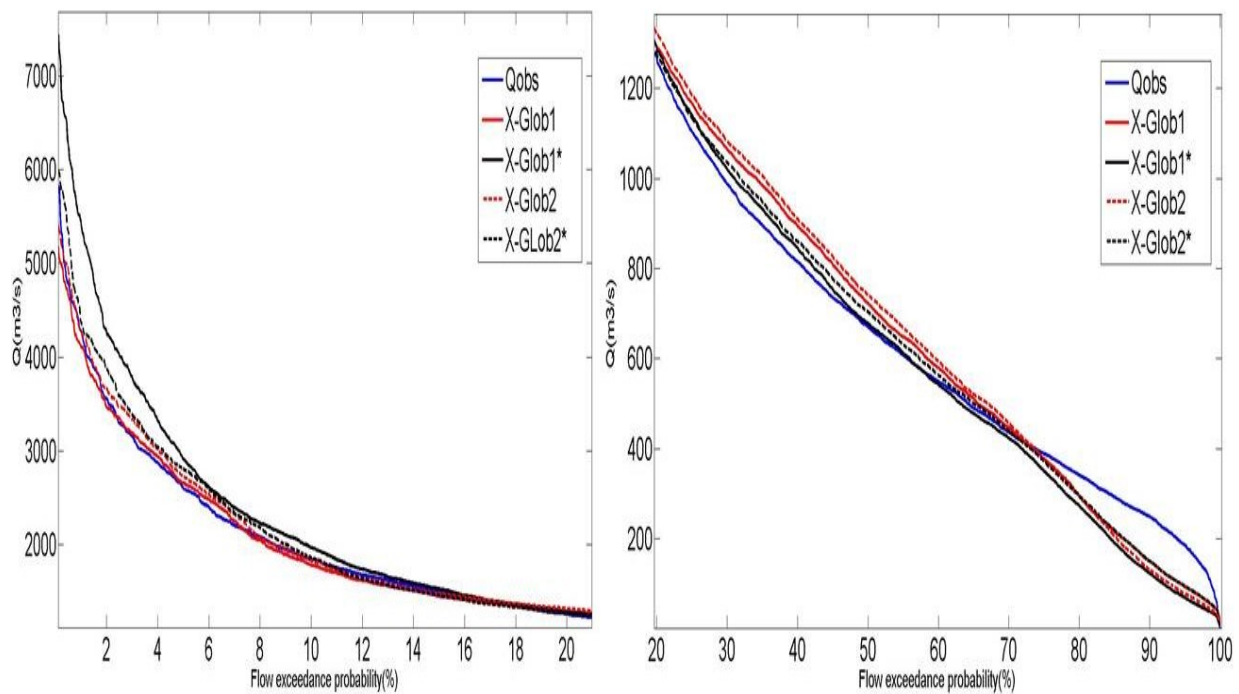


Figure 2-30 Courbe de débits classés observés et simulés avec la calibration multisite sur la station LSJ.

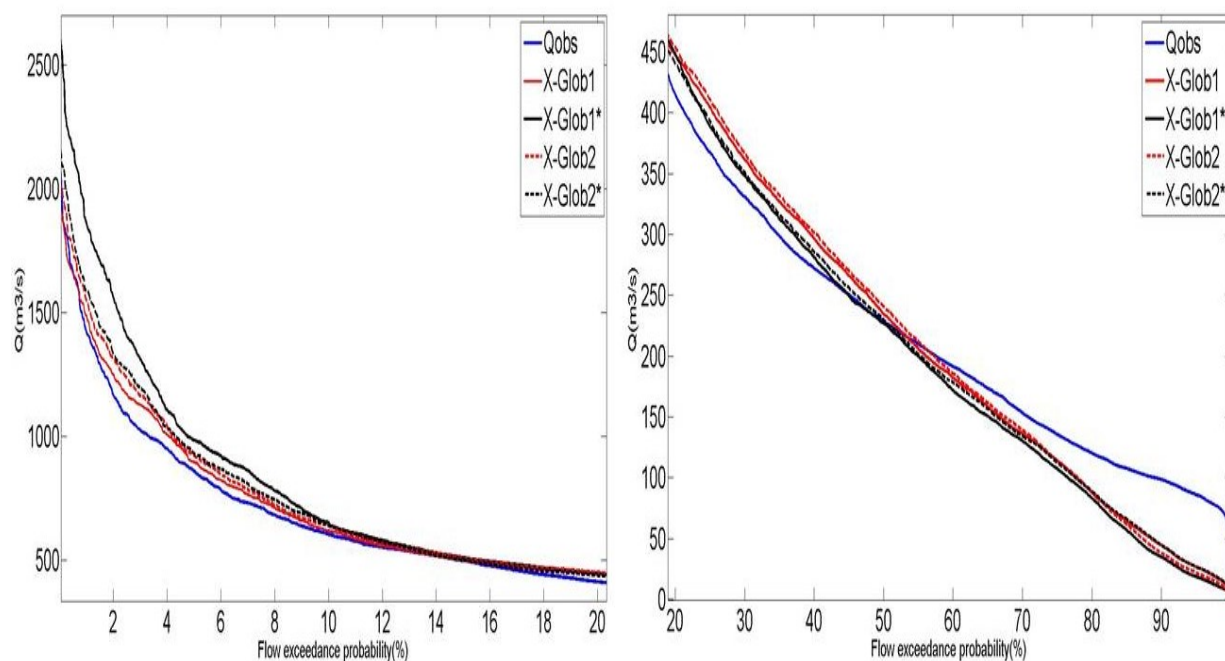


Figure 2-31 Courbe de débits classés observés et simulés avec la calibration multisite sur la station ASHA.

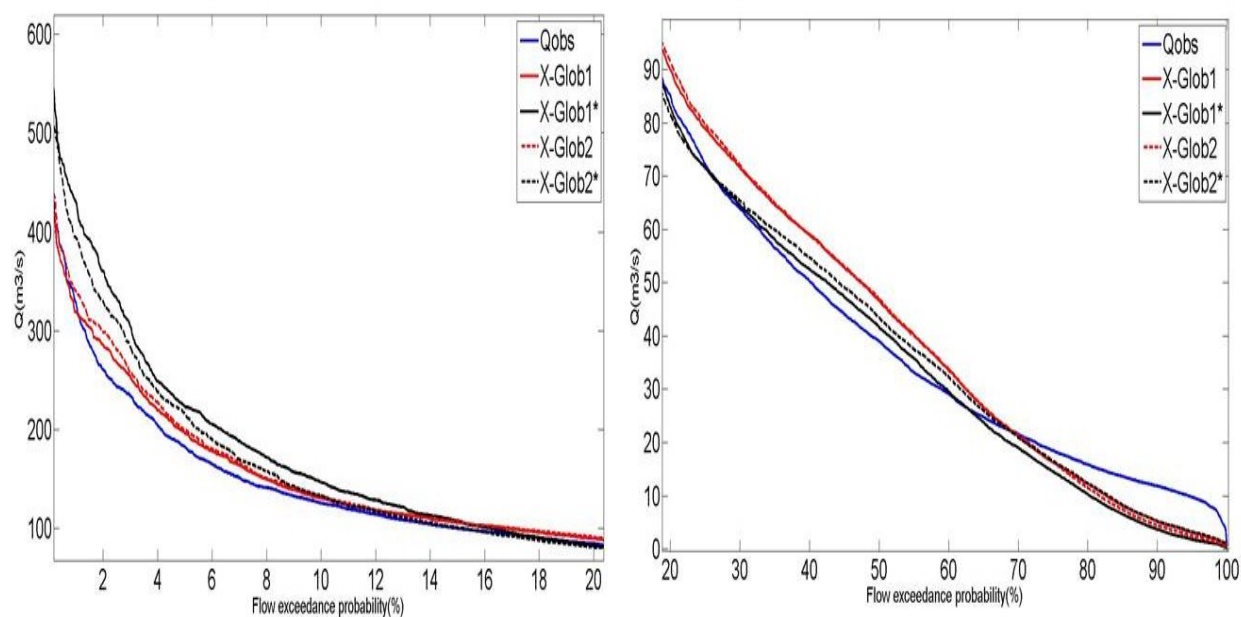


Figure 2-32 Courbe de débits classés observés et simulés avec la calibration multisite sur la station MISN12.

Le tableau 2-15 résume les valeurs du NSE obtenues pour les deux périodes de calibration pour la calibration globale avec la première stratégie et la calibration locale sur chaque station considérée. On présente aussi la performance du vecteur $\bar{X}$ qui est la moyenne des paramètres résultant de la calibration locale. X_Glob1 évalué sur les trois stations LSJ, ASHA et MISNI2 a la même performance que le vecteur de paramètres résultant de la calibration locale sur ces stations sur la première période. Sur la deuxième période, la calibration globale est moins performante que la calibration locale sur les stations LSJ et ASHA. Toutefois, la performance de X_Glob1 est acceptable avec NSE autour de 0.7 pour les deux stations. Sur la station MISNI2, le NSE est similaire. Ceci dit qu'une calibration en considérant différentes stations simultanément peut aboutir à un modèle performant similaire à un modèle calibré localement sur ces stations. Le modèle paramétré avec une moyenne des paramètres obtenus des calibrations locales maintient une performance similaire à la calibration locale. Toutefois, la calibration du modèle sur toutes les stations simultanément permet de gagner un temps de calcul considérable comparativement à celui associé à de multiples calibrations locales sur chaque station.

Table 2-15 NSE de la calibration globale et la calibration locale.

Station	Calibration 1998-2006			Calibration 2003-2011		
	X_Glob1	$\bar{X}_{stat_i}$	$\bar{X}$	X_Glob1	$\bar{X}_{stat_i}$	$\bar{X}$
LSJ	0.87	0.85	0.83	0.73	0.89	0.87
ASHA	0.86	0.87	0.86	0.68	0.82	0.81
MISNI2	0.78	0.78	0.76	0.64	0.61	0.79

6.5.3 Conclusion

Les modèles hydrologiques sont souvent calibrés sur la station à l'exutoire du bassin versant. Cette station peut être non révélatrice du régime hydrologique en entier d'un bassin versant à cause de la variabilité spatiale et l'hétérogénéité des conditions hydrologiques et physiographiques (Leta et al., 2017). Par conséquent, la calibration multisite permet de mieux contraindre le modèle à fin de capter les différentes représentations des processus hydrologiques dans le bassin versant (Zhang et al., 2008; Niraula et al., 2012; Saeidifrazad et al., 2014; Leta et al., 2017).

Dans cette partie de la thèse, deux calibrations multisites ont été comparées. La première consiste en une calibration sur trois stations simultanément. La deuxième s'étale sur deux étapes : une calibration par station pour délimiter l'espace des paramètres selon la variabilité spatiale des paramètres, ensuite une calibration simultanée sur les trois stations. Les deux approches ont été appliquées avec une validation croisée sur deux périodes de données.

Les résultats de la calibration par station révèlent le peu de variabilité spatiale ou temporelle de certains paramètres sur le bassin du lac St-Jean, notamment les hauteurs des seuils de déversement des réservoirs. Ces seuils peuvent être désormais fixés pour le bassin du lac St-Jean. On note une variabilité spatiale significative pour les paramètres de fonte qui sont des paramètres sensibles et qui dépendent aussi des conditions météorologiques, en particulier les températures observées. Les coefficients de vidange montrent aussi une variabilité spatiale et temporelle sauf le coefficient CVSB (coefficient de vidange basse du réservoir SOL).

La première stratégie de calibration est plus efficace et plus performante que la deuxième stratégie de calibration multisite. Elle permet de contraindre le modèle en considérant les différents aspects hydrologiques dans les stations considérées. Bien que cette stratégie semble être moins performante sur la deuxième période 2003-2011, le biais est similaire ou moins élevé sur certaines stations comparativement à celui de la deuxième stratégie (ex. ASHAM, MISNI). La deuxième stratégie est sujette au surajustement des paramètres dû à l'ajustement des paramètres à chaque station séparément lors de la première étape. Globalement, les résultats montrent la supériorité et l'avantage de la calibration multisite contre une calibration sur un seul site.

CHAPITRE 3 SIMULATION MONTE CARLO COUPLÉ AU TABU SEARCH POUR ESTIMER LES PARAMÈTRES D’UN MODÈLE HYDROLOGIQUE

Les travaux réalisés dans ce chapitre sont présentés sous forme d’article scientifique rédigé en anglais et dont le titre est :

Estimating hydrological model parameters by combining Monte Carlo Simulation with Tabu Search

Samah Larabi ^{*1}, André St-Hilaire ¹, Fateh Chebana ¹ and Marco Latraverse ²

¹INRS-ETE, Québec, Québec, Canada

²Rio Tinto Alcan, Saguenay, Québec, Canada

Cet article est soumis à *Advances in Water Resources Research Journal*. Cette version correspond à la version de l’article tel que soumis au journal.

Le contenu et les objectifs de cet article ont été discutés entre les différents auteurs. Le code du modèle hydrologique a été fourni par M. Latraverse. La définition de la méthodologie a été réalisée par le premier auteur en collaboration avec le professeur St-Hilaire et le professeur Chebana. La programmation de la méthode, l’analyse des résultats et la rédaction ont été effectuées par le premier auteur sous la supervision du professeur St-Hilaire. Une révision du contenu de l’article a été effectuée par les différents co-auteurs.

Abstract

Hydrological model calibration is a complex task that determines the performance of a parameterized hydrological model. Many performant calibration algorithms exist to address this issue. These algorithms randomly explore the parameter space within an upper and lower bounds set by the judgment of the modeller concerning feasible parameters values. Because of equifinality and the presence of unrealistic combinations, no global calibration method can provide one single best parameter set that adequately represents the physical processes of the basin. When the initial conditions and parameter space are not well defined, these algorithms may converge to unrealistic combinations even though they improve the objective function used as a performance metric in the process. The proposed approach in this paper to estimate model parameters consists in defining the parameters space first with a Monte Carlo simulation coupled with the concept of depth functions. Then, it explores the reduced parameters space with Tabu Search as a global metaheuristic search algorithm. The purpose of this algorithm is to provide a realistic set of parameters that can simulate flows adequately in the studied system. This type of algorithm traces its trajectory to provide one single parameter set. The proposed procedure is applied to calibrate the CEQUEAU model on four different basins located in Canada. The approach was applied with a split sample and a differential split sample procedure. The latter test aims to calibrate the model on a dry period and validate it on a wet period or vice-versa, in order to verify the robustness of the calibration. The Tabu Search algorithm is shown to be highly efficient, does not require a large number of iterations and converges to a realistic global optimum when the initial conditions are well defined. It performed better than an evolutionary global optimization algorithm that has been identified as particularly performant in an extensive study of automatic calibration method.

Keywords: Calibration, hydrological model, CEQUEAU, Tabu search, Monte Carlo Simulation, Depth functions.

1 Introduction

The estimation of hydrological model parameters is a complex task conditioning the performance of the model. Indeed, the quality of the simulations or the predictions of a model depends on how well the parameters were fitted (Duan et al., 1992). This task, namely calibration, consists of adjusting parameters based on historical data. It depends on different factors including the quality and the amount of input data, the objective function used, the strategy adopted and the initial conditions. It is essential to define all these factors adequately in order to obtain a model that reproduces the hydrological behaviour well of the watershed.

There are several studies that proposed algorithms and strategies for calibration. The traditional approach, which is still often used, is a manual calibration by trial and error. This approach consists in tuning the parameters based on graphical comparison of simulations and observations or depending on performance measures until a satisfactory solution is obtained. This procedure is subjective, depends on the expertise of the modeller, is not optimal and is time consuming (Arsenault et al., 2013). The difficulty of such an approach rises with the complexity of the model and the number of parameters.

Another alternative to estimate model parameters is to use automatic or semi-automatic approaches. In this case, calibration is considered as an inverse optimization problem (Moradkhani and Sorooshian; 2009). The automatic approach implies (Gupta et al., 1999): (1) selecting a period of calibration data; (2) making an initial guess of parameters ranges values; (3) selecting an objective function (measure of agreement between the simulations and observations); (4) selecting an automatic optimization procedure (a search algorithm). Hence, calibration is affected by each element selected to define the procedure. The first search algorithms that were introduced for automated and semi-automatic calibration were based on local search strategies. These algorithms cannot overcome the calibration difficulties such as the non-linearity of the hydrological models, the presence of multiple local optima (Duan et al., 1992), nor find a global optimum and they also depend on the initial conditions.

With the improvement of computer technologies, more complex and global optimization algorithms demanding computational power have emerged. In hydrological modelling, random search algorithms are widely used (Pechlivanidis et al., 2011). Most of these algorithms are population based. In particular, genetic algorithms (Holland, 1975; Seibert, 2000), or

evolutionary algorithms such as Shuffled Complex Evolution (SCE-UA) (Duan et al., 1992), and Covariance Matrix Adaptation Evolution strategy (CMA-ES) (Hansen and Ostermeir, 1996; Arsenault et al., 2013) are often used. These algorithms decrease calibration difficulties, as they allow finding an optimum closer to the global optimum. In the literature, many studies have introduced new algorithms and carry out a comparison with more traditional approaches (in particular SCE-UA) to prove their efficiency (e.g., Arsenault et al., 2013; Duan et al., 1992; Tolson and Shoemaker, 2007; Zhang et al., 2009). Most of these algorithms consist of exploring the entire bounded space randomly. Vrugt et al. (2007) proposed a multi-method approach called AMALGAM that combines different strategies to enhance the search. However, the resulted optimum solution depends on the bounds that are set depending on the modeller judgment of feasible parameter values. In addition, different sets of parameters located in different regions of the parameter space can provide the same solution, i.e. same objective function value; none of these global algorithms provide a single best parameter set. This phenomenon is termed equifinality. This implies the presence of some optimal unrealistic solutions that lead to a parameterized model not representing the natural system adequately and its hydrological processes. Because of all these reasons, these algorithms can converge to an unrealistic optimum. Hence, better constraining conditions for model parameters are necessary.

Bardossy and Singh (2008) showed that the set of the most adequate and most performing parameter vectors (i.e. vectors of parameters that lead to a performant model) are geometrically structured within the parameters space and proposed a procedure to investigate and identify the properties of this set with depth functions. Depth functions were introduced by Tukey in 1975 as a statistical measure that quantifies how a point is central in a data cloud in multidimensional space. Parameters leading to good model performance are considered performing. The goal of their study was not to find the parameter vectors which perform best for the calibration period but to find parameter vectors which: 1) lead to good model performance over the selected time, 2) lead to a hydrologically reasonable representation of the corresponding process, 3) are not overly sensitive: small changes of the parameters should not lead to very different results, 4) perform well for other time periods and might also perform well on other catchment (Bardossy and Singh; 2008).

The purpose of our current study is to set a procedure to estimate model parameters that are realistic and lead to a calibrated hydrological model that can simulate adequately the discharge of

the modelled catchment. The proposed procedure is a two-step calibration approach where the first step aims to fix the parameters bounds and give better constraining conditions subsequently used in a global optimization algorithm. As mentioned before, the calibration of the model can be highly affected by the boundaries imposed on the range of parameter values that are fixed subjectively. Based on the same concept proposed by Bardossy and Singh (2008), our proposed procedure uses depth function to define the optimal zone of the parameter space (called the optimal space) prior to calibrating the hydrological model. The assumptions of robustness, insensitivity and transferability over time periods of the best performing parameter vectors are investigated. A global sensitivity analysis is carried out before and after the identification of the optimal space to inspect the insensitivity of the parameters in the initial and the optimal space. The second step consists in adjusting the model parameters in the optimal space. The global metaheuristic algorithm Tabu Search (TS) is used at this stage. This type of search traces its trajectory within the restricted parameter space to find a global optimum. TS was used by Zhen and Wang (1996) to define the structure of the aquifer parameters. In fact, it was used to define an optimal zonation of the aquifer system. In the same study, TS did not require many iterations and objective function evaluations to perform well compared to the simulated annealing algorithm. The TS was also used to optimize the zonation and associated average groundwater pumping rates by Tung and Chou (2004) and in a multi-objective context for optimal design of groundwater remediation systems by Yang et al. (2013). In the present study, the TS is compared to an evolutionary algorithm: the Covariance matrix-Evolution strategy (CMA-ES) (Hansen & Ostermeir, 1996). The selection of CMA-ES as a comparative algorithm is based on a recent study (Arsenault et al., 2013) comparing many algorithms for automatic calibration of a number of models (including CEQUEAU used here), which showed that the Covariance matrix-Evolution strategy (CMA-ES) is a performant algorithm suited for large parameter space. Our proposed approach is applied with a split sample and a differential split sample procedure to calibrate 22 parameters of the conceptual semi-distributed CEQUEAU hydrological model. The first procedure consists of calibrating and evaluating the model performance on two independent periods of data with similar conditions; whereas the second procedure consists on calibrating and evaluating the model on two data sets representing different hydrological conditions and behaviour (Klemeš, 1986; Kirchner, 2006). Despite the fact that the latter form is more powerful

and incisive for evaluating model behaviour, it is not widely used (Kirchner, 2006; Seibert, 2003).

The paper is structured as follows. Section 2 describes the CEQUEAU hydrological model, the study areas, the data and the calibration evaluation structure. Section 3 describes the proposed procedure and the algorithms. Section 4 and section 5 discuss the application of the procedure to the case study and the results obtained.

2 Model, Study areas and data

In this section, the considered hydrological model and the study areas are presented. A structure of calibration/ evaluation illustrating the data used is described.

2.1 CEQUEAU model

CEQUEAU (Charbonneau et al., 1977; Morin and Paquet, 1995, 2007; St-Hilaire et al., 2003) is a deterministic semi-distributed hydrological model used to simulate and forecast water flow in any point of the modelled watershed. The model takes into account the physical characteristics and their variation in space and time by discretizing the watershed into equal squares called whole squares. For each whole square, the percentages of the areas covered by forest, wetlands, lakes and rivers are defined. Each whole square is subdivided into a maximum of four so-called partial squares. The model contains two essential components, a production function and a transfer function. The first aims to simulate the available water volume in each whole square and for each time step by calculating the hydrological budget. The vertical flow of water from the surface to the unsaturated and saturated soil horizons is conceptually modelled considering the ground as two interconnected reservoirs. The schematic representation is given in Figure 3-1. For each whole square at time t we have:

$$Q_t = P_t - ETP_t + (HU_t - HU_{t-1}) + HL_t - HL_{t-1} \quad (3.1)$$

where Q (mm) is the total runoff, P (mm) is liquid precipitation or snowmelt, ETP (mm) is the evapotranspiration, HU (mm) is the water accumulated in the upper reservoir, HL (mm) is the water accumulated in the lower reservoir and t is the model time step.

The second discretization into partial squares allows the model to simulate, with the transfer function, the horizontal flow considering the physical characteristics. With equation 3.2, the available water volume is routed downstream from a partial square to the other:

$$XKT_i = 1 - \exp\left(-\frac{EXXKT \text{ RMA3}_i}{\max(SI, Slac)} * \frac{100}{CEKM2}\right) \quad (3.2)$$

where XKT_i is the routing coefficient (fraction of available water routed downstream) of the i^{th} partial square, $EXXKT$ is the fitting parameter, $RMA3_i$ is the area of the basin upstream of the i^{th} partial square (km^2), SI is the area of surficial water upstream of the partial square, $Slac$ (km^2) is the area of surficial waters on the partial square i , and $CEKM2$ (km^2) is the area of the whole square (constant throughout de drainage basin).

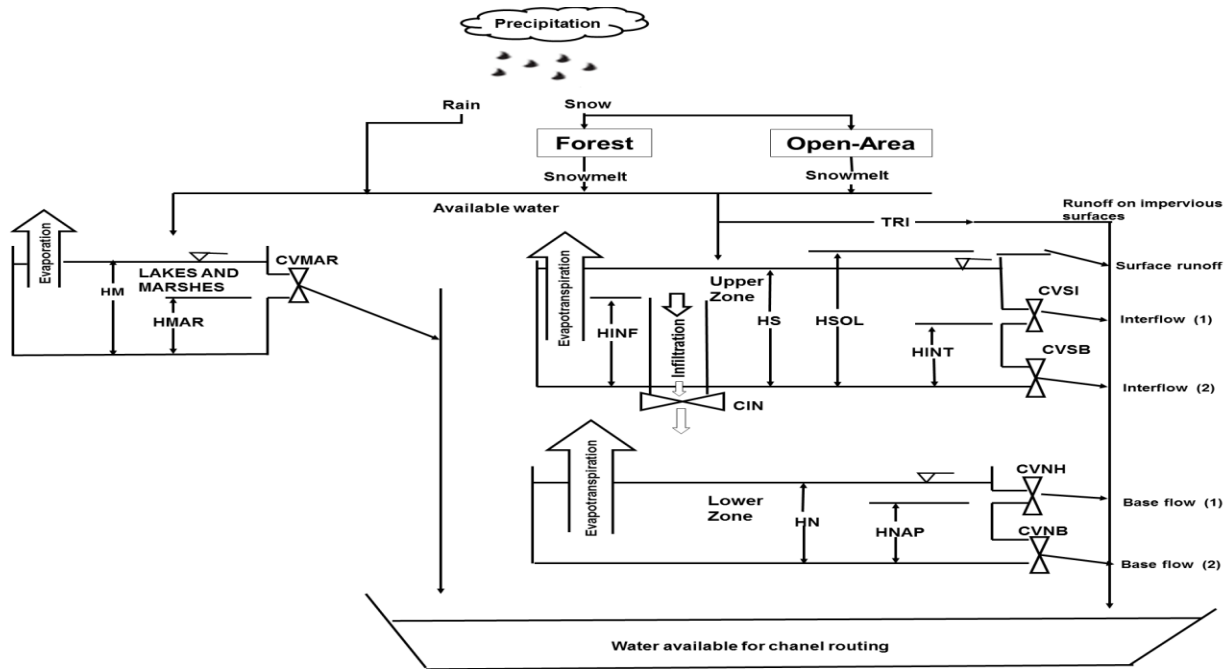


Figure 3-1 Diagram of the production function (adapted from Morin and Paquet 1995 and St-Hilaire et al; 2000).

CEQUEAU has a total of 28 describing the snowmelt process, infiltration, evaporation, the contribution of the base flow, interflow and surface runoff. These parameters are not known a priori and need to be adjusted in order to simulate flows that are similar as possible to the observed ones. Only 22 of these parameters described in Table 3-1 were used for this study. The

six parameters that are not calibrated in this study are the two coefficients of the Thornthwaite formula used to estimate potential evapotranspiration, temperature and precipitation lapse rate, percentage of impervious area in the basin. The required inputs are physiographic data, daily precipitation, the maximum and the minimum daily air temperature.

Table 3-1 CEQUEAU Parameters description and ranges.

Parameter	Description (unit)	Type	Range	
STRNE	Rain-Snow threshold (°C)	snow melt	-2	3
TFC	Melting rate in forested area (mm/°C/day)	snow melt	0	10
TFD	Melting rate in open area (mm/°C/day)	snow melt	0	10
TSC	Melting threshold in forested area (°C)	snow melt	-2	3
TSD	Melting threshold in open area (°C)	snow melt	-2	3
TTD	Cold content coefficient for the snowpack	snow melt	0	3
TTS	Priming threshold of the snowpack (°C)	snow melt	-5	2
CIN	Percolation coefficient from Upper Zone to Lower Zone	Interflow	0	1
CVMAR	Lakes and Marshes drainage coefficient	Interflow	0	1
CVNB	Lower- Zone lower drainage coefficient	Base flow	0	1
CVNH	Lower- Zone upper drainage coefficient	Base flow	0	1
CVSB	Upper- Zone lower drainage coefficient	Interflow	0	1
CVSI	Lower- Zone intermediate drainage coefficient	Interflow	0	1
XINFMA	Daily Maximum infiltration (mm/day)		0	40
HINF	Percolation threshold from the upper to the lower zone (mm)	Infiltration	0	100
HINT	Upper- Zone intermediate drainage threshold (mm)	Interflow	0	100
HMAR	Lakes and Marshes drainage threshold (mm)	Interflow	200	500
HNAP	Lower- Zone upper threshold (mm)	Base flow	20	200
HPOT	Threshold of evaporation at the potential rate (mm)	Evapotranspiration	0	80
HSOL	Height of the Upper Zone (mm)	Surface runoff	100	300
HRIMP	Quantity of water needed to runoff begins on sealed surfaces (mm)	Surface runoff	0	10
EVNAP	Percentage of daily Evapotranspiration taken from the Lower zone	Evapotranspiration	0	1

2.2 Study areas and data

The proposed methodology is applied to estimate the model's parameters for four catchments located in different regions of Canada with different hydrological properties. Table 3-2 presents some characteristics of the studied basins as well as the calibration and evaluation periods. Figures 3-2 and 3-3 show the location of these catchments. The time series of mean annual precipitations are derived from meteorological data observed during the entire period of calibration and evaluation given in Table 3-2. The two first drainage basins, namely Chute-Du-Diable and Passes-Dangereuses, are two subcatchments of the Lac St-Jean basin located in the region of Saguenay-Lac Saint-Jean in Québec. The Passes-Dangereuses is located upstream of Chute-Du-Diable. They empty into the St-Jean Lake which is the third-largest lake in the province of Québec. This area is mainly covered by forest. Catamaran brook, as the third catchment, is a small forested basin draining its water to the Miramichi River in the center of the province of New Brunswick, Eastern Canada. The fourth drainage basin Nechako River is located in British Columbia, Western Canada. This catchment is situated in the subboreal spruce biogeoclimatic zone and is described in more details by Macdonald et al. (2011). The main source of water in the four basins is the snowmelt occurring during the spring. The rainwater during autumn also contributes to streamflow but is less significant, particularly for the first two basins.

Table 3-2 Some characteristics of the basins studied.

Basin	Chute-Du-Diable (CD)	Passes Dangereuses (PD)	Nechako (Nech)	Catamaran (CAT)
Area(km ²)	9 701	11 485	47 200	52
Elevation(m)	215-615	450-914	986-1863	986-1863
Annual mean Precipitation(mm)	1000	1082	602	1180
Maximum Flow (observed date)	1 796 m ³ /s (12-05-1960)	2 339 m ³ /s (21-05-2014)	1338 m ³ /s (05-06-2007)	7.42 m ³ /s (30-10-2003)
Calibration Period	1985-1999	1985-1999	1985-1999	1998-2001
Evaluation Period	2000-2013	2000-2013	2000-2013	2002-2003
Wet Period	1989-1992	1999-2003	1987-1990	—
Dry Period	2007-2010	1996-1998	1995-1998	—

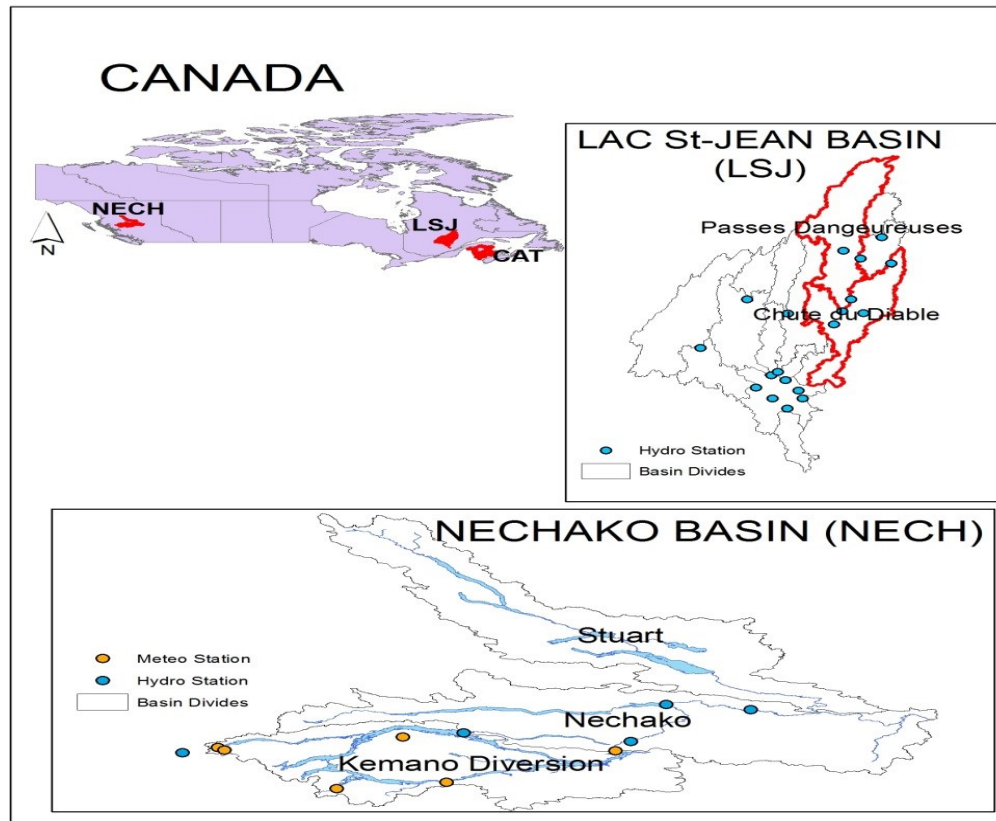


Figure 3-2 Map of the Nechako River in the Western Canada, Chute-Du-Diable and Passes-Dangereuses in Eastern Canada.

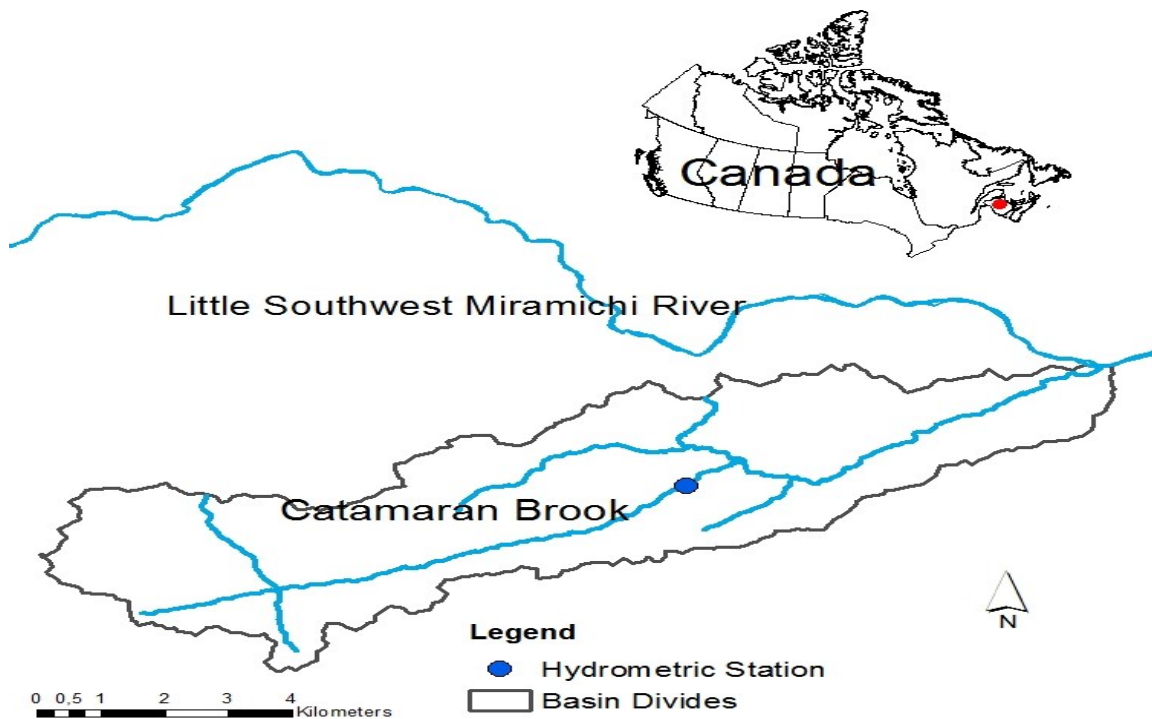


Figure 3-3 Map of the Catamaran Brook basin in the Eastern Canada.

2.3 Calibration/Evaluation structure

Two structures of calibration are used: the split sample and the differential split sample. The first one consists on dividing the data into two equal periods where one is used for calibration and the second is used for evaluation. The calibration and evaluation periods are given in Table 3-2 for each watershed using the first year of the calibration period as a spin up period for the model. The model is considered acceptable if it has similar performance in both periods. This condition is necessary and not sufficient (Klemeš, 1986). The differential split sample uses two independent data sets with different conditions. The model is calibrated on a dry period then validated on a wet period and vice-versa. This allows evaluating the robustness of the calibration through the ability of a model calibrated on a dry period (or wet period) to predict high flows (or low flows). Figure 3-4 presents an example of the chronological series of flows used for each period in the Chute-Du-Diable basin. Dry and wet periods used for each catchment are presented

in Table 3-2. This approach is not used to calibrate the model on the Catamaran Brook because of the short-long period from 1998 to 2003 of the available data.

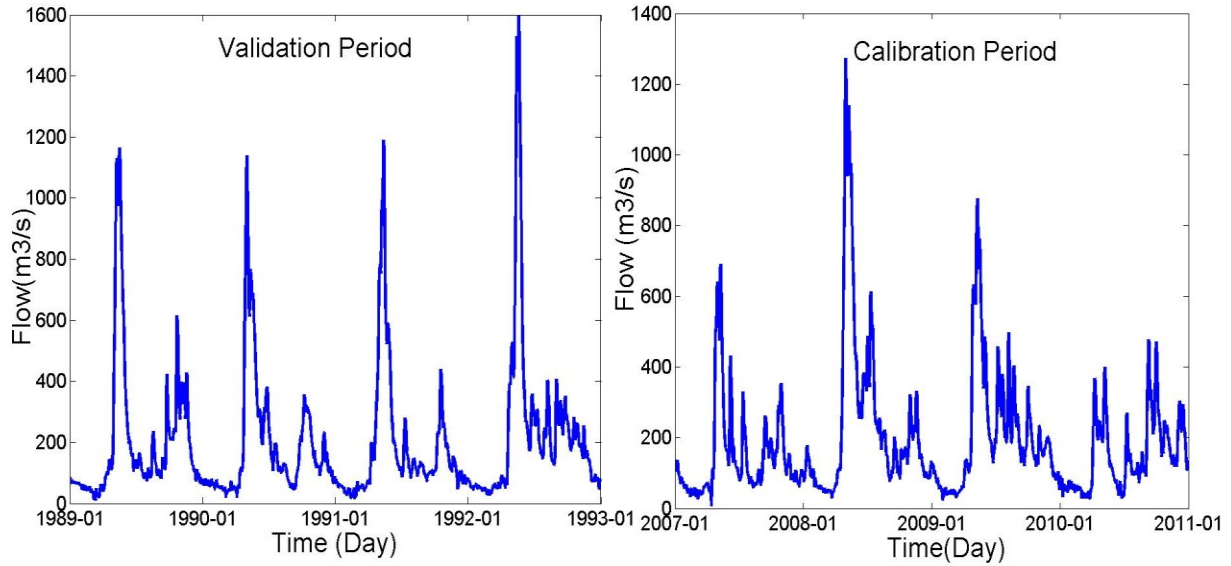


Figure 3-4 Chronological flow series used to calibrate the model with the differential split sample test on Chute-Du-Diable basin.

3 Methodology

This section describes the proposed procedure for parameter estimation and the algorithms used. The procedure consists of two parts. The first one aims to reduce the parameter space using a Monte Carlo simulation coupled with the concept of depth functions. This stage defines a new search space called the optimal space henceforth. This step of the methodology, as well as the global parameter sensitivity analysis in the resulting space, is presented in subsections 3.1 and 3.2. Subsections 3.3 and 3.4 present the global optimization algorithms TS and CMA-ES used to estimate the model parameters in the optimal space, which is the second step of the methodology. Subsection 3.5 presents the performance criteria used to evaluate and compare the models provided by TS and CMA-ES. The main steps of the proposed methodology are as follows:

1. Define the optimal bounds for each parameter (i.e. optimal space) with Monte Carlo simulations and Depth functions;
2. Explore the optimal space to define a set of an optimal parameter combination with a global optimization algorithm (TS or CMA-ES).

3.1 Refining the parameters space

In order to refine the parameters space and find realistic bounds of the parameters, depth functions are used within a Monte Carlo scheme. The notion of depth function was first introduced by Tukey in 1975 as a statistical tool to provide for a given F distribution on $\mathcal{R}^d$ a F -based center-outward ordering of points $\mathbf{x}$ in $\mathcal{R}^d$. The desirable properties of any depth function $D(\cdot; F)$ are (Zuo and Serfling, 2000):

- Affine invariance: depth $D(\mathbf{x}; F)$ does not depend on the coordinate system.
- Maximality at the center: depth $D(\mathbf{x}; F)$ is maximum only if $\mathbf{x}$ is the center (e.g. the point of symmetry with respect to some notion of symmetry).
- Monotonicity relative to the deepest point: depth decreases as points move away from the deepest point.
- Vanishing at infinity: depth of a point $\mathbf{x}$ goes to 0 when its norm $\|\mathbf{x}\|$ goes to infinity.

There are several depth functions (Tukey, 1975; Oja, 1983; Liu, 1990; Zuo and Sefling, 2000). The easiest one to compute is the Mahalanobis depth function which is based on the Mahalanobis distance (Mahalanobis, 1936). For a point $\mathbf{x}$ with respect to set $\mathbf{X}$ in the d dimensional space $\mathcal{R}^d$ the Mahalanobis depth is:

$$D(\mathbf{x}, \mathbf{X}) = \frac{1}{1 + (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})} \quad (3.3)$$

where $\boldsymbol{\mu}$ is the mean vector of $\mathbf{X}$ and $\boldsymbol{\Sigma}^{-1}$ is the covariance matrix of $\mathbf{X}$.

The Mahalanobis depth takes its value in $[0, 1]$. A low depth value, close to 0, means that the point is in the boundary while high depth close to 1 means that the point is close to the center (mean) of $\mathbf{X}$. In this study, the Mahalanobis depth is used because it can be computed for any dimension, the calculation is exact and not based on approximations. Other functions such as the halfspace depth are computationally very expensive for high dimensions or for a large number of points in $\mathbf{X}$ (Bardossy and Singh, 2008) or can be computed only in two or three dimensions

(e.g., Simplicial depth). In the present context, the dimension is very high (larger than 20), which makes the selection of the Mahalanobis distance appropriate. Moreover it allows the consideration of correlation between the parameters due to the covariance matrix. Thereby, it allows taking into account the interaction between the parameters during the identification of the optimal space.

Bardossy and Singh (2008) proposed an iterative procedure to identify the geometrical structure of the performing parameters using the depth function concept. The set of resulting parameters is robust, insensitive to small changes (i.e. small changes to parameters values should not lead to different model output) and transferable in time (i.e. similar performance in other periods).

In this paper, the same principle is adopted and the proposed algorithm is modified to refine the initial search space, such as the one defined in Table 3-1 for CEQUEAU. The main idea of the proposed procedure is to use a Monte Carlo simulation and update the parameter boundaries at each iteration, according to an objective function, such as the RMSE, or Nash-Sutcliffe Efficiency coefficient (NSE). Depth functions defining only the deepest points, in accordance with a given threshold, are selected to define the optimal boundaries. To illustrate the concept used, Figure 3-5 represents a projection of the parameters in a two-dimensional space.

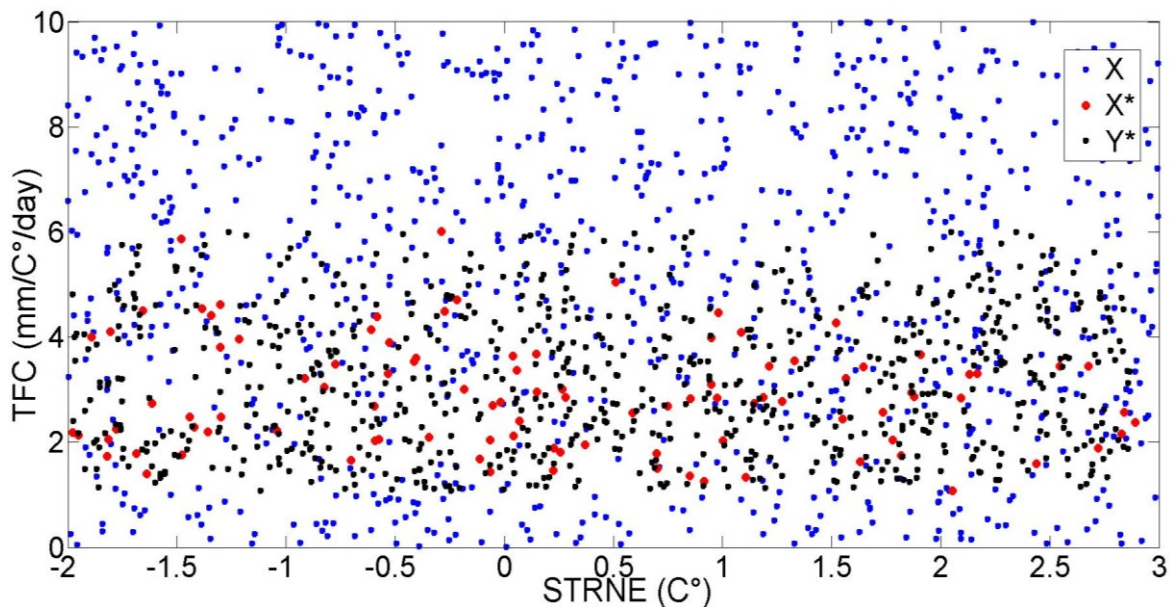


Figure 3-5 Scatter plot of model parameters, of the set X (N random snowmelt parameters), the set X* (top 10% of X), and the set Y* (the deepest points within the subset X*).

An initial ensemble $\mathbf{X}$ of a large number $N(=1000)$ of snowmelt parameters vectors is sampled uniformly using the bounds defined in Table 3-1 to initiate the Monte Carlo simulation. The mean NSE of this set is 0.46. The best performing vectors $\mathbf{X}^*$ (e.g., the top $P\%$, in our case $P=10$) are selected based on objective function (here the NSE). The boundaries of the parameters TFC and STRNE of $\mathbf{X}^*$ are respectively [1.07 5.18] and [-1.99 2.66] instead of the initial ranges [0 10] and [-2 3] respectively. New set $\mathbf{Y}$ of M parameters is generated between the new boundaries to optimize the attracted region of the space. With the Mahalanobis function, the set $\mathbf{Y}^*$ of the deepest points (points with depth superior to a threshold (here the first quartile) is selected. Generating points within the limits of $\mathbf{X}^*$ allows neglecting a region of the parameter space that does not include performant parameter values while generating other parameter combinations that lie in the vicinity of the performing parameters. The mean NSE of this new set $\mathbf{Y}^*$ is equal to 0.77, better than the previous value 0.46, and its standard deviation is 0.11. The performance is improved by sampling more parameters around $\mathbf{X}^*$. Actually, the deeper the points are within $\mathbf{X}^*$, the more robust the parameters are. In Table 3-3 we explore the performance of the set of deepest points according to different thresholds. The thresholds are fixed as 0.25, 0.5, 0.75, 0.95 quantiles of depths. The performance of $\mathbf{Y}^*$ (mean NSE) increase and the standard deviation of NSE decrease as we increase the threshold. We can conclude that the closer the points are to the center of $\mathbf{X}^*$, the more performing they are (Bardossy and Singh, 2008). The set $\mathbf{Y}^*$ becomes $\mathbf{X}$ and the process is repeated until the global performance does not improve. The attraction region of the parameters space is defined with limits of the final $\mathbf{Y}^*$.

Table 3-3 Performance of the set of the deepest points $\mathbf{Y}^*$ according to different thresholds fixed as quantiles.

Depth Threshold	Mean NSE	Standard deviation of NSE
Q 0.25	0.77	0.108
Q 0.5	0.82	0.055
Q 0.75	0.86	0.030
Q 0.95	0.88	0.020

The major difference between the ROPE algorithm and our proposed modified version is updating the parameter bounds at each iteration. In fact, in the ROPE algorithm the bounds are constant and the $\mathbf{Y}^*$ set is generated under the condition that the depth of each vector $\mathbf{Y}^*$ is

superior to a threshold. Satisfying this condition may considerably increase the computing time. Moreover, the deepest point within subspace $\mathbf{X}^*$ will have coordinates closer to those of $\mathbf{X}^*$ because of the ‘vanishing to infinity’ property of the depth function. Therefore, changing the bounds and selecting deepest points from those that are closer to $\mathbf{X}^*$ is efficient and not time consuming.

The algorithm inspired by the ROPE algorithm (Bardossy and Singh; 2008) is as follows:

1. Fix the parameter limits
2. Generate uniformly a set $\mathbf{X}$ of N (here $N=1000$) random parameter vectors within this limits
3. Evaluate every parameter vector performance with an objective function and identify the P % best parameters (in our case $P=10\%$)
4. Define the limits of the subset of best parameters selected in step 3
5. Generate a set $\mathbf{Y}$ of N random parameter vectors within the new ranges
6. Calculate the Mahalanobis depth for each vector in $\mathbf{Y}$ with respect to the top 10% best parameters, and select the vectors $\mathbf{Y}^*$ having a depth superior to a threshold r (we use the first quartile of depths as a threshold)
7. Evaluate every parameter vector of $\mathbf{Y}^*$ with an objective function
8. $\mathbf{Y}^*$ becomes $\mathbf{X}$, and the process is repeated until the performance of $\mathbf{Y}^*$ does not improve.

In order to quantify the size and allow an objective comparison between the bounds defined in Table 3-1 and the final optimal bounds inferred by the proposed procedure, we use an index calculated for each parameter i and defined as follows:

$$G_i = \frac{\text{length of optimal range of parameter } i}{\text{length of initial range of parameter } i} \times 100. \quad (3.4)$$

This criterion is used to check whether the best parameter sets are located in a small zone of the parameter space or can be found anywhere in the space defined according to expert judgment of feasible/reasonable parameter values. If the latter is true then the reduction of the parameter space would not be significant as the sampling is uniform and random when the Monte Carlo simulation is initiated and the bounds won’t change in the later iterations.

The resulting reduced space in this section is explored by the algorithms presented in the next subsections 3.3 and 3.4 in order to define a realistic performing parameter set that inherits the characteristics of the optimal space.

3.2 Sensitivity analysis

The deepest points are robust, insensitive to small changes and transferable in time (Bardossy and Singh; 2008). A simple global sensitivity analysis is carried to identify insensitive parameters in the optimal space. A global approach One-Factor-at-a-Time (OAT) coupled with a simulation by Latin-Hypercube (Griensven et al.; 2006) is used. This approach allows quantifying the sensitivity of each parameter in the entire space with a low cost. The OAT consists of a local sensitivity analysis repeated several times for different parameter vectors to cover the entire space by a Monte Carlo simulation. This type of sampling is very complex and needs many iterations, which motivates the use of a Latin Hypercube (LH) sampling. The LH strategy uses a Monte Carlo simulation with a stratified sampling. The space is subdivided into K -equal intervals. Each interval is sampled only once. So the required number of model simulations to evaluate the sensitivity of P parameters is $K*(P+1)$. For each interval j , and for each parameter x_i , the partial effect $S_{i,j}$ of the parameter i is:

$$S_{i,j} = 100 \times f_i^{-1} \times \left| \frac{M(x_1, \dots, x_i + f_i, \dots, x_p) - M(x_1, \dots, x_i, \dots, x_p)}{|M(x_1, \dots, x_i + f_i, \dots, x_p) + M(x_1, \dots, x_i, \dots, x_p)|/2} \right| \quad (3.5)$$

where M is the performance of the model, f_i is the perturbation fraction of the parameter i , j is the interval index of the LH subdivision. The global effect S_i of each parameter i is the mean of the partial effects $S_{i,j}$. Parameters with high value are the most sensitive.

3.3 Tabu Search algorithm

The Tabu search algorithm (TS), introduced by Glover in 1986, is based on the descent search concept integrating different techniques to improve the search. Tabu Search may be viewed as an iterative procedure that starts with an initial solution, evaluates its neighborhood $N(\mathbf{x}) \subset S$ (the search space), and replaces the current solution with the best solution in $N(\mathbf{x})$ even if it does not improve the objective function when compared to the initial solution until a stop criterion is

satisfied. A neighborhood is an application that associates to a solution $\mathbf{x}$ a subset $N(\mathbf{x}) \subset S$. Generally it is defined as an ensemble of moves m that switch from a solution $\mathbf{x}$ to $\mathbf{x}' \in N(\mathbf{x})$:

$$N(\mathbf{x}) = \{\mathbf{y} \in S \mid \mathbf{y} = \mathbf{x} \oplus m \text{ with } m \in M(S)\}. \quad (3.6)$$

where m is a move and $M(S)$ is the ensemble of the possible moves in the search space S . A move can be defined as changing the value assigned to a variable, deleting or adding an element or interchanging position of parameters (Glover, 1990). Replacing a solution $\mathbf{x}$ with a solution $\mathbf{x}'$ that might be less performant, prevents the search to stop at the first optimum found. This approach may lead the search to be trapped in a cycle because $\mathbf{x}$ might be the best solution in $N(\mathbf{x}')$. Then the search would go back to this point. Thereby the tabu list is a key element in the TS algorithm. This list noted $\mathbf{T}$ is a memory that saves the $|\mathbf{T}|$ last solutions explored. These solutions are considered forbidden and the search is prohibited to explore these solutions in the next iterations. Thus cycles of length less or equal to $|\mathbf{T}|$ are prevented. Practically, rather than saving these solutions that may be space consuming, the inverse of the moves that switches from a solution to another are saved (Glover et al., 1993).

In what follows, we describe a basic and often used form of Tabu Search that was applied to our parameter estimation problem. (Other more advanced forms of TS are described in the book by Glover and Laguna, 1997). In the context of our current problem, $\mathbf{x}$ is a combination of parameters. We define a move as a changing value assigned to a parameter i in $\mathbf{x}$ by a predefined fraction f_i of the range of the parameter i . Then every move is characterized by a set of parameter index and a sign corresponding to the direction of change (i.e. indicating whether this parameter was increased or decreased by f_i). The aim of such a strategy is to allow the search to adjust each parameter depending on whether its nominal value (value assigned in the current solution) needs to be decreased or increased in interaction with other parameters. For every m in $M(S)$, there is an inverse move $m^{-1} \in M(M(S))$ such us $(\mathbf{x} \oplus m) + m^{-1} = \mathbf{x}$. The tabu list is defined as a FIFO (First In First Out) list containing the inverse moves m^{-1} . These moves are banned for the next $|\mathbf{T}|$ (number of parameters to be adjusted) iterations. However, the search might make an undesirable prohibition by banning some moves that may lead to good solutions. An aspiration criterion is integrated in the search that aims to accept these moves. The aspiration criterion accepts a forbidden move only if it leads to an optimum better than the one found so far.

In order to make the search more intelligent, two different strategies named *intensification* and *diversification* are used. These two techniques allow the search to benefit from the information gathered from the previous iterations other than the objective function. The intensification aims to force the search to explore the promising region deeply and diversification stimulates the search to explore the unvisited regions. In this context, a memory $\mathbf{M}$ is used to save the moves that frequently lead to good solutions. For example, one form of intensification (which we have employed here) focuses on those parameters that have been changed and led to improving the model performance. It incites the algorithm to optimize only these parameters. In this approach the algorithm constructs the neighborhood $N(\mathbf{x})$ of the current solution $\mathbf{x}$ by modifying only the frequently changed parameters. These parameters are saved in a frequency memory $\mathbf{F}$. The other parameters are fixed. The intensification is followed by the diversification strategy, inciting the search to optimize the parameters that were not changed in the process so far. Thus the algorithm alternates among phases of learning, intensification and diversification. Figure 3-6 illustrates the flow chart of the algorithm.

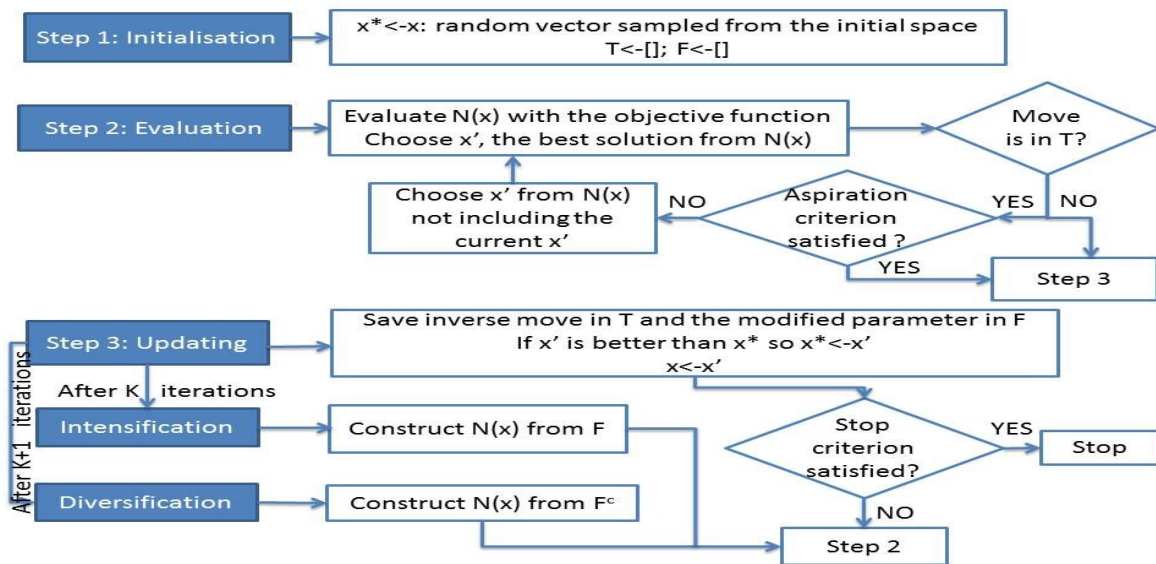


Figure 3-6 Tabu Search algorithm flow chart.

The Tabu Search algorithm stops if one of these criteria is met: 1) performance does not improve in 3 successive iterations; 2) the maximum number of iterations or the maximum number of function evaluations is attained.

3.4 Covariance Matrix Adaptation-Evolution strategy (CMA-ES)

The CMA-ES (Hansen and Ostermeir, 1996, 1997, 2001) is an evolutionary algorithm designed for the non-convex and non-linear problems. The class of algorithms represented by CMA-ES class of algorithm is inspired by the theory of biological evolution. The basic principle underlying CMA-ES consists of sampling a new population at each iteration of the best individuals according to an objective function under a new mean and a covariance matrix. Indeed, the procedure uses an adaptive process of the covariance matrix coupled with an evolution strategy. At each iteration, the algorithm samples a new population of the best individuals from the current population under a new mean and a covariance matrix until a stop criterion is satisfied. Details can be found in Hansen and Ostermeir (1996, 1997, 2001).

3.5 Performance criteria

The performance criterion used to calibrate CEQUEAU in the three procedures described above is the Nash Sutcliffe efficiency (henceforth denoted simply by NSE) defined as follows:

$$NSE = \left[1 - \frac{\sum_{i=1}^n (q_{ci} - q_{oi})^2}{\sum_{i=1}^n (q_{oi} - \bar{q}_o)^2} \right] \quad (3.7)$$

Where q_{ci} and q_{oi} are the simulated and observed flows at the time step i .

Two additional criteria, the Kling-Gupta efficiency (KGE) and a bias factor are used for general evaluation and comparison of the hydrological simulation results obtained using the parameters provided by the TS and CMA-ES. The bias used is defined as follows:

$$Bias = \left| \frac{\sum_{i=1}^n (q_{ci} - q_{oi})}{\sum_{i=1}^n q_{oi}} \right| \quad (3.8)$$

with q_{ci} and q_{oi} are the simulated and observed flows at the time step i .

Gupta et al (2009) discussed the utility of analysing the three components of the NSE: bias, correlation and variability when calibrating with this criterion. The intention of introducing

other performance criteria such as KGE is to allow for a more complete objective comparison between the solutions provided by both algorithms instead of a comparison based on a single criterion (the NSE). The KGE (Gupta et al., 2009) is a decomposition of the NSE to take into account its different components and overcome the propensity of NSE to penalize the performance mostly for poorly simulated high flows. Indeed, the KGE reflects the compromise between three different components the correlation, the bias and a measure of relative variability in the simulation and observation:

$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (3.9)$$

where $\alpha = \sigma_s / \sigma_o$ is the ratio of the standard deviation of simulation and observation; $\beta = \mu_s / \mu_o$ is the ratio between the mean of the simulation and observation and r the linear correlation between simulation and observation.

4 Results

This section presents the results of applying our proposed procedure with the two calibration/evaluation structures. Subsections 4.1, 4.2 and 4.3 describe the results obtained with the split sample approach and subsection 4.4, the results obtained with the differential split sample. The three algorithms: Monte Carlo Simulation coupled with depth function, TS and CMA-ES were easily adapted for parallel computing. Thereby each realization of the model was executed on a single processor. This allowed a considerable decrease in the time of execution.

4.1 Identification of the optimal space

For each basin, the algorithm starts with a set of parameters randomly generated within the initial ranges (Table 3-1). Figure 3-7 shows the improvement of the mean NSE through iterations until it becomes stationary for all basins. The algorithm starts with an initial set having a negative mean NSE converges to positive values over 0.8 in some basins. It converges in 23 (Nechako River) to 53 (Catamaran Brook) iterations.

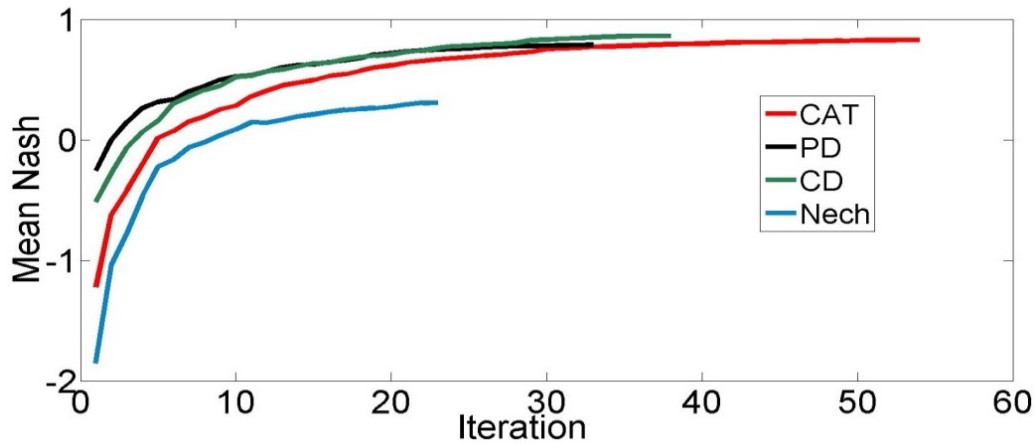


Figure 3-7 The evolution of the mean Nash of the ensemble Y^* in the identification of the optimal space process.

Table 3-4 summarizes the boundaries of the resulting optimal space and Figure 3-8 presents the G index illustrating the ratio between the length of the new range and the initial range for each parameter. Generally the initial space is reduced in average from 80 % to 50 % in all basins. The algorithm converges to realistic bounds. The bounds differ from one basin to another and concord with the conditions of each drainage basin. For the TFC parameter, for example, the optimal ranges are [0.9; 1.54] mm/°C/day at the Nechako River while at the Catamaran Brook it ranges in [2.6; 3] mm/°C/day. This means that the melting is faster in the Catamaran Brook than in the Nechako River. This behaviour is totally in agreement with the characteristics of both watersheds. The Nechako River is a watershed with altitude ranging between 986 and 1863 m while Catamaran Brook is a small forested basin located further south than the Nechako River, with altitudes ranging between 120 and 343 m. The space of some parameters (STRNE, TFC, and CVSB) was considerably reduced in all basins with comparable percentages. In fact, these parameters are the most sensitive.

Table 3-4 Parameters optimal ranges.

Parameter	Initial Range		Optimal Range							
			CD		PD		Nech		CAT	
STRNE	-2	3	-1.11	-0.11	-1.6	-1.1	2.14	2.79	-0.31	0.42
TFC	0	10	2.14	3.04	2.0	2.7	0.90	1.54	2.63	3.02
TFD	0	10	3.31	7.11	5.8	8.5	2.79	6.70	6.37	7.61
TSC	-2	3	0.35	1.47	0.4	1.2	1.43	2.58	1.10	1.47
TSD	-2	3	-0.67	1.10	-1.4	0.0	-1.14	1.05	-0.87	0.06
TTD	0	3	1.23	2.25	1.2	2.3	0.73	2.41	1.21	1.96
TTS	-5	2	-3.51	-1.95	-3.7	-1.8	-1.67	0.88	-2.42	-1.22
CIN	0	1	0.31	0.69	0.2	0.7	0.27	0.85	0.46	0.64
CVMAR	0	1	0.29	0.64	0.2	0.6	0.22	0.79	0.37	0.60
CVNB	0	1	0.07	0.09	0.1	0.6	0.09	0.62	0.07	0.08
CVNH	0	1	0.28	0.68	0.3	0.7	0.25	0.80	0.28	0.57
CVSB	0	1	0.04	0.05	0.03	0.05	0.08	0.16	0.03	0.04
CVSI	0	1	0.20	0.58	0.2	0.5	0.16	0.71	0.46	0.59
XINFMA	0	40	12.89	26.51	11.2	26.9	11.58	30.95	14.90	20.41
HINF	0	100	57.03	80.74	67.0	83.6	19.86	73.25	56.29	63.32
HINT	0	100	63.12	84.18	67.6	88.5	38.83	86.28	53.53	60.41
HMAR	200	500	291.0	420.31	293.9	442.5	280.22	451.37	300.89	377.25
HNAP	20	200	89.97	152.26	73.7	154.6	53.55	157.18	115.11	145.92
HPOT	0	80	23.37	41.35	32.5	50.7	50.28	66.41	15.77	26.12
HSOL	100	300	142.4	213.00	156.0	235.5	137.50	241.25	165.19	209.94
HRIMP	0	10	3.27	7.03	2.8	6.8	2.36	7.76	3.02	5.87
EVNAP	0	1	0.34	0.70	0.3	0.7	0.22	0.80	0.38	0.66
Iteration			37		32		23		53	

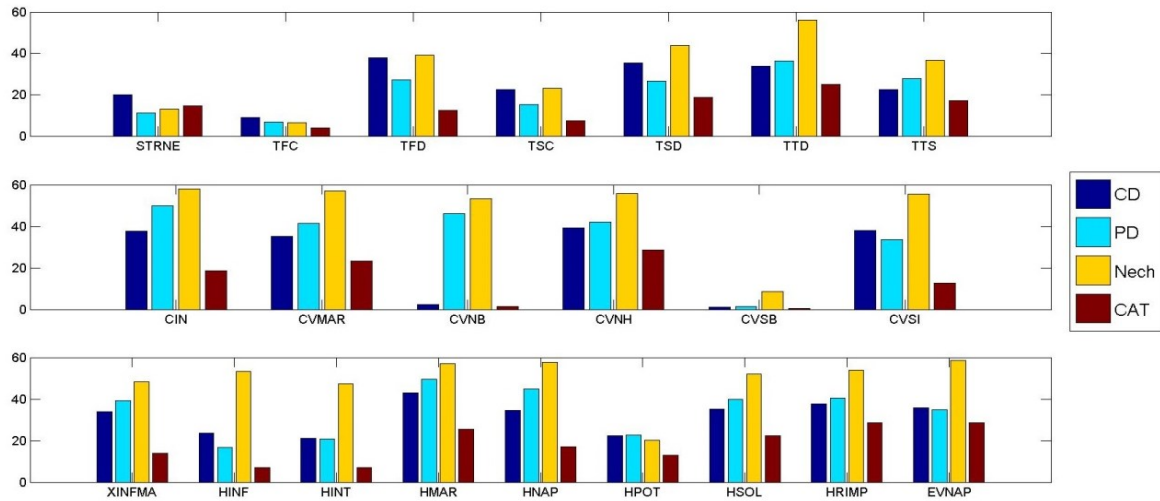


Figure 3-8 Percentage (G) of the initial space retained by parameter and by basin.

In Table 3-5, S and Sopt are the quantification of the sensitivity of each parameter in the initial space and in the optimal space. High values of S and Sopt indicate that a small perturbation of the parameter influence the performance of the model. These two indexes are used to rank the parameters by importance based on sensitivity. The drainage coefficient (CVSB) and the snow melt rate parameter in the forest area (TFC) are the most influential parameters standing with high values of S. The ranges of these parameters are reduced up from 90 to 99%. Only 1 % of the initial range of the CVSB parameter is retained in the Chute-Du-Diable, Passes-Dangereuses, Catamaran Brook, and 9% in the Nechako. This parameter controls the lower interflow in the upper (unsaturated) soil reservoir. For TFC, only 4% (Catamaran) to 9% (Chute-Du-Diable) of the initial range is retained. The drainage coefficients and snow melt parameters are also influential. The rest of parameters have negligible influence. More than 60% of the initial ranges of these parameters are rejected, depending on their ranking except for Nechako River where on average 60% of the initial space is retained. In the optimal space, the sensitivity is decreased for all parameters except two parameters: CVSB and CVNB. Despite the fact that the sensitivity of these parameters is increased in Nechako River and Catamaran brook, their ranges were considerably reduced.

Indeed in the Nechako River, an increase of the sensitivity of some parameters (e.g. STRNE) in the optimal space is noticed. Uncertainty on model input (daily precipitations and temperature) could explain in part this increase. CEQUEAU parameters show very low sensitivity in the Catamaran Brook basins compared to other basins. This can be due to the short period of data (only 5 years) used to evaluate the sensitivity of these parameters on this basin whereas in the other basins 13 years of data were used.

Table 3-5 Sensitivity of parameters in the initial space (S) and in the optimal space (Sopt).

Parameter	S				Sopt			
	CD	PD	Nech	CAT	CD	PD	Nech	CAT
STRNE	13	115	17	0.02	1	2	50	0.03
TFC	193	589	100	0.07	2	6	57	0.03
TFD	38	31	19	0.01	0	1	8	0.02
TSC	151	233	44	0.05	2	4	25	0.03
TSD	17	89	117	0.01	0	1	6	0
TTD	176	49	3	0.01	0	0	6	0
TTS	44	74	16	0.03	0	1	11	0.03
CIN	171	19	11	0.01	2	2	13	0.03
CVMAR	162	7594	2	0.04	1	1	0	0.01
CVNB	168	115	44	0.03	51	5	61	0.62
CVNH	6	0.04	0	0	0	0	0	0
CVSB	857	1140	287	0.17	81	60	470	1.60
CVSI	155	101	12	0.01	1	2	16	0.02
XINFMA	0	2	0	0	0	0	0	0
HINF	1	1	1	0	0	0	1	0
HINT	2	5	0	0	0	0	0	0
HMAR	0	0	0	0	0	0	0	0
HNAP	0	0	0	0	0	0	0	0
HPOT	4	8	1	0	0	0	1	0
HSOL	0	0	0	0	0	0	0	0
HRIMP	5	10	2	0	0	0	2	0
EVNAP	4	3	1	0	0	0	2	0

The most influential parameters have index in red.

Any solution taken from the optimal space will have NSE equal to at least the mean NSE of the ensemble Y^* obtained at the last iteration. Table 3-6 describes the distribution of the NSE of the parameters sets at each watershed. The standard deviation is small and the 0.05 and 0.95

quantiles values are close to the Mean NSE value. The parameters sets have equivalent performance. The NSE is over 0.8 in the Chute-Du-Diable and Catamaran brook basins. In the Passes-Dangereuses basin, it is around 0.79 and is around 0.31 in Nechako River. Even in some cases where the mean NSE is over 0.8, calibration is needed instead of taking any random solution. Although the entire space is reduced by at least 50%, the optimal upper bound of some parameters is not decreased considerably. For example, the optimal upper bound of EVNAP is fixed between 0.6 to 0.8, which, for these drainage basins, is considered a high value for a percentage of daily evapotranspiration taken from the lower reservoir.

Table 3-6 Performance of the inferred optimal set with the 0.05 and 0.95 quantiles by the first step of the proposed methodology.

Basin	Mean NSE	0.05 Quantile	0.95 Quantile	standard deviation
CD	0.86	0.83	0.88	0.015
PD	0.79	0.77	0.80	0.011
Nech	0.31	0.21	0.39	0.058
Cat	0.83	0.81	0.84	0.008

4.2 Calibration without refining the parameter space

To illustrate the problem mentioned whereby optimization algorithms can converge to non-realistic set of parameters when the ranges are not well defined, a calibration is carried out without refining the parameters space. Both algorithms CMA-ES and TS are executed in the space of parameters defined with the ranges in Table 3-1 to adjust the model parameters and simulate discharge on the four catchments. The sets of parameters resulted are noted OptTS and OptCMAES. Table 3-7 presents the performances of the set of parameters obtained. Both algorithms converge to sets of parameters with relatively equivalent performances in all basins.

Table 3-7 Performance of the set of parameters obtained with TS and CMA-ES executed in the initial space.

Basin		Calibration			Evaluation		
		NSE	Bias	KGE	NSE	Bias	KGE
CD	OptTS	0.91	0.03	0.94	0.89	0.04	0.92
	OptCMAES	0.89	0.03	0.94	0.89	0.03	0.94
	OptTS	0.83	0.06	0.89	0.81	0.04	0.89
PD	OptCMAES	0.81	0.06	0.89	0.80	0.05	0.89
	OptTS	0.71	0.04	0.84	0.69	0.20	0.73
Nech	OptCMAES	0.69	0.08	0.80	0.58	0.22	0.69
	OptTS	0.89	0.03	0.93	0.79	0.10	0.81
CAT	OptCMAES	0.85	0.01	0.92	0.76	0.08	0.84

These solutions have acceptable performances but are not always realistic. Indeed, some parameters were assigned values that do not match the behaviour of the basins. For example, in Passes-Dangereuses basin, CMA-ES fixed CIN to 0 meaning that there is no infiltration from the Upper reservoir to lower (see Figure 3-1). TFD is set to 0 which reflects that there is no melting in the open area. EVNAP was set to 1 in Chute-Du-Diable, Passes-Dangereuses and Catamaran by TS and CMA-ES. This parameter represents the fraction of daily evapotranspiration taken from the lower reservoir. Fixing this parameter to 1 means that the only source of evapotranspiration is coming from the lower reservoir, which is not the case. Usually, evapotranspiration of the surface and in the upper reservoir contributes more significantly to the daily evapotranspiration than the lower reservoir. In Chute-Du-Diable basin, CMA-ES fixed TFD to values over 8 mm/°C/day as a rate of snowmelt in the open area. This value is too high considering the characteristics of this basin, but there is very little open (i.e. unforested) area.

4.3 Calibration after refining the parameter space

Both algorithms TS and CMA-ES are executed in the optimal space derived from the first stage of the procedure. The resulting parameter sets are denoted by OptTS* and OptCMAES*. Table 3-8 presents the performances of these sets of parameters obtained for each basin. Both algorithms provide realistic solutions, simulating the catchment behaviour adequately. The optimal space is refined and excludes unrealistic values and the ranges of values are decreased, depending on the basin characteristics. Thereby, the risk of converging to an unrealistic combination of parameter values is eliminated. Both the set of parameters OptTS* and

OptCMAES* have NSE values superior to the Mean NSE provided in the first step except for Catamaran Brook.

Table 3-8 Performances of the set of parameters resulted of the calibration in the optimal space.

Basin		Calibration			Evaluation		
		NSE	Bias	KGE	NSE	Bias	KGE
CD	OptTS*	0.90	0.05	0.92	0.90	0.02	0.94
	OptCMAES*	0.89	0.01	0.94	0.86	0.05	0.90
	OptTS*	0.82	0.04	0.90	0.76	0.02	0.88
PD	OptCMAES*	0.83	0.04	0.90	0.81	0.03	0.90
	OptTS*	0.62	0.09	0.79	0.59	0.23	0.70
Nech	OptCMAES*	0.64	0.09	0.80	0.60	0.22	0.70
	OptTS*	0.79	0.05	0.87	0.72	0.09	0.74
CAT	OptCMAES*	0.80	0.04	0.88	0.74	0.10	0.75

The CVNB is intentionally set to 0 and the range of EVNAP is changed to [0 0.3] for all basins. Imposing these conditions leads to a deterioration of the NSE of any solution taken randomly from the optimal space. Because CVNB=0 and the ranges of EVNAP differ from the one defined in the optimal space (see Table 3-4), the NSE value of the model calibrated on Catamaran Brook with TS and CMAES (0.79 and 0.8 resp.) is less than 0.83 (the mean NSE of the optimal space). Without these conditions, calibration leads to the unrealistic situation where the water level simulated in the lower reservoir is almost always empty, except during the spring flood. In order to maintain a minimum water level in the lower (groundwater) reservoir, CVNB (lower outlet) is fixed at 0. Hence, levels in the lower reservoir vary only through the upper drainage outlet as determined by the CVNH and HNAP coefficients, or because of evaporation (see Figure 3-1). The upper bound of the EVNAP is fixed at 0.3 and the lower bound to 0, so that the algorithm would not compensate the closed CVNB by increasing EVNAP and empty completely the lower reservoir through evaporation, which is deemed unrealistic. Calibration results on the Nechako River are modest with NSE coefficient values around 0.6 for calibration and evaluation periods. The KGE is around 0.8 in calibration and is 0.7 in evaluation period. This means that the set of parameters obtained catches well the variability of the flows. Figure 3-9 shows a good fit between simulations and observations except for low flows in the Chute-Du-Diable, Passes-Dangereuses and Nechako River during calibration period. The low flows are underestimated in

these basins in the evaluation period too (see Figure 3-10). The model calibrated for Catamaran Brook simulates well both low and high flows during the calibration period, but underestimates high flows during the evaluation period.

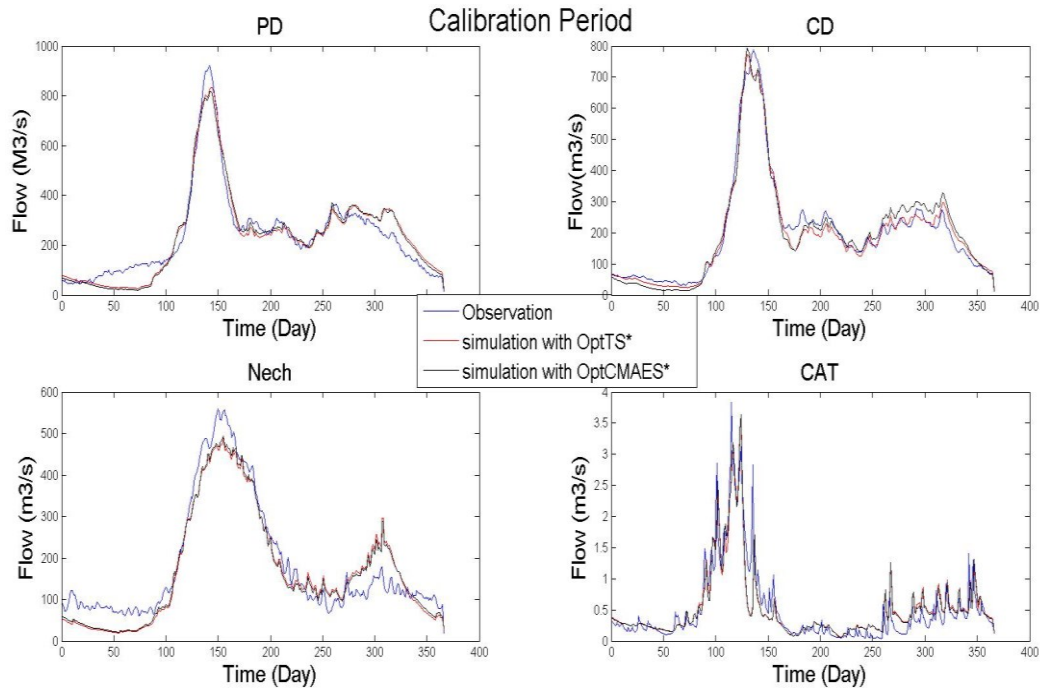


Figure 3-9 Interannual mean flow simulated and observed during the calibration period in all basins.

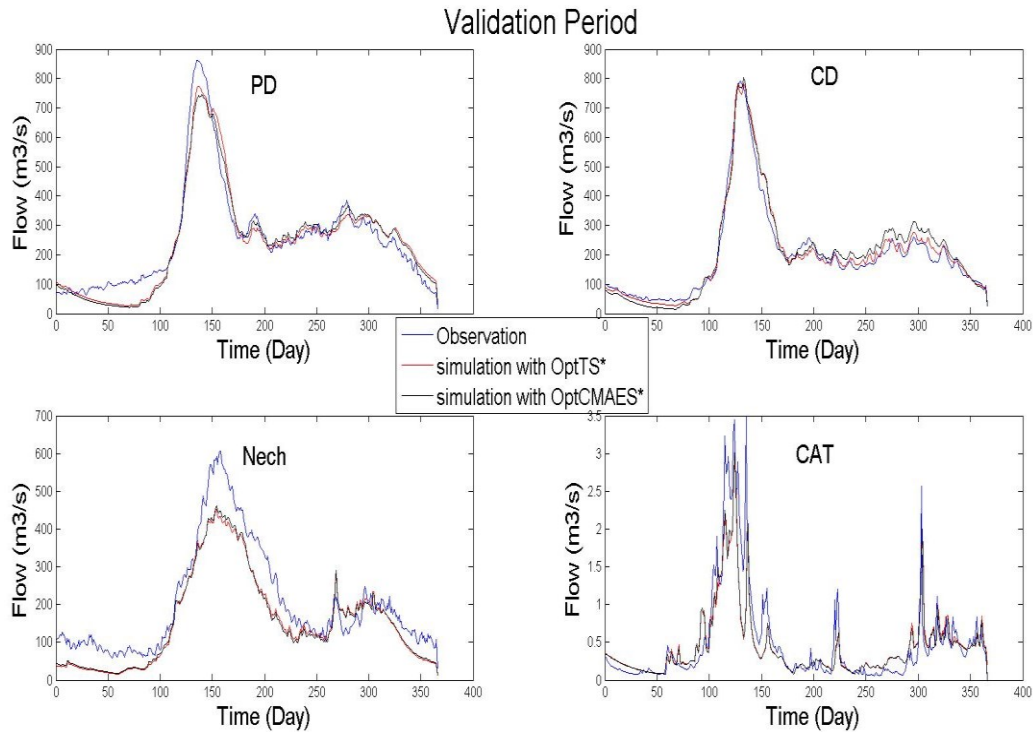


Figure 3-10 Interannual mean flow simulated and observed during validation period in all basins.

The seasonal NSE and bias (Table 3-9) confirm the poor estimation of the low flows during the winter with negative values of the NSE except the Catamaran basin and good estimation of flows during the rest of the year. The model structure obtained for Catamaran estimates relatively well the lows flows with a relative seasonal bias equal to 0.07 for OptTS* and 0.12 for OptCMAES*. Overall, the model catches the entire variability of the observed flows reflected by KGE which is over 0.87 by OptTS* and OptCMAES*.

Table 3-9 Seasonal NSE and bias of the set of parameters resulted of the calibration in the optimal space.

		Calibration							
		Seasonal NSE				Seasonal Bias			
Basin		JFM	AMJ	JOS	OND	JFM	AMJ	JOS	OND
	OptTS*	-2.90	0.82	0.73	0.87	0.13	0.23	0.10	0.04
CD	OptCMAES*	-0.96	0.89	0.71	0.80	0.42	0.07	0.02	0.15
	OptTS*	-1.38	0.73	0.71	0.74	0.36	0.08	0.09	0.15
PD	OptCMAES*	-1.30	0.81	0.73	0.76	0.54	0.05	0.05	0.11
	OptTS*	-1.87	0.56	0.18	-1.14	0.54	0.14	0.10	0.32
Nech	OptCMAES*	-1.26	0.68	0.27	-0.93	0.58	0.09	0.08	0.26
	OptTS*	0.18	0.45	0.33	0.61	0.07	0.11	0.25	0.09
CAT	OptCMAES*	0.01	0.70	0.37	0.81	0.12	0.09	0.75	0.02
		Evaluation							
	OptTS*	-0.16	0.95	0.69	0.31	0.25	0.03	0.11	0.20
CD	OptCMAES*	-0.46	0.80	0.70	0.68	0.39	0.01	0.13	0.16
	OptTS*	-1.43	0.68	0.78	0.73	0.39	0.00	0.01	0.14
PD	OptCMAES*	-2.12	0.75	0.76	0.76	0.40	0.06	0.04	0.15
	OptTS*	-1.07	0.62	0.34	-0.10	0.67	0.19	0.19	0.13
Nech	OptCMAES*	-1.23	0.61	0.25	-0.04	0.56	0.19	0.24	0.12
	OptTS*	0.44	0.62	0.61	0.78	0.37	0.20	0.08	0.02
CAT	OptCMAES*	0.47	0.64	0.63	0.79	0.36	0.19	0.10	0.02

4.4 Calibration with differential split sample

In this section, the results of the differential split sample are not compared to a calibration without refining the parameter space and only TS is used because it provides sets of parameters with equivalent performances as CMA-ES in the optimal space. This test was not used to calibrate the model for Catamaran Brook because of the short length of the available data. The algorithm with Depth function converges to a less performant space compared to the space obtained with split sample for all catchments except for Nechako. For this basin, when calibrating on wet period, the algorithm converges to a space with mean NSE around 0.5 versus 0.3 (with the split sample). Table 3-10 exposes the performance of the sets of parameters obtained with TS in calibration and evaluation periods.

Table 3-10 Performance of the set of parameter resulted of the calibration with the differential split sample.

Basin	Calibration on Wet Period			Evaluation on Dry Period		
	NSE	Bias	KGE	NSE	Bias	KGE
CD	0.92	0.03	0.94	0.88	0.09	0.82
PD	0.85	0.01	0.92	0.87	0.03	0.91
NECH	0.60	0.09	0.78	0.44	0.21	0.63
Basin	Calibration on Dry Period			Evaluation on Wet Period		
	NSE	Bias	KGE	NSE	Bias	KGE
CD	0.93	0.04	0.93	0.87	0.08	0.88
PD	0.75	0.04	0.72	0.74	0.03	0.72
NECH	0.50	0.25	0.65	0.52	0.21	0.70

The performance of the calibrated model on Chute-Du-Diable during wet and dry periods are similar. The performance of this calibrated model drops in both cases to NSE over 0.87 during the evaluation period versus NSE over 0.92 during calibration. The same conclusion can be made when the KGE criterion is used. The calibrated model on Passes-Dangeureuses during a dry period is more performant (NSE=0.85 and KGE=0.92) than the one calibrated during a wet period (NSE=0.75 and KGE=0.72). In both cases the calibrated model maintains its performance during evaluation periods. In the Nechako River, the calibrated model during a dry period is performant only for the calibration phase. The NSE drops from 0.6 to 0.44 when the evaluation period is wet. On the contrary, the calibrated model on wet period has lower but consistent performance (NSE=0.5 during calibration, 0.52 for evaluation). The KGE has higher values (over 0.63 in both tests during calibration and evaluation), which reflects a good simulation of the variability of flows.

5 Discussion

On the optimization of the parameters space

The algorithm defines an optimal space that allows better constraining conditions to estimate realistic parameters for subsequent use of a global optimization method. The convergence time of the algorithm depend on the studied basin, the number of parameters and the model calibrated. When adjusting few CEQUEAU parameters (7 to 15) on the Chute-Du-Diable drainage basin, the algorithm converges after 9 to 12 iterations. In the case study illustrated by Bardossy and

Singh (2008), the algorithm used converged in only 4 iterations to define the geometric structure with the Halfspace depth of 9 parameters of the HBV model. Using other depth functions may give different results of the number of the selected deepest points. Hence, it affects the size of the resulted space.

For example Figure 3-11 presents the set Y^* , of the deepest parameters within the set X^* , selected from the same set Y by the Mahalanobis depth function (used in this study), and the Halfspace depth. The parameters laying around the boundaries of X^* are rejected by the Halfspace depth as they have depth equal to 0. In fact, the Halfspace of a point p with respect to a finite set M in a multivariate space is defined as the minimum number, calculated over all possible hyperplanes, of points of the set M laying in one side of a hyperplane through the point p (Bardossy and Singh, 2008). The structure of the depth function explains the rejection of parameters in the boundaries of the set X^* .

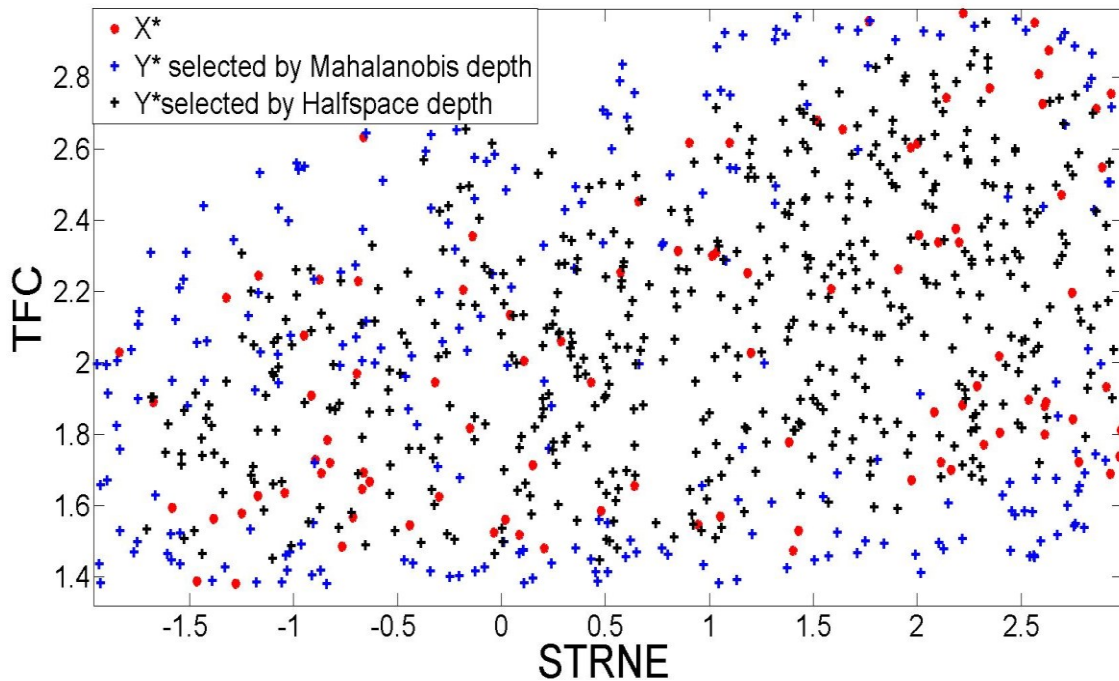


Figure 3-11 Deepest points Y^* within the set X^* selected by Mahalanobis and Halfspace depth functions.

In this study, the purpose of using a depth function is to allow a strategic sampling in the Monte Carlo simulation to select points that lay around any performant parameter. The Mahalanobis depth was an appropriate choice as it does not neglect the points in the boundaries and it is easily calculated for any dimension. Some depth functions may not be adapted for high dimensions and increase the complexity of the optimization problem in particular if the model is complex. CEQUEAU is considered as a complex model compared to lumped models. So the choice of the depth function must consider the dimension of the space and the complexity of the model.

On the performance of algorithms TS and CMA-ES

In the optimal space both algorithms converged to realistic sets of parameters with equivalent performances. None of these algorithms is capable of providing a solution that is equally performant for both high and low flows. Even though OptCMAES* had better seasonal NSE values for some periods, the seasonal relative bias calculated for the same period is higher or the same as the bias of OptTS*. For example, in Chute-Du-Diable basin, the NSE coefficient calculated for the winter is -2.90 for OptTS*, inferior to the NSE calculated for OptCMAES*, which is still negative (-0.96). But the bias of OptCMAES* is 0.42 against 0.13 for OptTS*. In the Nechako River, Seasonal NSE of OptCMAES* for all periods are better than the values obtained for OptTS*. In the other hand, the biases are relatively the same for both set of parameters in calibration and evaluation period. In the initial space, TS converged to better and robust results. For example OptTS in the Nechako River have NSE equal to 0.71 in calibration and 0.69 in evaluation. In the same catchment, OptCMAES has a NSE equal to 0.69 in calibration and 0.58 in evaluation. In fact, TS was started with relatively good solutions and improved them in less than 100 iterations. The TS strategy is varying parameters of the same solution in order to improve the objective function. Consequently, the TS converges after a few iterations to a better solution when it is started with a performing solution. If it cannot find a better solution it will propose the initial solution as an optimum. When the bounds of the space are not well defined, both algorithms tend to converge to unrealistic solution, as shown in the results of calibration without refining the space. Indeed, the model tries to compensate parameters in order to match the observations without considering the realism of the

combination. This is a major constraint when calibrating a conceptual model, such as CEQUEAU.

As the space is already reduced and contains only performing parameters, using TS is more adequate. This algorithm takes into consideration the sensitivity of parameters in defining its trajectory. In the phase of exploration, TS examines the space of parameters and save in the frequency list the parameters that improved the objective function. This frequency list will mostly contain the sensitive parameters and the one in interaction with them. So in the phase of intensification, it adjusts only these parameters trying to find a compromise between the interacted parameters. In the diversification stage, the insensitive parameters will be examined.

On the performance of the proposed approach

The proposed procedure allows finding realistic performing set of parameters over calibration and evaluation period with split sample and differential split sample. The first stage, refining the parameters space, provides good starting solutions for global optimization algorithms. The spring flood was well simulated in Chute-Du-Diable and Passes-Dangereuses basins with high values of seasonal NSE in calibration and evaluation period and low relative seasonal bias (around 0). The global NSE and bias were affected by the adequacy of the simulated spring flood which, in these two basins, is very important (flow values exceed $800 \text{ m}^3/\text{s}$). The global NSE was not affected by the poor estimation of low flows in Chute-Du-Diable and Passes-Dangereuses basins, unlike the Nechako River. In Catamaran Brook, low flows were relatively well simulated by both algorithms. The seasonal bias calculated for the winter are 0.07 for OptTS* and 0.12 for OptCMAES*. In fact, in this basin, high flows are not as important as in the other basins as the peak reach values around $3.5 \text{ m}^3/\text{s}$ against values over $600 \text{ m}^3/\text{s}$ in the other basin (see Figure 3-9). Discrepancies between simulated and observed peak flows in Catamaran Brook are not penalizing as much as for the largest basins. The underestimation of low flows in the other catchments can be explained by the use of a single objective function (NSE) to calibrate the model. In fact, a single objective function is unable to catch the entire characteristics of observations and tends to represent one aspect (low flows or high flows) at the expense of the other. As pointed out in several studies, the NSE tends to weigh less importance to differences in simulated and observed low flows and is incapable to catch the variability of the flows (e.g.,

Krause et al., 2005; Gupta et al., 2009). Indeed, using other performance criteria may give different results as each criterion is often more sensitive to one specific aspect of the hydrograph (low flows, high flows, variability, etc.) than another. With split sample structure, the procedure was able to provide realistic performing set of parameters in all catchments regardless of the characteristics of the catchment and hydro-meteorological conditions. Generally, models fail the differential split sample test, but it is an instructive test to illustrate the risks related to the model and its failures (Seibert, 2003).

6 Summary and Conclusion

This study proposes a new approach to calibrate hydrological models based on a two-step procedure. The first step uses a Monte Carlo simulation coupled with the concept of depth function to reduce the parameter space, which gives better constrain conditions to a global optimization algorithm. Indeed, it is shown that when the ranges of parameters are not well defined, global optimization algorithms converge to some non-realistic parameter values. The size of the resulting space depends on the depth function that must be chosen considering the dimension of the space and the complexity of the model. The second stage consists in using a global optimization algorithm to explore the redefined parameter space. In this study, it is shown that the Tabu Search is more suited to use with the proposed approach as it converges in few iterations when initiated with a good solution and takes into consideration the sensitivity of parameters in order to improve the objective function. It provides the same (if not better) performance results as an evolutionary algorithm like CMA-ES.

The proposed procedure was applied to calibrate the CEQUEAU model on four different catchments using a split sample and differential split sample. The model calibrated was robust i.e. the performance is relatively similar in calibration and evaluation period in all basins and reflects their hydrological behaviour. However, the low flows are poorly estimated in the basins where the flood spring is very important. This is mainly because a single objective function (NSE) is used in calibration. Other factors can contribute to the low model performance, such as the structure of the model and the uncertainty of meteorological input data or of the response system measurements (i.e. flows). Further research is needed to estimate parameter sets that are capable of simulating the entire variability of hydrographs. Additional opportunity for obtaining

improved results may come from applying more advanced forms of Tabu search as described in Glover and Laguna (1997).

Acknowledgments

Funding from NSERC and Rio Tinto Alcan for this project is acknowledged. Authors are also thankful for the technical assistance provided by the hydrological forecast team at Rio Tinto Alcan.

References

- Arsenault, R., Poulin, A., Côté, P. and Brissette, F. 2013. A comparison of stochastic optimization algorithms in hydrological model calibration. *J. Hydrol. Eng.*, 1943-5584.
- Bardossy, A and S.K Singh. 2008. Robust estimation of hydrological model parameters, *Hydro. Earth Syst. Sci.*, (12), 1273–1283.
- Charbonneau, R., Fortin, J-P. and Morin, G. 1977. The CEQUAU model: Description and examples of its use in problems related to water resource management/ Le modèle CEQUEAU: description et exemples d'utilisation dans le cadre de problèmes reliés à l'aménagement. *Hydrological Sciences bulletin*, 22(1), 193-202.
- Duan, Q., Sorooshian, S. and Gupta V. 1992. Effective and Efficient Global Optimization for conceptual Rainfall-Runoff Models. *Water Resour. Res.*, 28 (4),1015-1031.
- Glover, F. 1990. Tabu Search a tutorial. *Interfaces* 20, 74-94.
- Glover, F. and Laguna, M. 1997. Tabu Search. *Kluwer academic. Springer*.
- Glover, F., Taillard ,E. and De Werra D.1993. A user's guide to tabu search, *Annals of Operations Research*, 41, 3-28.
- Griensven, A.V, Meixner, T., Grinwald, S., Bishop, T., Diluzio, M. and Srinivasan, R. 2006. A global sensitivity analysis tool for parameters of multi-variable catchment models. *J. Hydrol.*, 324 10-23.
- Gupta, H., Sorooshian, S. and Yapo, P. O. 1999. Status of Automatic Calibration for Hydorlogic Models: Comparison with Multilevel Expert Calibration. *J. Hydrol. Eng* , 4(2), 135-143.
- Gupta, H., Kling, H., Yilmaz, K. and Martinez, G. F. 2009. Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. *J. Hydrol.*, 80-91.
- Hansen, N. and Ostermeier A. 1996. Adapting Arbitrary Normal Mutation Distributions in Evolution Strategies : The Covariance Matrix Adaptation. In: Proceedings of the 1996 IEEE Intern. Conf. on Evolutionary Computation (ICEC '96), 312-317.

- Hansen, N. and Ostermeier A. 1997. Convergence Properties of Evolution strategies with the Derandomized covariance Matrix Adaptation: The $(\mu/\mu_1, \lambda)$ -CMA-ES. EUFIIT'97, 5th Europ. Cong. On Intelligent Techniques and Soft Computing, Proceedings: 650-654.
- Hansen, N. and Ostermeier A. 2001. Completely Derandomized Self-adaptation in Evolution strategies. By the Massachusetts institute of Technology Evolutionary Computation, 9(2), 159-195.
- Holland, J. H. 1975. Adaptation in Natural and Artificial Systems. University Michigan Press, Ann Arbor, Michigan, USA.
- Kirchner, J. W. 2006. Getting the right the answers for the right reasons: Linking measurments, analyses, and models to advance the science of hydrology. *Water Resour. Res.*, 42, W03S04.
- Klemeš, V. 1986. Operational testing of hydrological simulation models. *Hydrol. Sci. J.*, 31(1), 13-24.
- Krause, P., Boyle, D. and Bāse, F. 2005. Comparaison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 89-97.
- Liu, RY. 1990. On a notion of data depth based on random simplies. *Annals of statistics*. 18(1), 405-414.
- Macdonald, J.S., Morrison, J. and Patterson, D.A. 2011. The efficacy of Reservoir Flow Regulation for Cooling Migration Temperature for Sockeye Salmon in the Nechako Watershed.
- Mahalanobis, P.C. 1936. On the generalized distance in statistics. *Proc. Natl. Acad. Sci., India*, 12, 49-55.
- Moradkhani, H. and Sorooshian, S., 2009. General review of Rainfall-Runoff Modeling: Model calibration, Data Assimilation, and Uncertainty Analysis. In *Hydrological Modelling and the Water Cycle: Coupling the Atmospherique and hydrlogical Models*, Sorooshian, S., Hsu, K.-L., Coppola, E., Tomasseti, B., Verdecchia, M., Visconti, G. (Eds.) Springer, 1-24.
- Morin, G. and Paquet, P. 1995. Le modèle de simulation de quantité et de qualité CEQUEAU, Guide de l'utilisateur. Version 2.0 pour Windows. INRS-Eau, rapport de recherche no 435, 54 pp.
- Morin, G. and Paquet, P. 2007. Modèle hydrologique CEQUEAU. *INRS-ETE, Rapport de recherche no R000926*, 458p.
- Oja, H. 1983. Descriptive Statistics for multivariate distributions. *Statistics & Probability Letters*, 1, 327-332.
- Pechlivanidis, I., Jackson, B., McIntyre, N. and Wheeler, H. 2011. "Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Global NEST Journal*, 13(3), 193-214.
- Seibert, J. 2000. Multi-criteria calibration of conceptual runoff model using a genetic algorithm. *Hydrol. Earth Sys. Sci.*, 4 (2), 215-224.
- Seibert, J. 2003. Reliability of Model Predictions Outside Calibration conditions. *Nordic Hydrology*, 34 (5), 477-492.

- St-Hilaire, A., Morin, G., El-Jabi, N. and Caissie, D. 2000. Water temperature modelling in a small forested stream: implication of forest canopy and soil temperature. *Can. J. Civ. Eng.*, 27, 1095-1108.
- St-Hilaire, A., El-Jabi, N., Caissie, D., Morin, G. 2003. Sensitivity analysis of a deterministic water temperature model forest canopy and soil temperature in Catamaran Brook (New Brunswick, Canada). *Hydrol. Process.*, 17, 2033-2047.
- Tolson, B.A. and Shoemaker, C.A. 2007. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resour. Res.*, 43, W01413.
- Tukey J.W. 1975. Mathematics and picturing data. In Proceedings of the International Congress on Mathematics, Canadian Mathematics Congress.
- Tung, C.P. and Chou, C.A. 2004. Pattern classification using tabu search to identify the spatial distribution of groundwater pumping. *Hydrogeology Journal*, 12, 488-496.
- Vrugt, J.A and Robinson, B.A. 2007. Improved evolutionary optimization from genetically adaptive multimethod search. *Proc. Natl. Acad. Sci. U.S.A.*, 104, 708-711.
- Yang, Y., Wu, J. Sun. X, Wu, J. and Zheng. C. 2013. A niched Pareto tabu search for multi-objective optimal design of groundwater remediation systems. *J. Hydrol.*, 490, 56-73.
- Zhang, X., Srinivasan, R., Zhao, K., and Liew, M.V. 2009. Evaluation of global optimization algorithms for parameter calibration of a computationally intensive hydrologic model. *Hydrol. Process*, DOI: 10.1002/hyp.7152.
- Zheng, C. and Wang, P. 1996. Parameter structure identification using Tabu search and simulated annealing. *Advances in Water Resources*, 19(4), 215-224.
- Zuo, Y and Serfling, R. 2000. General notions of statistical depth functions. *The annals of statistics*, 28(2), 461-482.

CHAPITRE 4 ANALYSE DE DONNÉES

FONCTIONNELLES POUR CALIBRER ET

ÉVALUER UN MODÈLE HYDROLOGIQUE

Les travaux réalisés dans ce chapitre sont présentés sous forme d'article scientifique rédigé en anglais et dont le titre est :

Using functional Data Analysis to calibrate and evaluate hydrological model performance

Samah Larabi ^{*1}, André St-Hilaire ¹, Fateh Chebana ¹ and Marco Latraverse ²

¹INRS-ETE, Québec, Québec, Canada

²Rio Tinto Alcan, Saguenay, Québec, Canada

Cet article est soumis à Journal of Hydrologic Engineering. Cette version correspond à la version de l'article tel que soumis au journal.

Les objectifs de cet article ont été discutés entre les différents auteurs. Le code du modèle hydrologique a été fourni par M. Latraverse. La méthodologie a été proposée et codée par l'auteur principal. L'analyse des résultats et la rédaction ont été effectuées par l'auteur principal en collaboration avec le professeur St-Hilaire. Une révision du contenu de l'article a été effectuée par le professeur Chebana.

Abstract

The performance of a hydrological model depends strongly on the calibration procedure, in particular, the goodness-of-fit measure used. It is widely recognized that traditional goodness-of-fit measures such as the Nash-Sutcliffe Efficiency (NSE) are biased toward securing a particular aspect of a hydrograph (high flows in the case of NSE). This paper proposes a new strategy for the model calibration that evaluates the ability of the model to simulate the complete shape, timing, and variability of the observed hydrographs. The methodology is based on the comparison of the simulated and observed the whole annual hydrograph as a single curve using the functional data analysis (FDA) framework. FDA is a recent statistical framework that considers observations as curves or functions. The hydrograph is a particular example of such functions. This study uses three Canadian watersheds with different hydrologic regimes to compare the performance of the proposed method versus a traditional calibration with NSE. The results show that the proposed procedure provides similar results for high flows on watersheds with large spring snowmelt and improves the consistency of the model.

Keywords: Calibration, CEQUEAU, Functional Data Analysis, Objective function.

1 Introduction

The performance of a hydrological model depends on its calibration quality, the amount and quality of input data, the procedure and the objective function used during the calibration process. The latter is a performance/goodness-of-fit measure reflecting the ability of a model to reproduce the observed system response. The majority of the performance measures aggregate or average the squared residuals (i.e. differences between the simulated and observed outputs) such as the Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970), the Root Mean Square Error (RMSE) and the index of agreement. Although the NSE is widely used and more intuitive to hydrologists (Pechlivanidis et al., 2011), this measure is not a reliable basis to evaluate a model in absolute terms (Schaeffli and Gupta, 2007). Moreover, when calibration is based on these performance measures it will typically lead to a better fit for high flows at the expense of low flows (Krause et al., 2005).

Generally, any single objective function used alone is inadequate to reflect properly the entire characteristics of the observed time series (Vrugt et al., 2003; Pechlivanidis et al., 2011). On the one hand, the different objective functions emphasize different aspects of the hydrograph and yield to different optimal parameters sets; consequently one must sacrifice the fit to one aspect of the hydrograph to fit another (Wagner et al., 2004). Globally, the tradeoff is between the fit to high flows and to low flows. On the other hand, it is very difficult to reduce all the information about the characteristics of the flow regime to a single piece of information reflected by the objective function (Gupta et al., 2008). In order to counter this problem and include different flow characteristics to find the optimal calibration of a model, the multi-objective approach was proposed by Gupta et al. (1998) and is widely used (Vrugt et al., 2003; Fencia et al., 2007; Yilmaz et al., 2008; Asadzadeh et al., 2014; Pechlivanidis et al., 2014). It consists of finding a trade-off between different criteria. In this regard, two main approaches are available in the literature. A first method is an aggregation of different objective functions expressed as a single one (e.g. the Kling-Gupta Efficiency (KGE), Gupta et al.(2009)). In this case, the results are strongly dependent on the aggregation. A second approach is based on the Pareto concept which consists of identifying a set of non-dominated solutions through optimization of different objective functions simultaneously (Yapo et al., 1998; Vrugt and Robinson, 2007; Pechlivanidis et al., 2011; Asadzadeh et al., 2014). However, to our knowledge, there is no method that fully

defines the optimal number or the type of objective functions to use in the multi-objective approach (Krause et al., 2005).

For real world applications, in particular the flood forecasting, the aim of using a model is often to simulate the entire hydrograph and the ability of the model to reproduce the amplitude and timing of observed peak flows is critical. When using the conventional statistics as objective functions (NSE, RMSE, R^2 , mean absolute error) in a single or multi-objective framework, the capacity to reproduce the hydrograph shape and temporal streamflow variability is often limited. There is therefore a need for statistics and performance measures that conserve and describe the entire shape of hydrographs. Functional data analysis (FDA, (Ramsay and Silverman, 2002,2005)) is a recent statistical framework dealing with observations within a space or temporal framework as curves or functions. In the present study, the hydrograph is a particular example of such functions. Using FDA comes with a number of advantages. The primary one is that the FDA allows detecting the timing of events in the time series, in addition to extracting additional information in the functions and their derivatives (Levitin et al., 2007). The FDA is used in different fields including economy (Ramsay and Ramsey, 2002), environmental sciences (Sguera et al., 2016) and meteorology (Ramsay and Silverman, 2002). It was first introduced in hydrology by Chebana et al. (2012) in a context of hydrographs exploratory analysis and outlier detection. It was subsequently used for streamflow hydrograph classification by Ternynck et al. (2015). Both studies showed that FDA has more advantages than a multivariate approach as it allows taking into account the entire shape and variability of a hydrograph without the subjectivity of choosing the characteristics of a basin to be studied.

The aim of this paper is to propose a calibration strategy that overcomes the limitations of both calibrations with a single or multi-objective function. The proposed approach uses FDA in order to compare the simulated and observed hydrographs as temporal functions instead of discrete data (daily flows). By doing so, the model is evaluated on its ability to reproduce the same shape, variability and timing of the observed hydrograph in one single step. The method is applied to calibrate the CEQUEAU model in three Canadian watersheds and compared to a conventional calibration with the NSE. The intent of this comparison is to test the potential of the proposed calibration procedure to improve simulated hydrograph components other than the annual peak flows (e.g. low flows), while providing the same if not better results for high flows, as obtained when using conventional calibration approaches such as the NSE. The model is evaluated using

the NSE, and several hydrological signatures as well as functional statistics. Hydrological signatures, which are indices used to describe the dynamic behaviour of a watershed (Hartmann et al., 2013; Gunkel et al., 2015), assess the degree of model consistency to reproduce the entire hydrological regime. Whereas, the purpose of using functional statistics is to detect when observations and model simulations are significantly different and do not overlap in a more formal mathematical setting than by mere observations of simulated and observed time series. Consequently, it allows testing the hypothesis motivating the present study: does a calibration considering an entire hydrograph as a single observation improve the ability of a model to simulate the complete hydrological behaviour of the basin?

The present paper is organized as follows. Section 2 introduces the proposed methodology, including the elements of FDA used, calibration procedure and the criteria selected for model evaluation. In section 3 a brief description of the model, study areas and the data are presented in addition to the results of the calibration with the proposed approach. A discussion is presented in section 4, followed by a conclusion in section 5.

2 Methodology

This section first presents the initial step of FDA, which consists of converting raw data into continuous functions via basis function expansion as well as the functional statistics used in this study. Then, the calibration procedure and the proposed objective function are presented in subsections 2.2 and 2.3. The evaluation criteria are presented in subsection 2.4.

2.1 Data smoothing and functional statistics

The first step in FDA is to convert the raw data (daily discharge in the present study) into smoothed curves. In other words, each annual daily flow series $\mathbf{Q}_i = \{Q_{i,1}, \dots, Q_{i,365}\}$ is transformed into one continuous temporal function or curve $x_i(t)$ $\{x_i(t), t \in [1, T = 365]\}$ representing a smoothed hydrograph. The curves $x_i(t)$ are constructed generally via a basis expansion (Ramsay and Silverman, 2002):

$$x_i(t) = \sum c_k \phi_k, \text{ where } \phi \text{ are the basis functions and } c \text{ are the coefficients.} \quad (4.1)$$

There is a variety of basis function (constant, polynomial, spline, Fourier series, wavelets...) adapted for each case study. In this context case, because the data are periodic, the natural choice is a Fourier basis (similar to Chebana et al., 2012). The curves are constructed as follows:

$$x_i(t) = c_1 + c_2 \sin \omega t + c_3 \cos \omega t + c_4 \sin 2\omega t + \dots c_K \cos K\omega t, \quad (4.2)$$

where $\omega = 2\pi/T$, T is the period (here $T=365$) and K is the number of the basis functions.

The curves are fitted to the data by estimating the basis coefficients and fixing a number of bases K . The latter impact how optimally the curves will fit the data and the degree of smoothing. A smaller K leads to smoother curves (low variance) at the expense of the capacity to capture sharper features in the time series; while a greater K leads to better fit to data (lower bias) (Levitin et al., 2007). The trade-off is catching most of the variability in the observed data by using a relatively rich basis functions (i.e. fixing a relatively large K value), while leaving out some of the spurious noise. To do so, a penalty least squared errors (Ramsay and Silverman, 2002) can be used to estimate the coefficients:

$$\text{PENSSE} = \sum_{j=1}^T (Q_{ij} - x_i(t_j))^2 + \lambda J[x_i], \quad (4.3)$$

where Q_{ij} is the flow for the time step j year i , x_i the curve to be estimated for year i , λ is the smoothing parameter, $J[x_i] = \int \left(\left(\frac{2\pi}{T} \right)^2 \frac{dx(u)}{du} + \frac{d^3 x(u)}{du^3} \right)^2 du$ is the roughness penalty and $T=365$ is the period. The smoothing parameter λ can be estimated by an ordinary or generalized cross validation. The latter is favoured as it is not time-consuming and tends to smooth more. The greater λ is, the smoother the curves are. Figure 4-1 presents an illustration of converted daily discharges to curves/ smoothed hydrographs to be analyzed with FDA. The dimension of data is decreased from $n= 365 \times$ number of years to $n=$ number of years where each annual hydrograph is a single observation. The model is calibrated on the new simulated $x_{sim}(t)$ and observed $x_{obs}(t)$ variables. Consequently, each simulated annual hydrograph during the calibration process is also converted into a curve $x_{sim}(t)$ and compared to the corresponding observed one $x_{obs}(t)$. The number of the basis function K and the smoothing parameter λ are set according to the observed streamflow for each catchment. The same set (K, λ) is used to convert the simulated streamflow while calibrating the model on the specific catchment.

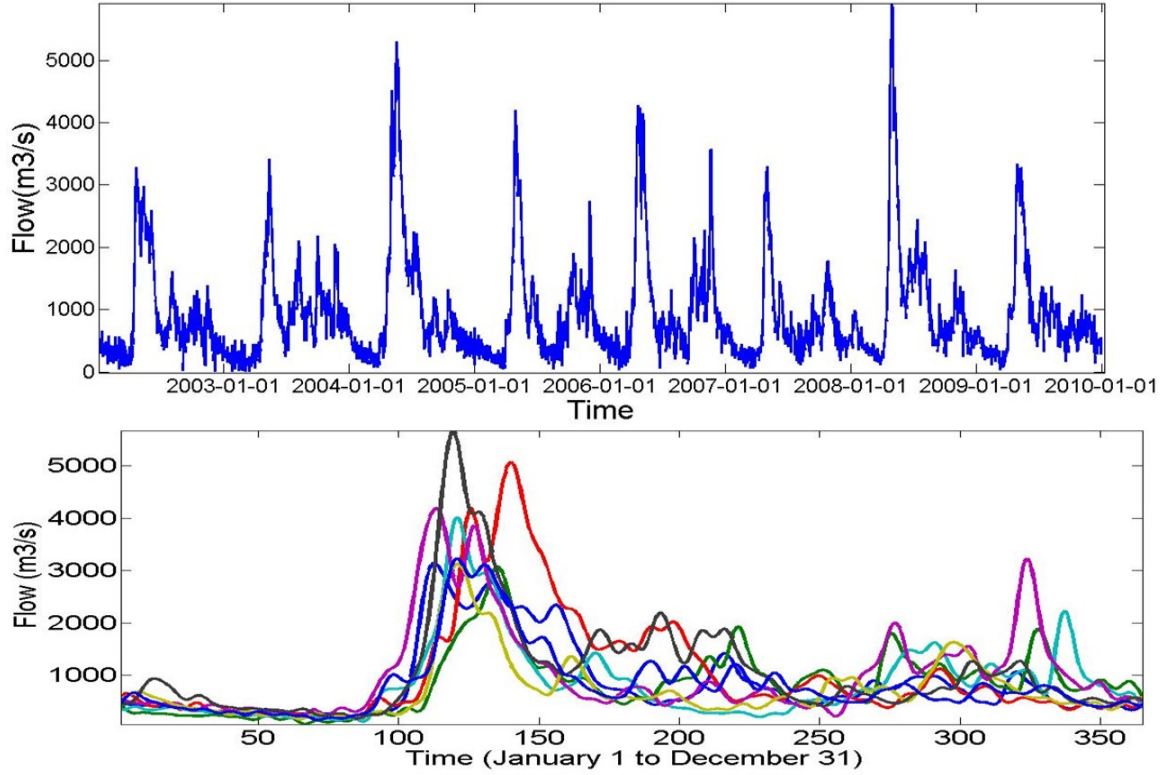


Figure 4-1 Transformation of flow time series into curves representing smoothed hydrographs. Daily observed flows in the Lac St-Jean basin from 2002 to 2009 are plotted in the upper figure and in the lower figure, the associated smoothed hydrographs are plotted, where each hydrograph is a temporal curve.

In this study, the mean function and the functional Student test are used. For a sample of functions $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ defined in the domain D (here $D = [1, 365]$), the mean function is as follows for each time step:

$$\bar{x}(t) = 1/n \sum_{i=1}^n x_i(t), \quad t \in D. \quad (4.4)$$

The functional Student test statistic (Ramsay and Silverman, 2006) for two samples of curves $\mathbf{x}_1$ and $\mathbf{x}_2$ is defined as follows for each time step t :

$$t - statistic(t) = \frac{|\bar{x}_1(t) - \bar{x}_2(t)|}{\sqrt{\frac{1}{T} \{Var[x_1(t)] + Var[x_2(t)]\}}} \quad \forall t \in D \quad (4.5)$$

This test verifies the hypothesis $H_0: \bar{x}_1(t) = \bar{x}_2(t) \forall t \in D$ for a 0.05 significance level. As the statistic is provided for each time step, this test is used to detect when the simulated and observed hydrographs are significantly different. Figure 4-2 illustrates an example of mean functions of two samples with the resulting functional Student test. The days when the black curve (which is the t-statistic) is over the dashed line (point-wise, representing the p-value), the hypothesis that the two samples have the same average is rejected. On the contrary, if the black curve is under the point-wise p-value dashed line, the hypothesis that the two samples are not significantly different is accepted.

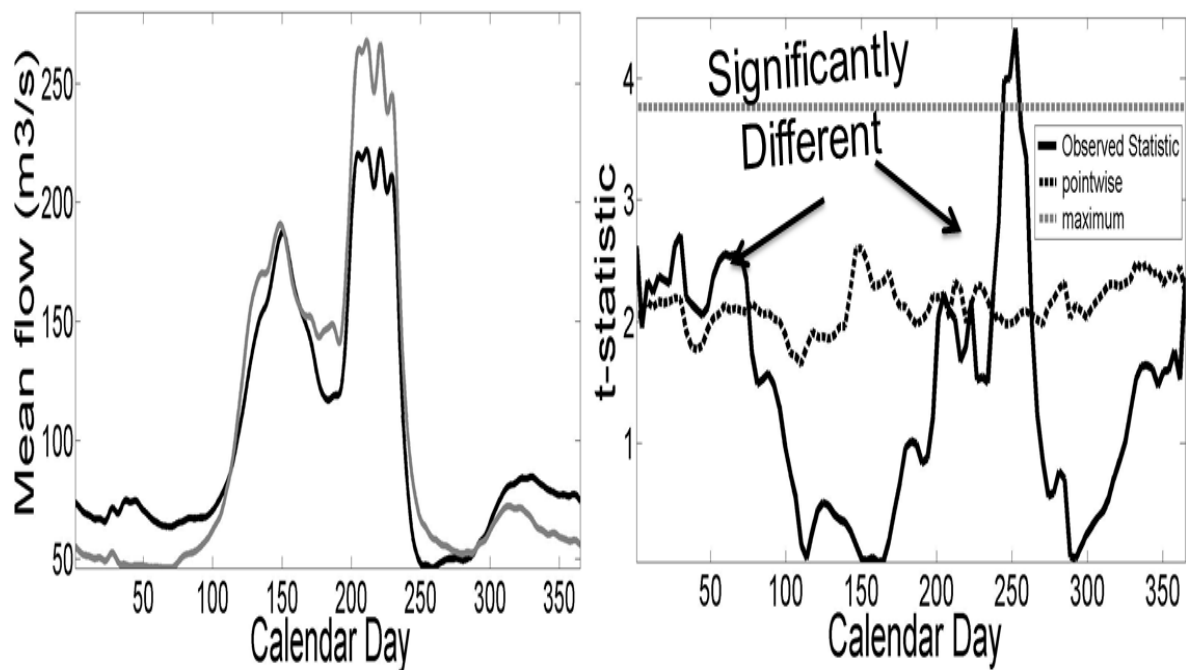


Figure 4-2 Example of mean functions (left Figure) with the associated functional Student Test (right Figure).

2.2 Calibration procedure

The calibration procedure used is a two-step approach where an optimal parameter subspace is first defined, then explored by a global optimization algorithm. Figure 4-3 presents the general flow chart of the calibration procedure. In the subsections below we define the purpose and strategy of each step.

➤ Step 1: Defining the optimal parameters space

The aim of this step is defining the region of attraction of the parameters space. To do so, an adaptation of the ROPE algorithm proposed by Bardossy and Singh (2008) is used. The method is based on a Monte Carlo simulation using a depth function (Tukey, 1975) to restrain the sampling space. The depth function is a statistical measure to quantify the centrality of a point within a multidimensional set. At each iteration, a set of best performing parameters vectors according to the objective function is selected. A new sample is generated from a uniform distribution within the bounds of this set and only the deepest points according to a threshold (e.g. 1st quartile of depths) form the new sample for the next iteration. Therefore, the method samples exclusively new parameter sets located in the region of the parameters space including performing parameters until the overall performance indicator ceases to improve.

➤ Step 2 : Adjusting the parameters in the optimal space

In this step, the optimal space defined in Step 1 is explored deeply using the Tabu Search algorithm. The Tabu search is a global optimization algorithm introduced by Glover (1990) based on the descent search concept. It starts with an initial configuration (in this case a vector of parameters), evaluates its neighboring set of candidate configurations $N(x)$, and replaces it with the best candidate (according to the objective function) until a stop criterion is satisfied. The algorithm uses a First in First out (FIFO) list T called the Tabu list. The purpose of this list is to forbid the search from re-evaluating configurations or revisiting the region of the space previously explored. Thus, it prevents the search to be trapped in cycles of length $|T|$. An intensification technique is integrated to direct the search to explore the promising region deeply of the space. A diversification technique is also used to prompt the search to explore the unvisited regions of the space. Hence, the entire search space is explored. As the space is already

optimized, we define a basic algorithm but more advanced strategies can be incorporated in the Tabu Search (Glover and Laguna, 1997).

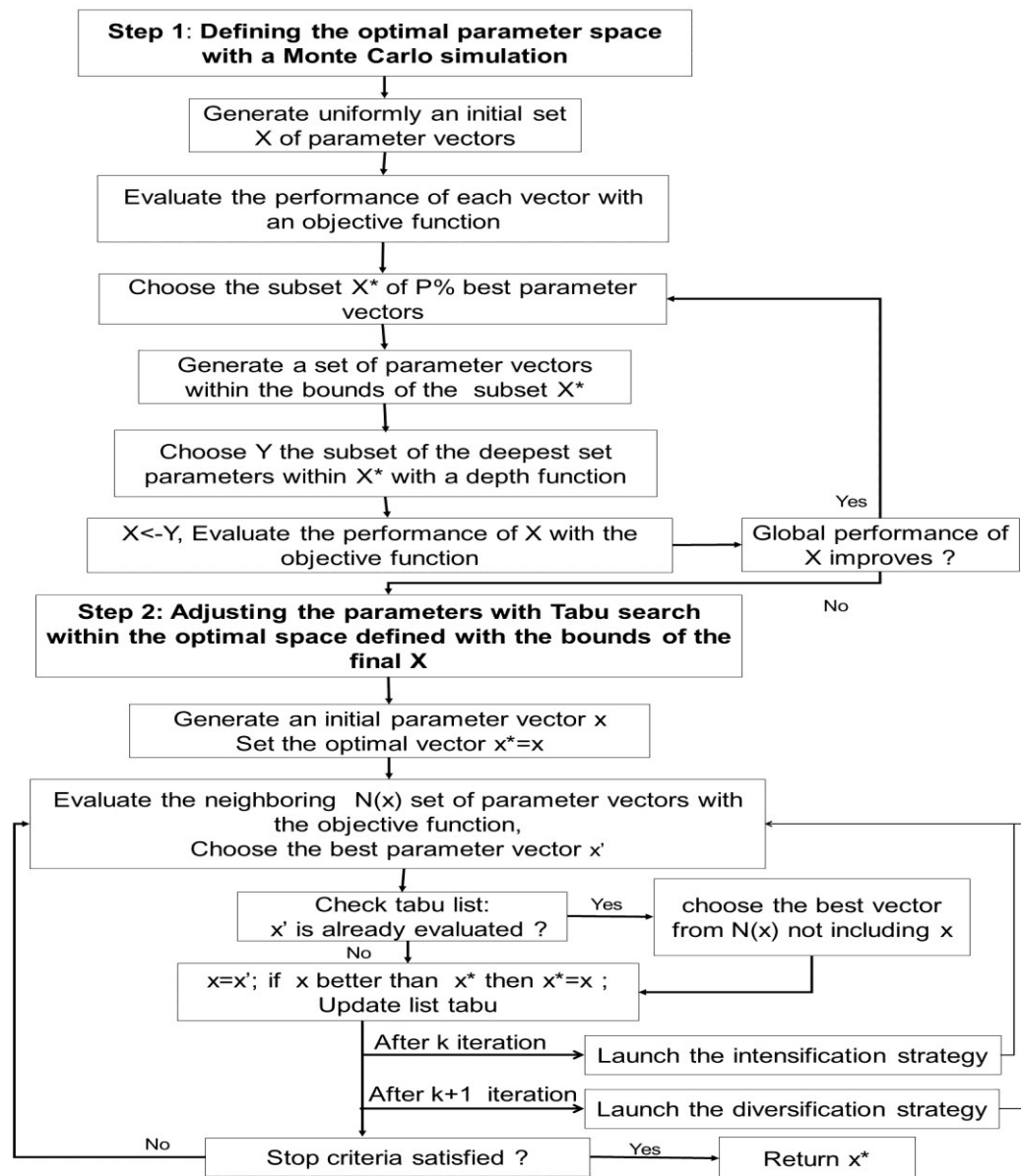


Figure 4-3 Flow chart of the calibration procedure.

2.3 Proposed objective function

The proposed objective function is based on a comparison between the simulated and the observed fitted curves represented as smoothed hydrographs. The proposed approach minimizes the distance between the simulated streamflow $x_{i,sim}$ and the observed $x_{i,obs}$ during each year i . In order to compare the timing of the hydrological events (flood, peak, drought, etc.), the distance between 1st derivatives ($x'_{i,sim}$, $x'_{i,obs}$) is also minimized by the objective function. Through a comparison of slopes the timing of hydrological events (flood, peak flow, and drought) can be extracted as illustrated in Figure 4-4 and as follows: The intersection points of the derivative curves with the x-axis are the days when the maxima or the minima are reached. Periods when the values of the derivative curve are around 0, which means that the curve is almost flat, would typically reflect a spell of relatively constant discharge, often associated to low flows.

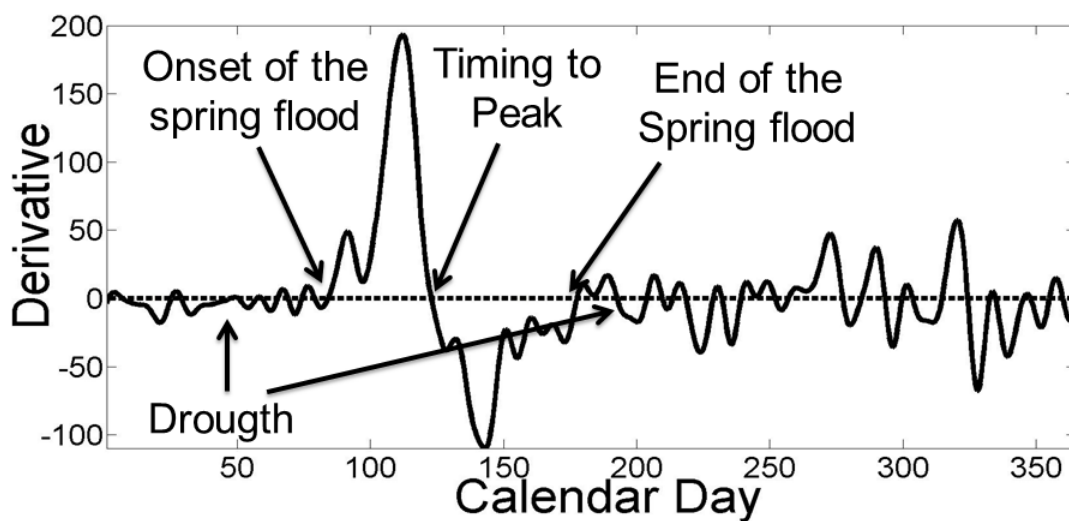


Figure 4-4 Extracting the timing of hydrological events through a change of slope from the first derivative.

The principle of the objective function consists then in minimizing the error in the amplitude and the phase for each year i :

$$\varepsilon_i = \int_{t=1}^T |x_{sim,i} - x_{obs,i}|^2 dt \quad (4.6)$$

where: $x_{sim,i}$ and $x_{obs,i}$ are the simulated and observed curves.

$$\varepsilon'_i = \int_{t=1}^T |x'_{sim,i} - x'_{obs,i}|^2 dt \quad (4.7)$$

where: $x'_{sim,i}$ and $x'_{obs,i}$ are the derivatives of $x_{sim,i}$ and $x_{obs,i}$.

In order, to minimize the effect of the part of the curve representing the flood, the objective function is weighted by the average of the t -statistic. The set of parameters not representing the observed hydrographs adequately on average will have a high average of t -statistics value as illustrated in Figure 4-2. Hence, using the weighting of the objective function with this criterion allows rejecting the set of parameters that is not able to represent the entire shape of the hydrographs (both high and low flows).

The calibration procedure aims first at defining an optimal space of parameters. Thus, an objective function that includes the features described in equations 4.6 and 4.7 is defined as follows in order to constrain a space where the parameters are able to adequately simulate the average hydrograph:

$$Mof = S * \int_{t=1}^T (|\bar{x}_{sim}(t) - \bar{x}_{obs}(t)|^2 + |\bar{x}'_{sim}(t) - \bar{x}'_{obs}(t)|^2) dt \quad (4.8)$$

where : $\bar{x}_{sim}(t)$ and $\bar{x}_{obs}(t)$ are the average curve of the simulated and observed curves, $\bar{x}'_{sim}(t)$, $\bar{x}'_{obs}(t)$ their derivatives, and S is the mean of t -statistic calculated for $x_{i,sim}$ and $x_{i,obs}$, $T = 365$.

At the second step of the calibration procedure, a deep exploration of the resulting parameter space is conducted and the ability of the model to reproduce adequately for each year the shape, volume and timing of the observed hydrograph is evaluated. Hence, the objective function is written as follows:

$$of = S * \sqrt{\sum_{i=1}^N (\varepsilon_i + \varepsilon'_i)} \quad (4.9)$$

where ε_i and ε'_i are as defined in Eq.4.6 and Eq. 4.7, S is the mean of t -statistic calculated for $x_{i,sim}$ and $x_{i,obs}$, and N is the number of years of calibration.

2.4 Evaluation criteria

To provide a complete evaluation of the calibrated model to simulate a hydrological regime, several hydrological signatures are used. The signatures used are descriptive statistics quantifying the distribution flow indices (e.g. mean, median to describe the central values) as well as the Student Test presented in subsection 2.1. To evaluate the frequency and the variability of flows, the following indices are used (Clausen and Biggs, 2000; Westerberg and McMillan, 2015):

- Q_{hF6} (high flow frequency): average number of days per year having flow exceeding 6 times the median flow.
- Q_{hF3} (high flow frequency): average number of days per year having flow exceeding 3 times the median flow.
- Q_{lF} (low flow frequency): average number of days per year having flow not exceeding 0.02 times the median flow.
- Q_{hV} (high flow variability): ratio between the mean annual maximum divided by the median.
- Q_{lV} (low flow variability): ratio between the mean annual minimum divided by the median.
- Q_{gV} (global variability): coefficient of variation which is the standard deviation divided by the mean flow.

In addition to the aforementioned indices and the NSE criterion, the flow duration curve (FDC) is used, which is also a hydrological signature widely used and representing simply the historical distribution of flow variability (Pechlivanidis et al., 2012). Different studies propose a decomposition of the FDC into segments to analyse high and low flows (Yilmaz et al., 2008; Casper et al., 2012; Euser et al., 2013; Shafii and Tolson, 2015). As proposed by Yilmaz et al (2008) and done by Sawicz et al. (2011) we adopt the partition of the FDC into three segments. The used three segments of the curve, reflect the watershed response to extreme events (high flow segment [0-0.02] flow exceedance probabilities), the behaviour of the catchment during

moderate events (mid-segment]0.2-0.7] flow exceedance probabilities), and long-term sustainability of flow during extended dry periods (low flow segment]0.7-1] flow exceedance probabilities).

3. Calibration of the hydrological CEQUEAU model

3.1 CEQUEAU model

The model used in the present study and briefly described in this subsection is the CEQUEAU hydrological model (Charbonneau et al., 1977; St-Hilaire et al., 2000; Morin and Paquet, 2007). CEQUEAU is a deterministic conceptual semi-distributed hydrological model. It takes into account the physical characteristics of the drainage basin and their variation in space via a partition of the watershed into equal areas called whole squares. For each whole square a hydrological budget is calculated by the production function (Figure 4-5). This function simulates the vertical flow considering the ground as two interconnected reservoirs. For each whole square at time t , we have:

$$Q_t = P_t - ETP_t + (HU_t - HU_{t-1}) + HL_t - HL_{t-1} \quad (4.10)$$

where Q (mm) is the total runoff, P (mm) is liquid precipitation or snowmelt, ETP (mm) is the evapotranspiration, HU (mm) is the water accumulated in the upper reservoir, HL (mm) is the water accumulated in the lower reservoir and t is the model time step.

Each whole square is subdivided into a maximum of four so-called partial squares according to sub-basin divides in order to simulate the upstream-downstream flow with a transfer function considering the physical characteristics of the basin. With equation (4.11) the available water volume is routed downstream from a partial square to the other:

$$XKT_i = 1 - \exp\left(-\frac{EXXKT \cdot RMA3_i}{\max(SI, Slac)} * \frac{100}{CEKM2}\right) \quad (4.11)$$

where XKT_i is the routing coefficient (fraction of available water routed downstream) of the i^{th} partial square, $EXXKT$ is the fitting parameter, $RMA3_i$ is the area of the basin upstream of the i^{th} partial square (km^2), SI is the area of surficial water upstream of the partial square, $Slac$ (km^2) is the area of surficial waters on the partial square i , and $CEKM2$ (km^2) is the area of the whole square (constant throughout de drainage basin).

Table 4-1 Model key parameters.

Parameter	Description	Range	
CIN	Percolation coefficient from Upper Zone to Lower Zone	0	1
CVMAR	Lakes and Marshes drainage coefficient	0	1
CVNB	Lower- Zone lower drainage coefficient	0	1
CVNH	Lower- Zone upper drainage coefficient	0	1
CVSB	Upper- Zone lower drainage coefficient	0	1
CVSI	Upper- Zone intermediate drainage coefficient	0	1
XINFMA	Daily Maximum infiltration (mm/day)	0	40
HINF	Infiltration threshold from the upper to the lower zone (mm)	0	100
HINT	Upper- Zone intermediate drainage threshold (mm)	0	100
HMAR	Lakes and Marshes drainage threshold (mm)	200	500
HNAP	Lower- Zone upper threshold (mm)	20	200
HPOT	Threshold of evaporation at the potential rate (mm)	0	80
HSOL	Height of the Upper Zone (mm)	100	300
HRIMP	Quantity of water needed to runoff begins on sealed surfaces (mm)	0	10
EVNAP	Percentage of daily Evapotranspiration taken from the Lower zone	0	1
STRNE	Rain-Snow threshold (°C)	-2	3
TFC	Melting rate in forested area (mm/°C/day)	0	10
TFD	Melting rate in open area (mm/°C/day)	0	10
TSC	Melting threshold in forested area (°C)	-2	3
TSD	Melting threshold in open area (°C)	-2	3
TTD	Cold content coefficient for the snowpack	0	3
TTS	Priming threshold of the snowpack (°C)	-5	2

3.2 Study areas and data

The CEQUEAU hydrological model is calibrated on catchments located in three different regions of Canada. Figures 4-6 and 4-7 present the catchments and their location. The first catchment is the Lac St-Jean drainage basin located in the province of Québec extending over an area of 45 261 km². Catamaran brook, the second catchment, is a small forested basin with an area of 52 km². This basin is located in the province of New Brunswick, Eastern Canada. The third catchment is the Nechako River in British Columbia, Western Canada. This basin is extending over an area of 47 200 km². The main source of water flow in the three catchments is the snowmelt accumulated during the winter. The peak of the hydrograph is caused by the spring snowmelt. In this study, the entire annual (from 1st January to 31 December) series are used. Table 4-2 presents the average annual precipitation (P), minimum and maximum flows and the calibration/ validation periods.

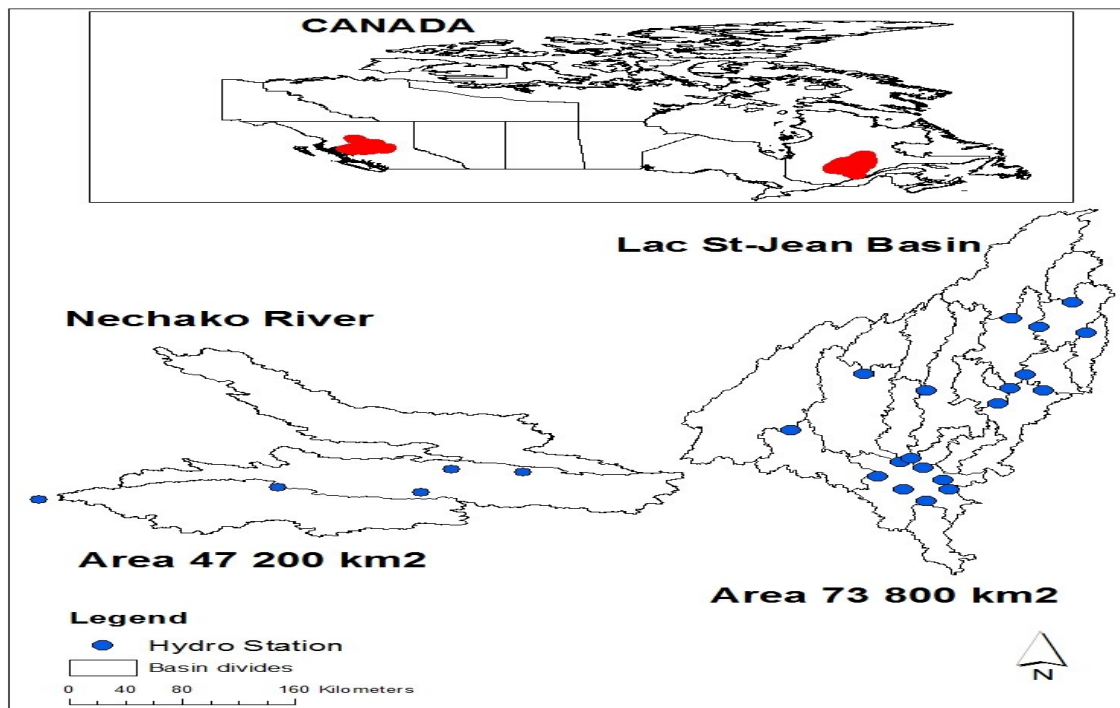


Figure 4-6 Lac St-Jean drainage basin and the Nechako river basin.

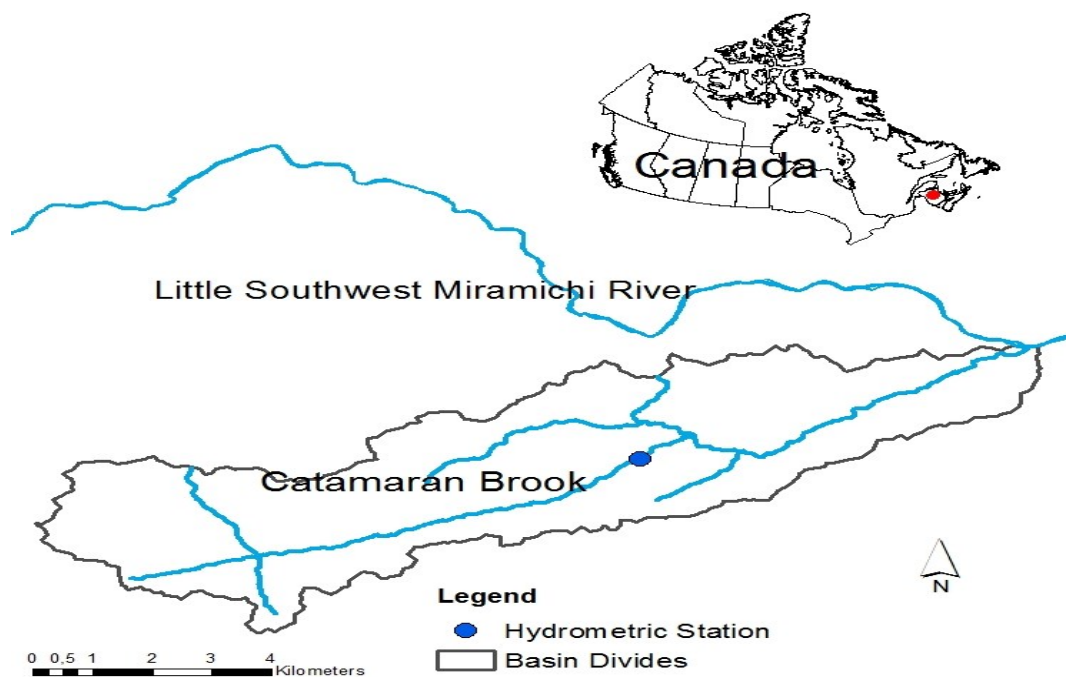


Figure 4-7 Catamaran Brook Basin (adapted from (Caissie et al., 1998)).

Table 4-2 Characteristics of the basins used.

Catchment Name (Notation)	Lac St-Jean (LSJ)	Nechako River (Nech)	Catamaran Brook (Cat)
Area (km ²)	73 800	47 200	52
Altitude (m)	88-792	986-1863	120-343
Annual Mean Precipitation (mm)	914	602	1180
Mean Annual maximum flow (m ³ /s)	4187	771	5,21
Annual mean flow (m ³ /s)	878	190	0,54
Calibration Period	2002-2009	1998-2005	1999-2001
Evaluation Period	2010-2013	2006-2012	2002-2003

3.3 Results

The results of the calibration with NSE as objective function are noted OptNSE and those with the approach based on FDA are noted OptFDA. The NSE values of the calibrated model and the relative bias between the simulated and observed hydrological signatures are reported in Table 4-3. In the Lac St-Jean basin, OptFDA and OptNSE have similar NSE values. OptFDA shows better performance than OptNSE to simulate the flow distribution signatures, high flow frequency and variability as well as global variability. Both calibrations failed to reproduce the frequency of low flows (i.e. average number of days per year having flow $< 0.02 \times$ median flow has a relative bias of 100%). In the Catamaran Brook drainage basin, OptFDA is less performant according to NSE values than OptNSE. However, it has lower bias than OptNSE on almost all signatures. In the Nechako River, both calibrations are equivalent with similar NSE values and signature relative bias. The model performance in this basin drops during the evaluation period because of the differences in peak floods between the two periods. During the evaluation period the flows in the Nechako River reach a maximum of $786 \text{ m}^3/\text{s}$ on 30 June 2007, a value that was not observed during the calibration period, during which the maximum observed discharge was $403 \text{ m}^3/\text{s}$ on 13 May 2005.

Table 4-3 NSE and Signature Relative Bias (%) values for calibration and evaluation periods for both calibrations OptNSE and OptFDA.

			LSJ		CAT		Nech	
Signature		Period	OptNSE	OptFDA	OptNSE	OptFDA	OptNSE	OptFDA
	NSE	Cal	0.89	0.88	0.85	0.73	0.86	0.81
		Eval	0.93	0.93	0.80	0.60	0.67	0.66
Flow distribution	Mean	Cal	-2	-1	-10	-6	-4	-6
		Eval	-7	-6	-2	5	-12	-13
	Median	Cal	-4	-1	-42	-21	-24	-26
		Eval	-14	-8	-38	-1	-36	-38
Flow Variability	Qhv	Cal	7	6	43	31	28	29
		Eval	14	9	45	-5	31	28
	QIV	Cal	-36	-73	78	57	12	29
		Eval	-43	-62	75	61	13	37
	QgV	Cal	-1	-1	14	3	17	14
		Eval	5	4	9	-16	21	16
Flow Frequency	QhF3	Cal	0	-5	17	11	61	59
		Eval	19	8	34	14	42	42
	QhF6	Cal	34	24	34	31	100	100
		Eval	30	11	38	29	76	78
	QIF	Cal	100	100	100	0	0	0
		Eval	100	100	0	0	0	0

Figure 4-8 illustrates the results of the functional Student test for the observed and simulated hydrographs with OptFDA and OptNSE. In the Lac St-Jean, according to the Student statistics the simulated flows during winter (days 45 to 90) are significantly different with OptNSE and OptFDA. However, the t -statistic of OptFDA is lower than that of OptNSE which reflects a slight improvement of the simulation. In Catamaran Brook, the late fall flows (flows observed during days 330-350) are substantially improved with calibration with FDA. In the Nechako River, OptFDA simulate better flows during winter and late summer (days 200-220) while simulations during fall (days 300-350) are less accurate with OptFDA than with OptNSE.

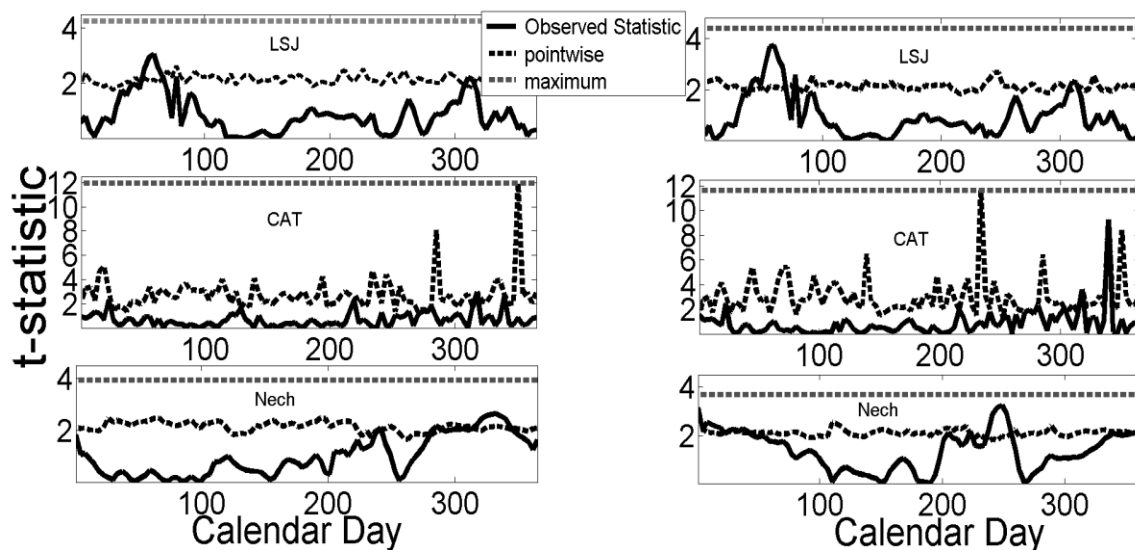


Figure 4-8 Results of the functional Student Test for model simulations during the calibration period. Right panel: results of OptNSE. Left panel, results of OptFDA.

Figures 4-9, 4-10, and 4-11 present the different portions of the FDC reflecting adequacy of high flows, mid-segment and low flows. The maximum observed flow is well estimated by OptNSE in Lac St-Jean drainage basin and Nechako River and well estimated by OptFDA in Catamaran Brook. The higher flows (exceedance probability less than 0.02) are well simulated with OptNSE and OptFDA in the Lac St-Jean. None of the calibrations is able to simulate well this complete portion of the FDC in the other two basins. In the Nechako River, both calibrations provide high

flows that are significantly different from the observed ones. The middle portion of the FDC is better simulated with OptFDA in both Lac St-Jean and Catamaran Brook, meaning that the calibrated model simulates adequately the response of the watershed during moderate flow events. Only flows with exceedance probability between 0.2 and 0.4 are well estimated by OptNSE in Nechako River. Both calibrations have difficulty to catch low flows with exceedance probability over 0.8. However, simulations for this flow range are improved with OptFDA in the Lac St-Jean and Nechako Rivers.

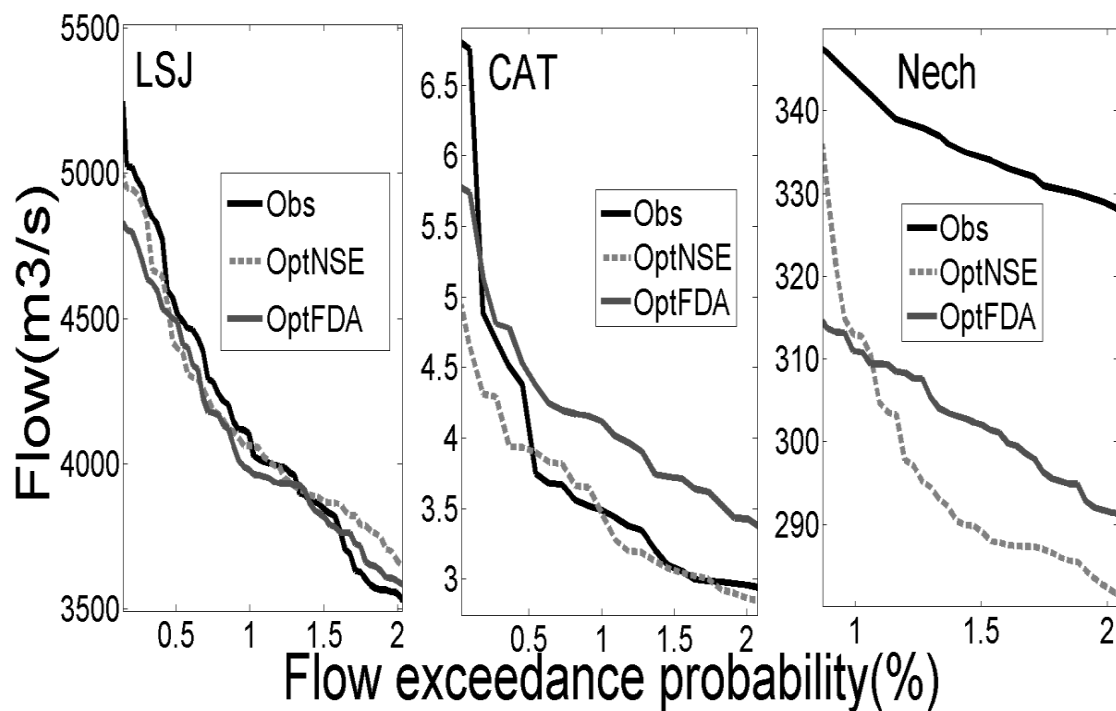


Figure 4-9 The upper part of the Flow Duration Curve representing the response of the catchment to extreme hydrological events.

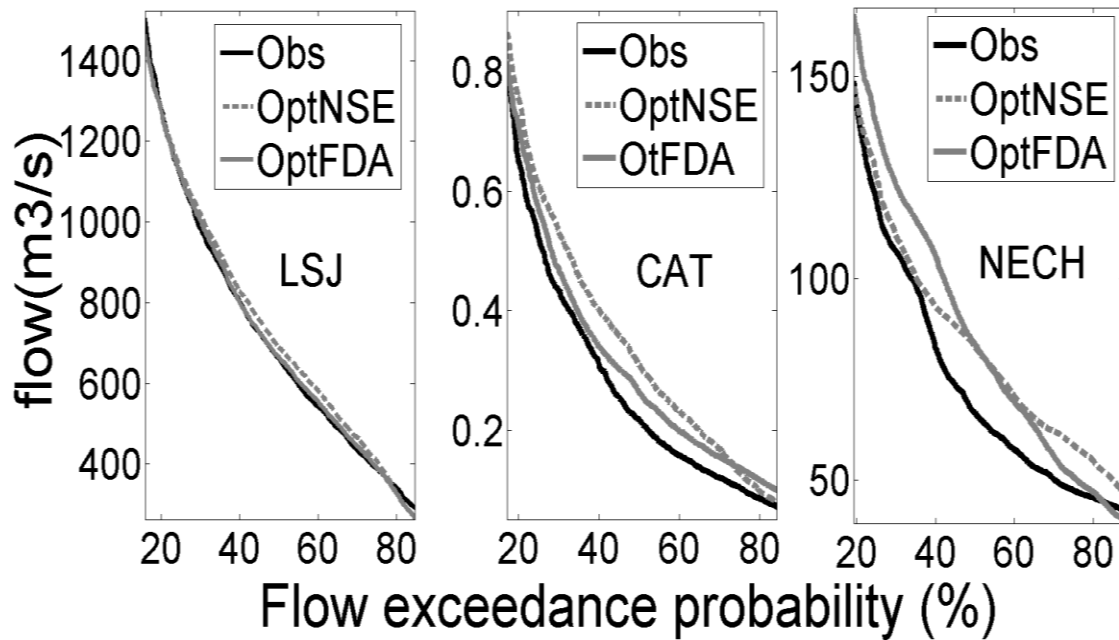


Figure 4-10 The mid-segment of the Flow Duration Curve representing the response of the catchment to moderate hydrological events.

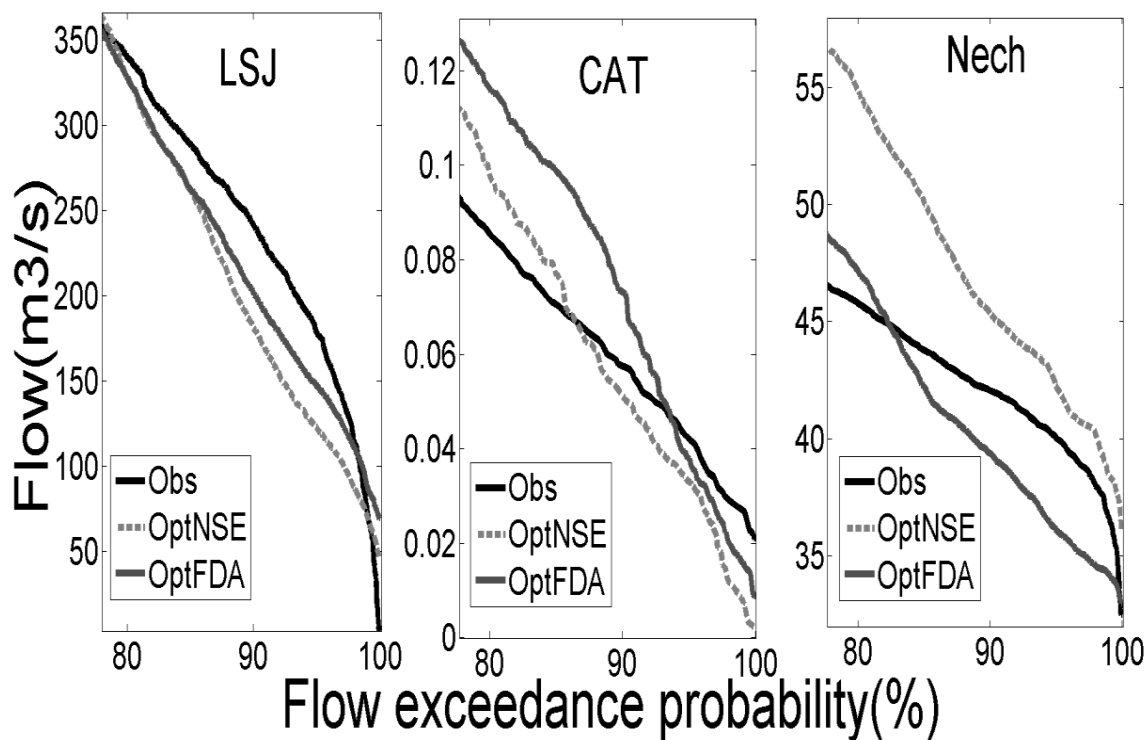


Figure 4-11 The lower portion of the Flow Duration Curve representing flows during drought events.

Table 4-4 provides the parameter values obtained for both calibrations. Depending on the watershed, both strategies fixed similar values to some drainage coefficient and thresholds. For example, in the lac St-Jean drainage basin, infiltration coefficient from Upper Zone to Lower Zone (CIN) is fixed to 0.53, Upper Zone lower drainage coefficient (CVSB) to 0.02 and Evaporation threshold (HPOT) to 41mm. The latter parameter is the threshold at which evaporation from the Upper Zone begins. Significant differences in the three watersheds are noticed for the evapotranspiration parameter EVNAP and lower-zone upper threshold (HNAP). The model adopts different behaviour according to the objective function. For instance, in Catamaran Brook, OptNSE sets lower parameters HINF and HINT (drainage and infiltration thresholds to the lower zone and larger HNAP (lower zone upper threshold) than OptFDA. In other words, OptNSE infiltrates more water to the lower zone than OptFDA. But 65% of the evaporated water comes from the lower zone (EVNAP= 0.65, which is a very high value) while only 19% comes from this zone with OptFDA. The hydrological behaviour simulated with optFDA seems more realistic, as the reduction of water stored in the lower reservoir occurs because of both baseflow contributions to river discharge and evapotranspiration to a lesser extent. In comparison, automatic calibration with OptNSE chiefly reduces the lower zone storage through evapotranspiration. The lower zone reservoir represents the saturated zone and its typical contribution to discharge is related to the baseflow (see Figure 4-5). In the lac St-Jean drainage basin, OptFDA evaporates 8% more from the lower zone than OptNSE, but it sets HNAP to a value 28% lower than OptFDA to supply the baseflow. In the Nechako River, OptFDA sets HNAP to 87mm (against 65mm for OptNSE) and evaporates less water from the lower zone (56%) than OptNSE (67%). In the three basins, OptNSE sets higher values to HSOL (height of the upper zone reservoir) than OptFDA. In fact, OptNSE supplies the runoff mostly from surface runoff and interflow to match the observed discharge during the flood period, but evaporates mostly from the lower zone to match low flows, while the calibration using OptFDA supplies some runoff from the lower reservoir, which corresponds to the baseflow contribution. A greater proportion of evapotranspiration from the upper reservoir, as obtained using OptFDA when compared to OptNSE is also more realistic.

Table 4-4 Calibrated parameter values by the proposed approach (OptFDA) and using the NSE (OptNSE) for the three drainage basins (Lac St-Jean, Catamaran Brook, Nechako River).

Parameter	LSJ		CAT		Nech	
	OptNSE	OptFDA	OptNSE	OptFDA	OptNSE	OptFDA
CIN	0.53	0.53	0.58	0.17	0.19	0.40
CVMAR	0.18	0.40	0.59	0.79	0.65	0.67
CVNB	0.05	0.08	0.07	0.05	0.19	0.31
CVNH	0.42	0.53	0.32	0.53	0.36	0.47
CVSB	0.02	0.02	0.03	0.02	0.01	0.03
CVSI	0.22	0.39	0.58	0.15	0.26	0.32
XINFMA (mm/day)	23.63	19.54	17.88	30.73	19.45	20.12
HINF (mm)	71.18	81.62	58.30	67.26	83.03	81.67
HINT (mm)	84.33	82.44	55.76	67.90	82.58	80.09
HMAR (mm)	341.48	355.99	358.26	453.80	403.85	334.71
HNAP (mm)	144.36	116.61	116.80	80.95	65.60	87.71
HPOT (mm)	41.49	41.59	15.83	17.24	13.60	2.56
HSOL (mm)	209.58	185.44	196.47	190.04	241.86	226.12
HRIMP	6.19	3.94	5.55	7.10	4.62	4.73
EVNAP	0.43	0.51	0.65	0.19	0.67	0.56
STRNE (°C)	-1.45	-1.55	0.04	2.02	1.99	1.93
TFC (mm/°C/day)	2.01	3.56	2.67	5.63	0.38	0.33
TFD (mm/°C/day)	8.10	8.00	6.67	7.74	4.16	3.95
TSC (°C)	0.12	1.13	1.23	2.44	0.08	1.16
TSD (°C)	-1.34	-1.11	-0.12	-0.90	0.37	1.07
TTD	1.65	0.08	0.25	0.73	0.85	0.62
TTS (°C)	-3.40	-3.11	-1.73	-3.36	-3.05	-1.36

4. Discussion

On the proposed calibration strategy

The results show the calibrated model with FDA provides similar results on high flows in the basins with large flood events such as the lac St-Jean basin and the Nechako River with an improvement of low flows. It also improved some aspects (flow frequency and variability) that the NSE is unable to catch like in the case of the Catamaran Brook. The use of the Student statistic during the calibration process allowed reducing the impact of the spring flood in order to fit low flows too.

The proposed methodology considers the entire hydrographs without the subjectivity of choosing specific aspects unlike the conventional multi-objective approaches. For instance calibration with the Kling-Gupta Efficiency (KGE, Gupta et al., 2009) focuses on the correlation, variability error and bias error or the signature based calibration that requires a pre-selection of the watershed patterns to be studied and their associated signatures (Gupta et al., 2008; Yilmaz et al., 2008; Hingray et al., 2010; Euser et al., 2013; Shafii and Tolson, 2015). The traditional statistics or the FDC are not reliable for an objective comparison between model structures. For example, in the case of the Nechako, the NSE and the bias on hydrological signatures are similar for both calibrations. However, the functional Student test, providing a statistic for each time step, favours the model calibrated with FDA over the calibrated model with OptNSE. In fact, this test provides a comparison between the mean simulated and observed flow for each time step (in our case for each day of the year). Thus, it allows detecting when the model fails to reproduce the observed flows for a better understanding and identification of potential uncertainty sources. In a conventional analysis, computing statistics for different time scales is necessary and information about the temporal variability of flow is partially or totally lost.

The proposed calibration scheme is based on exploring and comparing hydrographs over time. This idea was suggested by Schaefli and Zehe (2009) using a wavelet-domain approach in order to assess how closely the time-varying frequency content of simulated flow series is to the time-varying frequency of the observed flow series. They used an objective function based on the kolmogorov-Smirnov distance that measures differences in the autocorrelation properties at a given time step but does not measure the differences in the mean of two time series. To remedy

this problem, a penalization based on the bias was adopted. In this paper, using the FDA framework allowed to assess differences between two time series and compute different statistics (e.g. mean function) describing the entire flow variability according to time. This approach can be adapted to analyze selected parts of the hydrographs (e.g. only the spring flood hydrograph) or selected flows during some period of the year following the same steps provided in the methodology section.

The proposed method requires a pre-analysis of the observed flows to determine how the annual flow series are being smoothed (i.e. number of the basis function K and the smoothing parameter λ). An extra step is required during the calibration process to convert the model simulations to temporal functions. This extra step does not increase the computation time strongly during the calibration process. As hydrographs and statistics are compared per time step, having exactly the same time step and observations per year is required. This might limit the applicability of the method for watershed with missing flow measurements for different periods over the years.

In this paper, the proposed objective function is based on FDA expressed as a distance between simulated and observed hydrographs as well as their derivative weighted by the functional Student statistic. The aim of weighting the objective function by this criterion was to ensure that the model simulates the entire hydrograph adequately (low and high flows simultaneously). However, in the watershed with large snowmelt events such as the Lac St-Jean watershed and the Nechako River, the objective function penalized more high flows than low flows, although an improvement was noted in low flow simulations compared to the NSE. Other formulations based on FDA are possible (e.g. comparing the integrated simulated and observed volumes) and other weighting criteria could be used. Other studies would be needed to explore other formulations.

On the model performance

The calibrated model with FDA improved the flow simulation but it is limited by model conceptualization, which does not completely represent the hydrological processes of the basins. The poor performance of the model in the Lac St-Jean and the Nechako River watershed on low flows can be due to poor performance to simulate accurate evapotranspiration rates or infiltration to the lower zone. According to the Student test, the model is unable to reproduce the winter flows and flows during some days during fall (in the case of the Nechako River). This means that

the baseflow contribution is likely underestimated during these periods. It is possible that the component of the production function controlling infiltration simulates poorly the amounts of water being drained to the lower zone. Due to CEQUEAU's conceptual nature, it is difficult to constrain precisely the infiltration parameters. Indeed, with the proposed calibration procedure, the ability of the model to reproduce the internal hydrological processes was not evaluated and parameter compensation was not addressed. As shown in the results subsection 3.3, according to the objective function, the model tends to compensate different parameters in order to reproduce the observed water volume which yields to different hydrological behaviours. In some cases it can yield unrealistic hydrological behaviours. Calibration based on NSE evaporates high water amounts from the lower zone (up to 67%) which is unrealistic in the study areas. These issues explain in part the poor performance of the model. In this paper only the flow observations were used to calibrate the model. This procedure can be extended to a multi-variable calibration (i.e. using different types of observations such as baseflow and runoff separately) in order to improve the reliability and efficiency of the calibrated model.

5. Conclusions

This paper proposes a new strategy to calibrate a hydrological model that consists of comparing hydrographs as temporal functions instead of daily discharge using functional data analysis. The aim of this strategy is to conserve the temporal structure and the shape of the hydrograph during the calibration and evaluation process. This procedure is general and can be used with any calibration scheme. The main conclusions of this study are:

- Calibration based on FDA provides similar results for large flood events in snow-dominated watersheds as a calibration with NSE, but improves other characteristics of the simulated hydrograph.
- Using functional statistics allows comparing low and high flows in one single step as well as allowing for an objective comparison of periods when simulated and observed flows do not overlap.
- Using functional statistics allows conserving the shape of a hydrograph for a better understanding and distinguishing between the potential uncertainty sources.

- In this paper, the proposed objective function is a distance between the simulated and observed hydrographs and their derivatives. Further work is needed to explore other formulations of objective functions using FDA and more extensive comparisons with advanced calibration strategies.
- Further work is needed to integrate this concept (i.e. comparing hydrographs as temporal functions) in an uncertainty analysis.

Acknowledgement

Funding from NSERC and Rio Tinto Alcan for this project is acknowledged. Authors are also thankful for the technical assistance provided by the hydrological forecast team at Rio Tinto Alcan.

References

- Asadzadeh, M., Tolson, B. and Burn, D. H. (2014). "A new selection metric for multiobjective hydrologic model calibration." *Water Resour. Res.*, 50, doi:10.1002/2013WR014970.
- Bardossy, A. and Singh, S., (2008). "Robust estimation of hydrological model parameters." *Hydro. Earth Syst. Sci.*, (12), 1273-1283.
- Caissie, D., El-Jabi, N. and St-hilaire, A. (1998). "Modelling of stream water temperatures in a small stream using different air to water relations." *Can. J. Civ. Eng.*, 2(25), 250-260.
- Casper, M., Grigoryan, G., Gronz, O., Heinemann, G., Ley, R. and Rock, A. (2012). "Analysis of projected hydrological behaviour of catchments based on signature indices." *Hydrol. Earth Syst. Sci.*, (16), 409-421.
- Charbonneau, R., Fortin, J-P. and Morin, G. (1977). "The CEQUAU model: Description and examples of its use in problems related to water resource management/ Le modèle CEQUEAU: description et exemples d'utilisation dans le cadre de problèmes reliés à l'aménagement." *Hydrological Sciences bulletin*, 22(1), 193-202.
- Chebana, F., Dabo-Niang, S. and Ouara, T. (2012). "Exploratory functional flood frequency analysis and outlier detection." *Water Resour. Res.*, 48, W04514, doi:10.1029/2011WR011040.
- Clausen, B. and Biggs, B. (2000). "Flow variables for ecological studies in temperate streams: groupings based on covariance." *J. Hydrol.*, 237, 184-197.
- Euser, T., Winsemius, H.C., Hrachowitz, M., Fencia, F., Uhlenbrook, S. and Savenije, H.H.G. (2013). "A framework to assess the realism of model structures using hydrological signatures." *Hydrology Earth Syst. Sci.*, 17, 1893-1912.

- Fencia, F., Savenije, H., Matgen, P. and Pfister, L. (2007). "A comparison of alternative multiobjective calibration strategies for hydrological modeling." *Water Resour. Res.*, 43, W03434, doi:10.1029/2006WR005098.
- Glover, F. (1990). "Tabu Search a tutorial." *Interfaces* 20, 74-94.
- Glover, F. and Laguna, M. (1997). "Tabu Search." *Kluwer academic. Springer*.
- Gunkel, A., Shadeed, S., Hartmann, A., Wagener, T. and Lange, J. (2015). "Model signatures and aridity indices enhance the accuracy of water balance estimations in a data-scarce Eastern Mediterranean catchment." *J. Hydrol.*, 4(doi:10.1016/j.ejrh.2015.08.002), 487–501.
- Gupta, H., Kling, H., Yilmaz, K. and Martinez, G. F. (2009). "Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling." *J. Hydrol.*, 80-91.
- Gupta, H., Wagener, T. and Liu, Y. (2008). "Reconciling theory with observations: elements of a diagnostic approach to model evaluation." *Hydrol. Process.*, 2, 3802-3813.
- Gupta, V. H., Sorooshian, S. and Yapo, P. O. (1998). "Toward improved calibration of hydrologic models: multiple and noncommensurable measures of informaton." *Water Ressour. Res.*, 34(4), 751-763.
- Hartmann, A., Wagener, T., Rimmer, A., Lange, J., Brielmann, H. and Weiler, M. (2013). "Testing the realism of model structures to identify karst system processes using water quality and quantity signatures." *Water Resour. Res.*, 3345-3358.
- Hingray, B., Schaeffli, B., Mezghani, A. and Hamdi, Y. (2010). "Signature-based model calibration for hydrological prediction in mesoscale Alpine catchments." *Hydrol. Sci. J.*, 1002-1016.
- Krause, P., Boyle, D. and Bäse, F. (2005). "Comparison of different efficiency criteria for hydrological assessment." *Adv.ances in Geosciences*, 5, 89-97.
- Levitin, D., Nuzzo, R., Vines, B. and Ramsay, J. (2007). "Introduction to Functional Data Analysis." *Canadian Psychology*, 48(3), 135-155.
- Morin, G. and Paquet, P. (2007). *Modèle hydrologique CEQUEAU. INRS-ETE , Rapport de recherche no R000926, 458p.*
- Nash, J. and Sutcliffe, J. (1970). "River flow forecasting through conceptual models. Part I-A discussion of principles." *J. Hydrol.*, 10(3), 282-290.
- Pechlivanidis, I., Jackson, B., McIntyre, N. and Wheeler, H. (2011). "Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications." *Global NEST Journal*, 13(3), 193-214.
- Pechlivanidis, I., Jackson, B., Mcmillan, H. and Gupta, H. (2012). "Using an informational entropy-based metric as a diagnostic of flow duration to drive model parameter identification." *Global NEST Journal*, 14(3), 325-334.
- Pechlivanidis, I., Jackson, B., McMillan, H. and Gupta, H., 2014. "Use of an entropy-based metric in multiobjctive calibration to improve model performance." *Water Resour. Res.*
- Ramsay, J.O. and Silverman, B.W. (2002). *Applied Functional Data analysis: methods and Case Studies.* New York: Springer.

- Ramsay, J.O. and Silverman, B.W. (2005). “*Applied Functional Data Analysis: Methods and Case Studies.*” 2nd ed. New York: Springer.
- Ramsay, J.O. and Silverman, B.W. (2006). “*Functional Data Analysis.*” 2nd ed. New York: Springer.
- Ramsay, J. and Ramsey, J. (2002). “Functional data analysis of the dynamics of the monthly index of nondurable goods production.” *Journal of econometrics*, 327-344.
- Sawicz, K., Wagener, T., Sivapalan, M., Troch, P.A. and Carillo, G. (2011). “Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA.” *Hydrol. Earth Syst. Sci.*, 9(15), 2895–2911.
- Schaeffli, B. and Gupta, H. V. (2007). “Do Nash values have value?.” *Hydrol. Process.*, 21, 2075–2080.
- Schaeffli, B. and Zehe, E. (2009). “Hydrological model performance and parameter estimation in the wavelet-domain.” *Hydrol. Earth Syst. Sci.*, 13, 1921-1936.
- Sguera, C., Galeano, P. and Lillo, R. (2016). “Functional outlier detection by a local depth with application to NOX levels.” *Stoch. Environ. Res. Risk. Assess.*, 30, 1115-1130.
- Shafii, M. and Tolson, B. (2015). “Optimizing hydrological consistency by incorporating hydrological signatures into model calibration objectives.” *Water Resour. Res.*, 51, doi:10.1002/2014WR016520.
- St-Hilaire, A., Morin, G., El-Jabi, N. and Caissie, D. (2000). “Water temperature modelling in a small forested stream: implication of forest canopy and soil temperature.” *Can. J. Civ. Eng.*, 27, 1095-1108.
- Ternynck, C., Ben Alaya, M. L., Chebana, F., Dabo-Niang, S. and Ouarda, T. B.M.J. (2015). “Streamflow hydrograph classification using functional data analysis.” *Journal of hydrometeorology*.
- Tukey, J. (1975). “Mathematics and picturing data.” *Proceedings of the international Congress of Mathematics*. Vancouver
- Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W. and Sorooshian S. (2003). “Effective and efficient algorithm for multiobjective optimization of hydrology models.” *Water Resour. Res.* 39(8), 1214. doi:10.1029/2002WR001746.
- Vrugt, J. A. and Robinson, B. A. (2007). “Improved evolutionary optimization from genetically adaptive multimethod search.” *Proc. Natl. Acad. Sci. U. S. A.*, 3(104), 708-711.
- Wagner, T., Wheeler, H. and Gupta, H. (2004). “*Rainfall-Runoff modelling in gauged and ungauged catchments.*” London: Imperial College press.
- Westerberg, I. and McMillan, H. (2015). “Uncertainty in hydrological signatures.” *Hydrol. Earth Syst. Sci.*, 19, 3951-3968.
- Yapo, P.O., Gupta, H.V. and Sorooshian, S. (1998). “Multi-objective global optimization for hydrologic models.” *J. Hydrol.*, 204, 83-97.
- Yilmaz, K., Gupta, H. V. and Wagener, T. (2008). “A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model.” *Water Resour. Res.*, 44, W09417, doi:10.1029/2007WR006716.

CHAPITRE 5 COMPARAISON D’APPROCHES DE CALIBRATION AVEC ANALYSE D’INCERTITUDE

Les travaux réalisés dans ce chapitre sont présentés sous forme d’article scientifique rédigé en anglais et dont le titre est :

Comparison of different calibration approaches with uncertainty analysis and introducing the functional data analysis approach

Samah Larabi *¹, André St-Hilaire ¹, Fateh Chebana¹

¹INRS-ETE, Québec, Québec, Canada

Cet article est soumis pour publication.

L’auteur principale a proposé la méthodologie et a effectué l’analyse des résultats. Les co-auteurs ont commenté et révisé le manuscrit.

Abstract

Hydrological models are typically calibrated and evaluated using statistical metrics, performance measures that quantify the closeness between the model and the system output, or hydrological signatures that link between the model and the underlying processes. Each signature is constructed to reflect a specific aspect of the system behaviour and each performance measure is sensitive to a certain feature (high or low flows). A combination of these measures is usually used in a multi-objective framework to evaluate model performance to simulate key aspects of the flow regime of a watershed. However, the computation burden of the multi-objective approach limits the number of objective functions to be used. Moreover, when using these metrics valuable information, in particular, temporal streamflow variability, is lost by projecting from the high dimension of data to the single dimension of a summary metric. This paper develops a method that provides a complete and objective assessment of model performance by considering the hydrograph as a single curve. Functional data analysis (FDA), a statistical framework that allows analyzing observed data series as continuous functions (curves) instead of n observations over time, is used for this purpose. The proposed method is compared to a single objective, multi-objective and signature-based calibration. The uncertainty induced by each approach on model outputs is evaluated. The results show that calibration based on FDA can achieve a high level of hydrological consistency and strong performance without the subjectivity associated with the selection of a limited number of performance metrics in a multi-objective calibration or a signature-based calibration.

Keywords: Calibration, Functional Data Analysis, Hydrological Signatures, Multi-objective Calibration, Uncertainty analysis

1 Introduction

Hydrological models are fundamental tools used operationally for forecasting, planning, and management. They are also used for academic and research matters to enhance the knowledge about the hydrological processes and the interaction between climate and land-surface hydrology (Singh and Woolhiser, 2002; Moradkhani and Sorooshian, 2009). These models use several parameters to describe the processes of the real world system. These parameters cannot, in general, be measurable or obtained directly from quantities of catchment characteristics and hence model calibration is needed (Madsen, 2000). The purpose of calibration is to find the parameter set that enables the model to emulate the watershed response as best as possible.

The traditional calibration approach, which is still often used, is a manual calibration by trial and error. This approach consists in tuning the parameters based on visual comparison of simulations and observations or depending on performance measures until a satisfactory solution is obtained. This procedure is subjective, depends on the expertise of the modeler, is not optimal and is time-consuming (Arsenault et al., 2013). The difficulty of such approach rises with the complexity of the model and the number of parameters.

Another alternative to estimate model parameters is to use automatic or semi-automatic approaches. In this case, calibration is considered as an inverse optimization problem (Moradkhani and Sorooshian, 2009). In other words, it is an optimization algorithm based search for a set of model parameter values that minimize the model prediction errors relative to available measured data for the system being modelled (Tolson and Shoemaker, 2007). This approach does not rely on human knowledge or expertise and is fast and objective. However, automatic methods have not yet matured to the point to entirely replace manual methods due to the difficulty of constructing objective functions and optimization algorithms which replicate the human judgment (Pechlivanidis et al., 2011). In fact, the main challenge in automatic calibration is the dependency of the model results on the type and the number of the objective functions used (Kamali et al., 2013). This study will focus on issues around defining an objective function or criteria for model calibration or evaluation.

The widely used criteria for model evaluation and calibration are statistical metrics (goodness-of-fit) called performance measures that quantify the closeness between simulated and observed system output (Van Werkhoven et al., 2009; Pechlivanidis et al., 2010). The majority of these

measures are defined as the sum of squared errors like the Root Mean Squared Error (RMSE) or the widely used Nash-Sutcliffe efficiency (NSE, Nash and Sutcliffe, 1970). This type of performance measures cannot evaluate properly the entire characteristics of the observed time series. The NSE or RMSE, for instance, are sensitive to high flows and lead to a better fit to high flows at the expense of low flows (Krause et al., 2005; Schaefli and Gupta, 2007; Gupta et al., 2009). Because these measures are sensitive to large values, good performance in some periods with high flows is able to dominate the global performance and mask the poor performance in other periods (Zhang et al., 2011). Krause et al. (2005) compared 9 different performance measures and concluded that none of the efficiency criteria performed ideally and each has its advantages and disadvantages that should be taken into account during model evaluation. They also showed that coefficient of determination (R^2) should not be used alone because it can produce high values for very bad model results. Moreover, these performance measures do not provide consistency on model form or indicate how well the model replicate the hydrological functioning of a watershed (Schaefli and Gupta, 2007; Gupta et al., 2009; Van Werkhoven et al., 2009; Pechlivanidis et al., 2011; Shafii and Tolson, 2015).

To remedy this issue, Gupta et al. (2008) urge a model diagnostic approach for an extensive model evaluation. This approach aims at examining the extent to which a model can be reconciled with data and points toward model components which, when assumed to be functioning properly, explain the discrepancy between the simulated and observed system outputs (Gupta et al., 2008). This approach is based on analyzing hydrological signatures that describe the functioning of the watershed and link between the model and the underlying process (Yilmaz et al., 2008; Hingray et al., 2010; Westerberg et al., 2011; Euser et al., 2013). Flow duration curve (FDC) is the most commonly used signature for model calibration and evaluation (e.g. Yilmaz et al., 2008; Westerberg et al., 2011; Pokhrel et al., 2012; Euser et al., 2013; Hartmann et al., 2013; Vrugt and Sadegh, 2013; Pfannerstill et al., 2014; Zhang et al., 2014; Shafii and Tolson, 2015).

Although hydrological signatures are better interpretable than performance measures in terms of underlying processes, they suffer from the same drawbacks. Each signature represents a certain aspect of the catchment at the expense of others because it is constructed to reflect a specific aspect of the system behaviour (Euser et al., 2013). For instance, the FDC provide information on the model performance to simulate the right distribution of flows with no information on

accurate flow timing. Van Werkhoven et al. (2009) recommend using a combination of performance measures with hydrological signatures to capture the different parts of the hydrograph and their timing.

Overall, any single objective function, whether it is a performance measure or signature, used alone is inadequate to reflect properly the entire characteristics of the observed time series (high and low flows) (Gupta et al, 1998; Yapo et al, 1998; Vrugt et al., 2003; Wagener et al, 2004; Pechlivanidis et al., 2011). The reason for this shortcoming is the loss of valuable information by projecting from the high dimension of data to the single dimension of a summary statistic (Gupta et al., 2008). Therefore, the use of several hydrological signatures and performance measures in a multi-objective framework is necessary to describe key features of the system.

The multi-objective approach uses either an aggregation of the considered objective functions or the Pareto Concept. The disadvantage of using an aggregation of multiple objectives is the subjectivity of choosing weights, the dependency of the results on the aggregation and the fact it hides the competitions among the conflicting objective functions (Efstratidias and Koutsoyiannis, 2010). In contrast, the Pareto concept shows the trade-off between the conflicting objectives but comes with a computational burden. This burden might be due to the complexity of single run in complex models or the number of objective functions used. The computational burden of this approach grows quickly with the number of objective functions and the set of Pareto optimal solutions tends to be impractically extended (Efstratidias and Koutsoyiannis, 2010; Pokhrel et al., 2012). To our knowledge, there is no method that defines how to construct a set of optimal objective functions that are conflicting and complementary.

Another limitation of performance measures and hydrological signatures is the loss of information about the hydrograph shape, consequently the representation of the temporal streamflow variability. Actually, the source of uncertainty is different depending on the period of the year, the location and climate of the watershed. For instance, the source of uncertainty for a Nordic watershed on high flows during spring is more likely due to the model component simulating the snowmelt, snow accumulation and measurement error of precipitation or snow water equivalent. On the other hand, the sources of uncertainty on high flows during fall, for the same watershed, are more likely due to rainfall measurement error and the model component(s) simulating infiltration. Hence, it is important to conserve the temporal streamflow variability and

compare observed and simulated streamflows as time series to distinguish between the different sources of uncertainty. A visual comparison between observed and simulated time series is common practice to evaluate model performance and is considered as a modeling success (Schaeffli and Gupta, 2007). This method, however, cannot be used in an automatic calibration process, is subjective and not practical for model evaluation on long time series or for communicating the model performance to potential end-users. For these reasons, there is a need for summary statistics and performance measures that conserve the hydrograph shape.

The objective of this paper is to introduce an evaluation and calibration approach that considers all features of a hydrograph in one single step using statistics that conserve the shape of the hydrograph. To do so, functional data analysis (FDA), introduced by Ramsay and Silverman (2002, 2005), is used. This mathematical framework allows analyzing observed series within a temporal or spatial framework as continuous functions or curves instead of n observations over time. FDA is used in different fields including economy (e.g. Ramsay and Ramsey, 2002), psychology (e.g. Levitin et al., 2007), environmental sciences (e.g. Sguera, et al., 2016) and meteorology (e.g. Ramsay and Silverman, 2002). In hydrology it was used for hydrograph exploratory analysis and outlier detection by Chebana et al. (2012), for streamflow hydrograph classification by Ternynck et al. (2015) and for streamflow forecasting by Masselot et al., (2016).

The key concept of FDA is that we observe at discrete points in time $\{t_j\}$ observations y_j on a continuous process $x(t_j)$. Formally, the observed outcome of the process y_j is as follows $y_j = x(t_j) + \varepsilon_j$ where ε_j is the unobserved error component (Ramsay and Ramsey, 2002). Therefore, the first step of FDA involves nonparametric smoothing techniques to estimate functional data $x(t), \{x(t), t \in \mathcal{T} \subset \mathbb{R}^+\}$. In this paper, functional data represent smoothed hydrographs. Functional data as well as the statistics are represented in a Cartesian coordinate system with time along the x-axis and the value of the dependent variable, here streamflows, along the y-axis as illustrated in Figure 5-1. By doing so, one can analyze both sources of variability that comprise functional data (Levitin et al., 2007). The two sources of variability are amplitude variability, related to the shape which is reflected by peaks and valleys along the y-axis, and phase variability reflecting shifts in timings of these features along the x-axis. Thereby, FDA allows detecting the timing of events in time series and describing the changes in model errors in different hydrological conditions. In the context of this paper, it allows detecting *when* a

hydrological model does not provide reliable simulations in formal mathematical settings than by mere observation and pinpoint toward the potential associated sources of uncertainty.

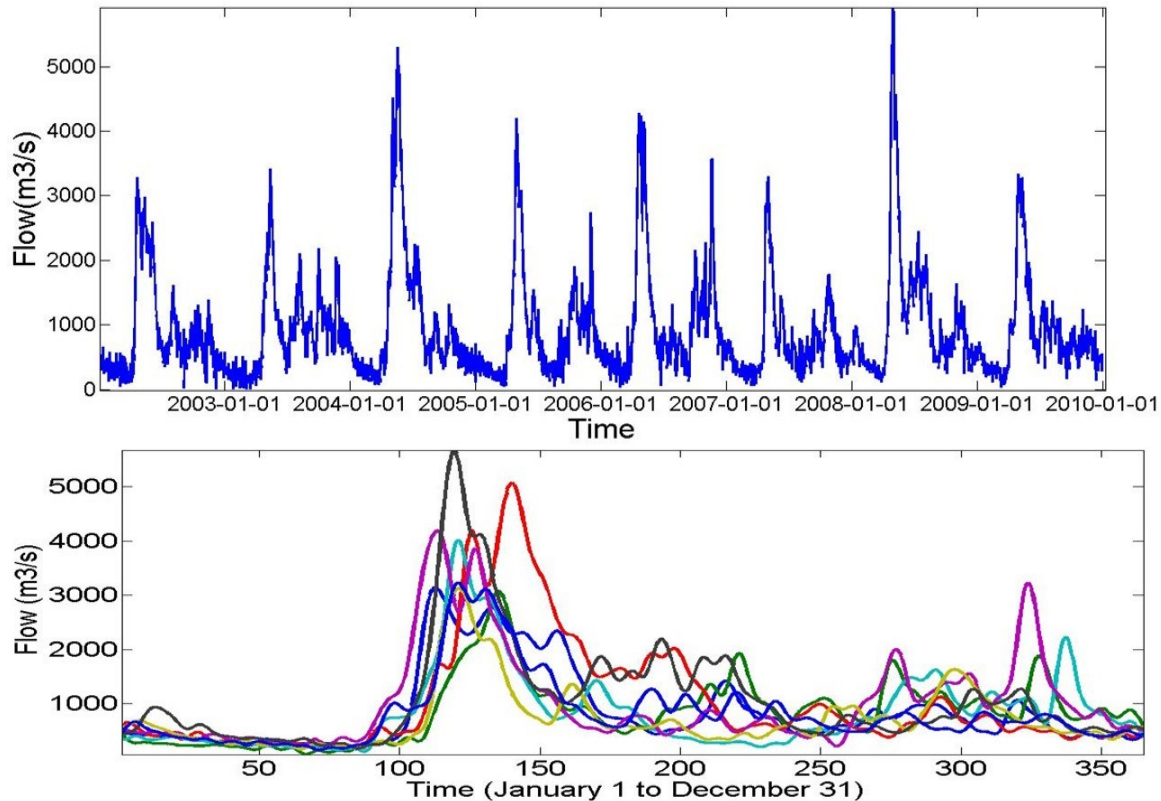


Figure 5-1 Converting raw data to continuous functions to be used with FDA. In the upper panel are the original daily flow series. In the lower panel are the associated annual smoothed hydrographs where each hydrograph is a single observation.

The efficiency of the proposed functional approach is compared to conventional approaches that use single performance measure, multi-objective performance measures and signature-based approach. The impact of the method and the choice of the objective function on model results is also evaluated through an uncertainty analysis. The calibration approaches are compared using the hydrological CEQUEAU model on a Nordic watershed. Functional statistics and the functional student test are used to evaluate model performance and assess how accurate the model outputs are with the system output for each time step.

The present paper is organized as follows: Section 2 describes the methodology adopted for each calibration concept. Section 3 presents the case study. Section 4 presents the results, followed by a discussion and possible ways forward to use FDA in section 5. Conclusions are provided in Section 6.

2 Methodology

Subsection 2.1 presents the methodology adopted for the conventional calibration approaches. Then steps for the functional-based calibration using FDA are presented in subsection 2.2. Calibration criteria selected for each approach are used with a Monte Carlo simulation uniform random search scheme presented in subsection 2.3.

2.1 Calibration with classical approaches

Functional-based calibration is assumed to catch the entire shape and variability of a hydrograph. Therefore, objective functions used for classical approaches are selected so that they include a measure of timing (usually measured by correlation), bias, variability, high flows and low flows. Selection of objective functions based on these criteria allows analyzing the efficiency of each approach to catch adequately the dynamic of the observed flow regime.

2.1.1 Classical single and multi-objective calibration

Single objective calibration is conducted using the widely used performance measure the Nash-Sutcliffe efficiency (NSE, Nash and Sutcliffe, 1970). This performance measure is efficient to capture the timing to peak flow and amplitude of peak flows. The study area, a snow-dominated watershed, motivates this choice. The Nash-Sutcliffe efficiency (NSE) is defined as follows:

$$NSE = [1 - \frac{\sum_{i=1}^n (Q_{sim,i} - Q_{obs,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \bar{Q}_{obs})^2}] \quad (5.1)$$

where: $Q_{sim,i}$ and $Q_{obs,i}$ are the simulated and observed flow at time step i , $\bar{Q}_{obs}$ is the observed mean flow.

Multi-objective calibration is conducted using the Kling-Gupta Efficiency (KGE, Gupta et al., 2009). This criterion is the decomposition of NSE proposed by Gupta et al. (2009) to overcome

the limitations of NSE. Calibration with KGE focuses on optimizing correlation and minimizing variability and bias error as separate criteria. KGE is reported to capture flow variability and peak flow with a high correlation between simulated and observed flows. This criterion is defined as the distance from the ideal point as follows:

$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (5.2)$$

where r is the linear correlation coefficient, $\beta = \mu_{sim}/\mu_{obs}$: the ratio between the simulated and observed mean flows, $\alpha = \sigma_{sim}/\sigma_{obs}$: the ratio between the simulated and observed variance.

2.1.2 Signature based calibration

In this study, 12 hydrological signatures are considered and summarized in Table 5-1. These signatures describe the discharge distribution (mean, median, and variance), global flow variability, high and low flows.

Table 5-1 Hydrological signatures used in this study.

Hydrological Signature	Signature Description	Signature computation based on time series of flows Q
Streamflow distribution	Mean flow	$\sum_{t=1}^N Q_t / N$, t time step
	Median flow	Median flow of the entire flow data Q
	Standard deviation of flows	
	Lag-1 autocorrelation coefficient	$\sum_{t=1}^{N-1} (Q_t - \bar{Q}) / (Q_{t+1} - \bar{Q}) / \sqrt{\sum_{t=1}^N (Q_t - \bar{Q})^2}$
Flow variability	Global variability	Coefficient of variation: ratio between standard deviation and mean flows
	High flow variability	Ratio of mean annual maximum to the median flow
FDC	FDC mid-segment slope	$\log(Q_{m1}) - \log(Q_{m2})$, where $m1$ and $m2$ are the lowest and highest flow exceedance probability within the mid-segment of FDC (here 0.2 and 0.7) (Yilmaz et al., 2008; Shafii and Tolson, 2015).
	FDC high flow volume	$\sum_{h=1}^{H} Q_h$, where $h=1, 2, \dots, H$ are flow indices located within the high flow segment (exceedance probability lower than 0.02) (Yilmaz et al., 2008, Shafii and Tolson, 2015).
	FDC low flow volume	$\sum_{h=1}^{H} Q_h$, where $h=1, 2, \dots, H$ are flow indices located within the low flow segment with H is the index of the maximum flow(exceedance probability higher than 0.07) (Yilmaz et al., 2008; Shafii and Tolson, 2015).
Annual flow indices	Annual mean maximum flow	Mean of the annual maximum flow
	Annual mean flow	Mean of the annual mean flow
	Annual mean volume	Mean of the annual volume

Different signature-based calibration and evaluation approaches exist in the literature. Shafii and Tolson (2015) proposed a signature-based approach that considers explicitly the number of signatures adequately reproduced by the model according to an acceptable threshold. They showed that when satisfaction of signatures is considered explicitly in the search for highly consistent parameters, the high-consistency regions in the parameter search space are explored more efficiently. Thus, this approach results in a larger number of consistent parameters sets and a higher level of hydrological consistency compared to the other traditional signature-based calibration approaches. For these reasons, we consider this approach that consists on optimizing the number of signatures reproduced according to an acceptable threshold through a consistency metric.

The consistency metric is computed as follows:

- For each hydrological signature a deviation is computed as a relative bias:

$$D_i = \frac{S_i^{sim} - S_i^{obs}}{S_i^{obs}} \times 100 \quad (5.3)$$

where S_i^{sim} and S_i^{obs} are the simulated and observed signature.

- Each signature is scored with a continuous scoring function as follows:

$$C_i = \begin{cases} 1, & |D_i| \leq |D_i^*| \\ 0, & |D_i| \geq |D_{i,max}| \\ \frac{|D_{i,max}| - |D_i|}{|D_{i,max}| - |D_i^*|}, & otherwise \end{cases} \quad (5.4)$$

where $D_{i,max}$ is selected as 100%, D_i^* is an acceptance threshold and D_i is the signature deviation.

- Signatures are sorted in a descending order and overall consistency metric is defined as the sum of the highest $n^* + 1$ score values as follows :

$$SIG_{OF} = \sum_{i=1}^{n^*+1} C_i, \text{ where } n^* \text{ is the number of the satisfied signatures.} \quad (5.5)$$

The overall consistency metric (Eq. 5.5) to be maximized provides information about the number of the hydrological signatures that are satisfactorily reproduced by simulations (the integer part representing the number of signatures having a scoring equal to 1 with Eq. 5.4) as well as some

information about the amount of deviation from the acceptability threshold obtained from unsatisfactory scoring functions (fractional part) (Shafii and Tolson, 2015). Because 12 signatures are considered in this study, this criterion takes values between 0 and 12. The value of 0 reflects a highly inconsistent model (i.e. no signature is satisfied according to the acceptance threshold) while the value of 12 means that the model is fully consistent (i.e. all signatures are reproduced satisfactorily).

2.2 Calibration with functional data analysis

The first step in FDA (Ramsay and Silverman, 2002) is smoothing the data by converting raw data into continuous functions via basis function expansion. There is a variety of basis function (constant, polynomial, spline, Fourier series, wavelets...) adapted for each case study. In our case, because data are periodic the natural choice is a Fourier basis (similar to Chebana et al., 2012). In this section, the procedure for smoothing as well as functional data statistics used and calibration criteria are presented.

2.2.1 Converting raw data to functional data

The first step in FDA is to convert the original time series of discharge (daily, in the present study) into smoothed curves. For a set of daily flows observed over n years, each annual series $\mathbf{Y}_i = \{Y_{i,1}, \dots, Y_{i,365}\}$ where $i=1, \dots, n$, is transformed into one continuous temporal function $x_i(t)$ $\{x_i(t), t \in [1, T = 365]\}$ or curve representing the smoothed hydrograph. Each continuous function $x_i(t)$ is generally defined as a linear combination of functions (Ramsay and Silverman, 2002):

$$x_i(t) = \sum_{k=1}^K c_{i,k} \phi_k(t) = \mathbf{c}_i' \boldsymbol{\phi}(t) \quad (5.6)$$

where $\boldsymbol{\phi}(t)$ is the K -vector of basis functions and $\mathbf{c}_i$ is the coefficient vector of length K .

In our case, the curves are constructed from Fourier basis functions as follows:

$$x_i(t) = c_{i,1} + c_{i,2} \sin \omega t + c_{i,3} \cos \omega t + c_{i,4} \sin 2\omega t + \dots c_{i,K} \cos K\omega t, \quad (5.7)$$

where $\omega = \frac{2\pi}{T}$, T is the period and here $T=365$, K is the number of basis functions.

The basis coefficients $\mathbf{c}_i$ are estimated so that the resulting curves are smooth to a certain degree and adequately fit the data. In fact, the number of basis controls both smoothing and how

optimally the curves will fit the data. The greater the value of K is, the better the fit of the constructed curves with the data (the lower the bias is). A smaller K value leads to smoother curves (low variance) but at the expense of the capacity to capture sharper features in the time series (Levitin et al., 2007). So, the trade-off consists in selecting a relatively rich basis functions (large K) to catch most of the variability in the observed data, while leaving out some of the spurious noise. In order to find a trade-off between low bias and high variance a penalty least squared errors (Ramsay and Silverman, 2002) is used to estimate the coefficient:

$$\text{PENSSE}_m(\mathbf{y}_i|\mathbf{c}_i) = \sum_{j=1}^T (y_{i,j} - x_i(t_j))^2 + \lambda \text{PEN}_m[x_i] \quad (5.8)$$

where $y_{i,j}$ is the observed flow at day t_j of the i^{th} year, $x_i(\cdot)$ is the curve to be estimated for the i^{th} year, λ is the smoothing parameter, $T=365$ days is the period and PEN_m is the roughness penalty defined as the integral of the squared m^{th} derivatives of the function $x_i(t)$.

When λ approaches 0, the curve is a rough sinusoidal interpolation of data and the greater λ is, the smoother the curve is. This parameter is estimated by a generalized cross validation which is not time-consuming and tends to smooth more than an ordinary cross validation. Once data are converted, each constructed curve represents an annual smoothed hydrograph, which constitutes a single observation instead of the original 365 observations from which it is constructed. In this context, the noise that is associated with measurement error is reduced through smoothing.

2.2.2 Functional statistics and calibration criterion

This sub-section presents functional statistics used in this study. For a sample of functions $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ defined in a domain D (here $D = [1, 365]$), the mean function is:

$$\bar{x}(t) = 1/n \sum_{i=1}^n x_i(t), \quad t \in D. \quad (5.9)$$

The variance is defined as:

$$\text{var}_X(t) = 1/(n-1) \sum_{i=1}^n (x_i(t) - \bar{x}(t))^2, \quad t \in D \quad (5.10)$$

A coefficient of variation function can be defined as the ratio between the standard deviation curve and the mean curve. Thus, it allows evaluating the flow variability for each time step.

The functional version of the Student statistic (Ramsay and Silverman, 2006) for two samples of functional $\mathbf{x}_1$ and $\mathbf{x}_2$ is a function of time and is defined as follows:

$$t - statistic(t) = \frac{|\bar{x}_1(t) - \bar{x}_2(t)|}{\sqrt{\frac{1}{T}\{var[x_1(t)] + var[x_2(t)]\}}} \quad \forall t \in D \quad (5.11)$$

Functional Student test compares the means of two groups over time. In our case, this test will be used to verify the accuracy of the simulated flows for each time step and detect *when* the simulated and observed flows are significantly different.

One way to compare between two curves is to compare the areas under each curve. In our context, evaluation of the model is conducted by comparing the area under the simulated and the observed hydrograph (i.e. the simulated and observed integrated annual volume). The objective function is defined as follows for each year i :

$$of_i = \left| 1 - \int_{t=1}^{T=365} Q_{i,sim}(t) / \int_{t=1}^{T=365} Q_{i,obs}(t) \right| \quad (5.12)$$

where : $Q_{i,sim}(t)$ and $Q_{i,obs}(t)$ are the simulated and observed smoothed hydrographs at the i^{th} year as described with equation (5.7).

To evaluate the performance of a model over n years of data during the calibration process, each model simulated smoothed hydrographs are compared to the observed one through the Student test. Then, the sum of of_i with i in $\{1, \dots, n\}$ is weighted by the mean Student statistic to compute the global criterion. This weighting criterion aims at amplifying the performance measure value, to be minimized, of parameter vectors not reproducing the flow adequately for each time step (i.e high and low flows). The global criterion is defined by equation (5.13):

$$FDA_OF = S \times \sum_{i=1}^n of_i \quad (5.13)$$

where : of_i is as defined in Eq. 5.12, S is the mean of the student test statistic given by Eq. 5.11, and n is the number of years of observations.

2.3 Parameter search procedure

To define an optimal ensemble of acceptable sets of parameters with each approach and analyze the uncertainty of the inferred space, a Monte Carlo simulation uniform random search with a depth function is used. This method is an adaptation of the Robust Parameter Estimation algorithm (ROPE) proposed by Bardossy and Singh (2008). The notion of depth function was first introduced by Tukey (1975) as a statistical measure to quantify the centrality of a point

within a multidimensional set. More precisely: ‘Associated with a given distribution P on R^d , a depth function is designed to provide a P -based center-outward ordering (and thus a ranking) of points x in R^d . High depth corresponds to ‘centrality’, low depth to ‘outlyingness’”(Liu et al., 2006). The aim of using a depth function within the Monte Carlo simulations is to guide the search for the parameter space containing the best performing parameters (i.e. sets of parameters leading to a performant model).

The principle of the methodology consists of ranking at each iteration, a new independent generation of parameters sets with a depth function based on its closeness to the center of the best sets of the last generation. The deeper is a point (having a high depth value) the closer it is to the best sets of the previous generation. The method starts by generating uniformly an initial generation of N (here N is 1000) parameter sets. According to the objective function, a subset of the best performing parameter sets is selected. A new generation is generated from a uniform distribution within the bounds of this set and only the deepest points according to a threshold (e.g. 1st quartile of depths) form the new generation for the next iteration. Therefore, the method samples exclusively parameter sets located in the region of the parameter space including performing parameters until the overall performance indicator ceases to improve.

3 Case study

3.1 Hydrological model description

The model used in the present study and briefly described in this subsection is the CEQUEAU semi-distributed hydrological model (Charbonneau et al., 1977; St-Hilaire et al., 2000; Morin and Paquet, 2007). The model takes into account the physical characteristics of the drainage basin and their variation in space via a partition of the watershed into equal areas called whole squares. For each whole square, a hydrological budget is calculated by the production function (see Figure 5-2) which simulates the vertical flow considering the soil as two interconnected reservoirs. Each whole square is subdivided into a maximum of four so-called partial squares according to the sub-basin divides in order to simulate the upstream-downstream flow with a transfer function. The model requires inputs including physiographic data for each whole square (altitude of the southwest corner of the square, forested area, lake area, and marsh area), daily

precipitation, minimum and maximum daily air temperatures. The 18 parameters considered for model calibration and presented in Table 5-2 are associated with the processes of snowmelt, infiltration, direct surface runoff, base flow, and evapotranspiration.

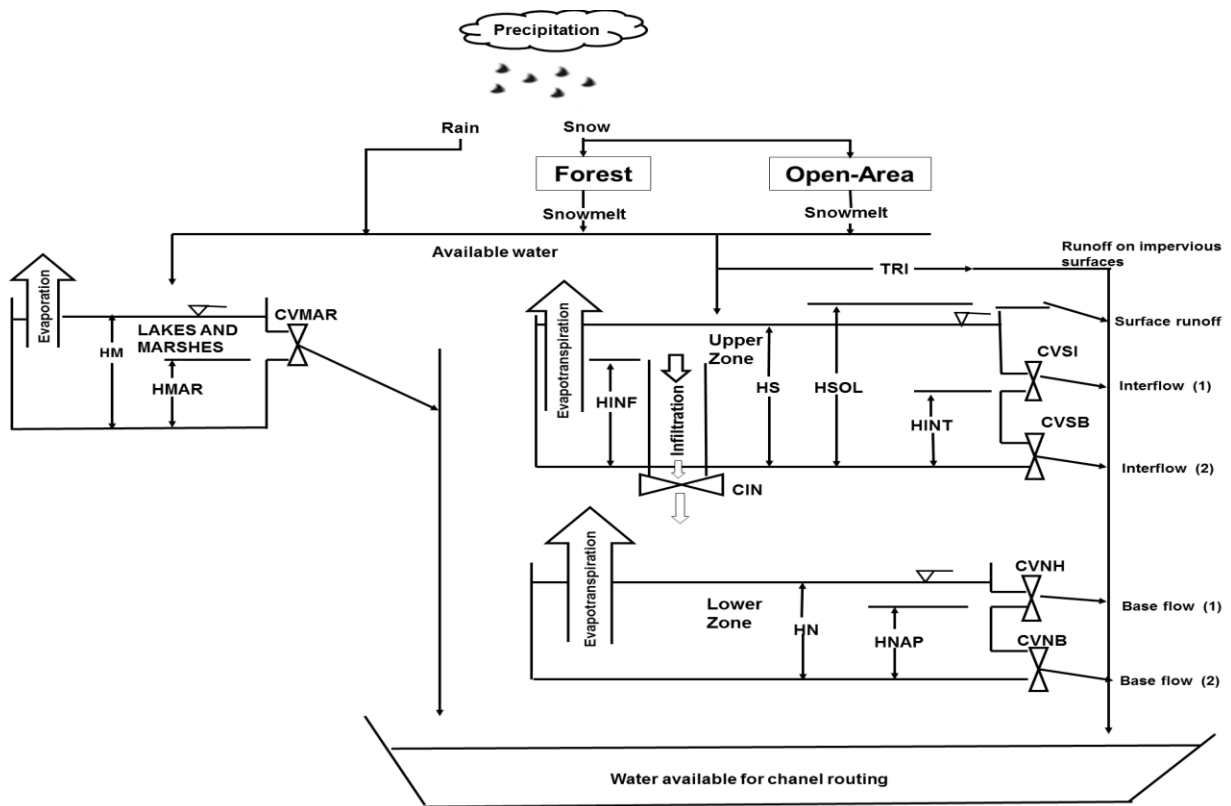


Figure 5-2 Production function of CEQUEAU model (adapted from Morin et Paquet 1995).

Table 5-2 Model key parameters.

Parameter	Description	Range	
STRNE	Rain-Snow threshold (°C)	-2	3
TFC	Melting rate in forested area (mm/°C/day)	0	10
TFD	Melting rate in open area (mm/°C/day)	0	10
TSC	Melting threshold in forested area (°C)	-2	3
TSD	Melting threshold in open area (°C)	-2	3
TTD	Cold content coefficient for the snowpack	0	3
TTS	Priming threshold of the snowpack (°C)	-5	2
CIN	Infiltration coefficient from Upper Zone to Lower Zone	0	1
CVNH	Lower- Zone upper drainage coefficient	0	1
CVSB	Upper- Zone lower drainage coefficient	0	1
CVSI	Lower- Zone intermediate drainage coefficient	0	1
XINFMA	Daily Maximum infiltration (mm/day)	0	40
HINF	Percolation threshold from the upper to the lower zone (mm)	0	100
HINT	Upper- Zone intermediate drainage threshold (mm)	0	100
HNAP	Lower- Zone upper threshold (mm)	20	200
HPOT	Threshold of evaporation at the potential rate (mm)	0	80
HSOL	Height of the Upper Zone (mm)	100	300
EVNAP	Percentage of daily Evapotranspiration taken from the Lower zone	0	0.25

3.2 Study area and data

The present study is carried out on the Chute-Du-Diable drainage basin, which is a sub-basin of the lac St-Jean drainage watershed in the Saguenay-Lac-Saint-Jean region in Québec, Canada (Figure 5-3). The Chute-Du-Diable drainage area is 9 701 km² with altitudes ranging between 215 and 615m. The region is mainly forested (90% of the area). The flow regime is typical of the Nordic basin dominated with spring snowmelt with less significant floods during the fall. For the period 1985-2013, annual mean precipitation is around 1000mm and annual mean discharge is 204 m³/s. Discharge data from 1990 to 2013 are used for model calibration and evaluation. The

first year is used as a warm-up period, the next nine years (1991 to 1999) are used for calibration and the remaining years are used for model evaluation.

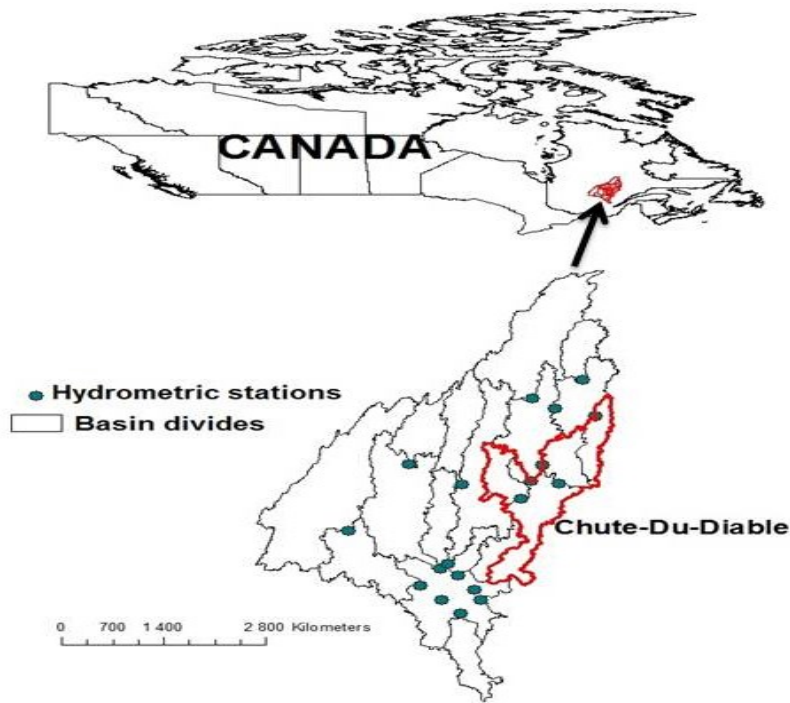


Figure 5-3 The localization of the Chute-Du-Diable watershed within the great lac St-Jean drainage basin.

4 Results

This section presents the results of the four calibration approaches. We note OptNSE the results of the calibration with NSE, OptKGE the results of the calibration with KGE, OptFDA are the results with the functional-based calibration. For the signature-based calibration, two acceptance thresholds D_i^* are fixed. The results of the acceptance threshold at 5% are noted OptSIG5 and results for acceptance threshold at 10% are noted OptSIG10.

Subsection 4.1 presents the impact of the calibration approach on the convergence of the algorithm. Subsection 4.2 presents the results of the global performance of each approach while subsection 4.3 compares the resulted parameter distribution and the performance of the

parameterized model with parameter mean values. Subsection 4.4 presents an evaluation of uncertainty inferred by each approach on the simulated flows. In subsections 4.3 and 4.4, the simulated flows are converted to functional data and analyzed with FDA.

4.1 Convergence of the algorithm

Figure 5-4 presents the convergence of the algorithm using each calibration approach. The algorithm converges in fewer than 25 iterations when using KGE or functional-based approach. It required more iterations to converge when using NSE or signature-based approach. The extra step (i.e. smoothing hydrographs for each simulation) required by the functional-based calibration does alter slightly the computation time. However, it remains reasonable and relatively similar to computation time required by the simplest approaches. In fact, it required 4 hours using a functional-based approach that converged after 25 iterations, whereas it required 3 hours using the KGE that converged after 22 iterations. The NSE calibration and signature-based calibration required more than 4 hours computation time. Note that CEQUEAU takes 2.15 seconds to simulate 10 years of data on an Intel i7 2.7 GHz processor on this watershed. The computation time increases with the number of the simulated years and in particular the number of whole squares to partition the watershed. This number depends highly on the watershed area.

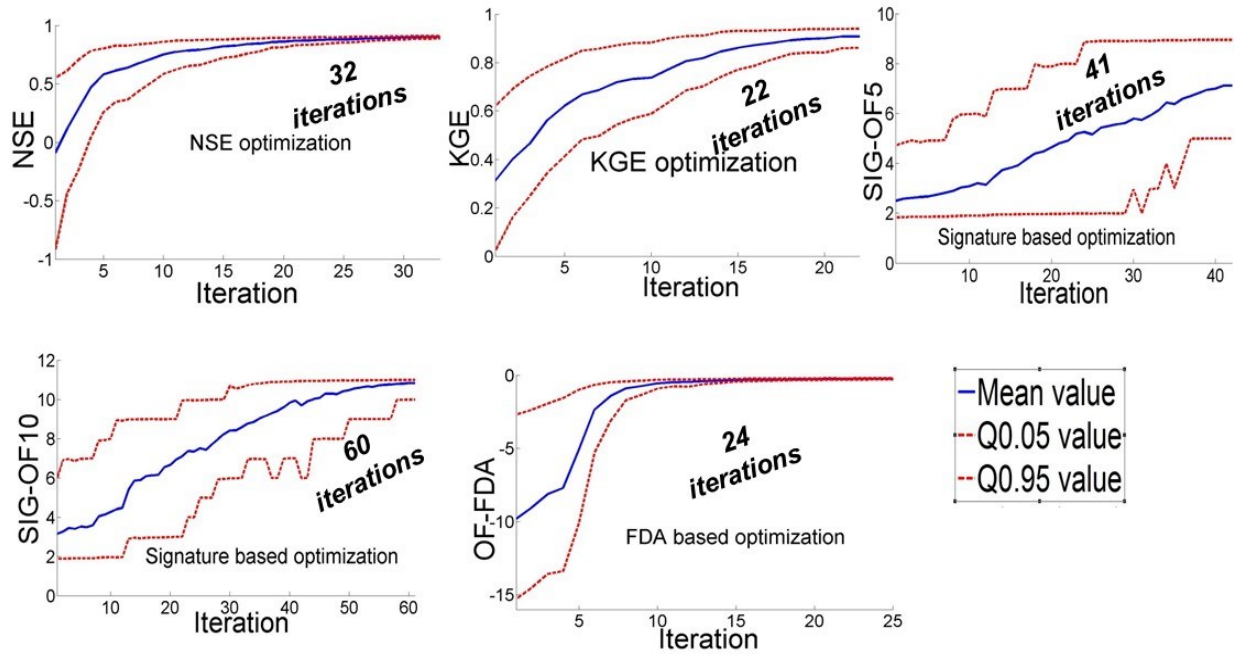


Figure 5-4 Convergence of the algorithm using different calibration approaches (OptNSE using NSE, OptKGE using KGE, OptSIG5 and OptSIG10 using the signature-based calibration at the 5% and 10% acceptance threshold resp., and OptFDA using the functional approach).

4.2 Comparison of the global performance of the four approaches

Each calibration approach yields to a parameter set that was evaluated in terms of performance (i.e. using NSE and KGE) and consistency. The latter is evaluated by considering the integer part of equation 5 representing the number of signatures being satisfied. The number of signatures being satisfied at the 5% acceptance threshold is noted SIG5 and at the 10 % acceptance threshold is noted SIG10.

Figure 5-5 and 5-6 show the scatter plot of each parameter set using the performance and consistency criteria. All calibration approaches are prone to equifinality with several parameter vectors having similar criteria values. These parameter vectors are not equally transferable in time in particular for OptNSE and OptKGE sets. In other words, parameter vectors having similar performance/consistency during calibration period does not have necessarily similar performance/consistency during the evaluation period. This corroborates the findings of

Bardossy and Singh (2008) that equally performing parameters are not necessarily equally transferable.

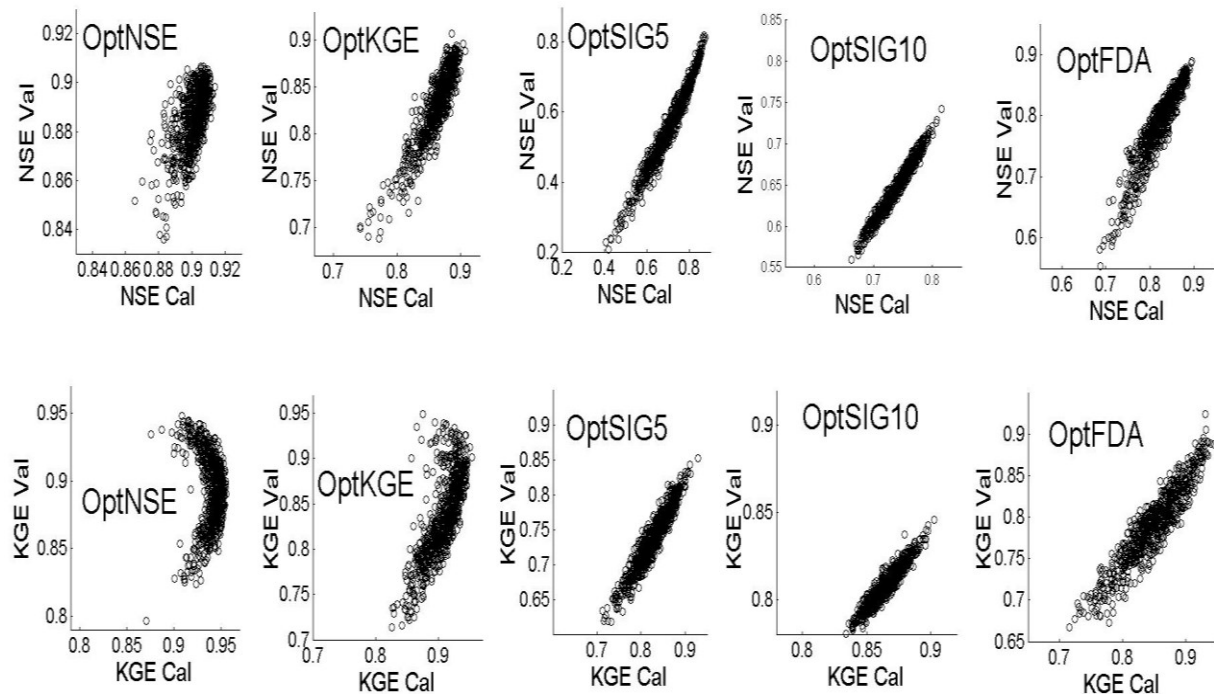


Figure 5-5 Model performance during calibration period vs model performance during the evaluation period.

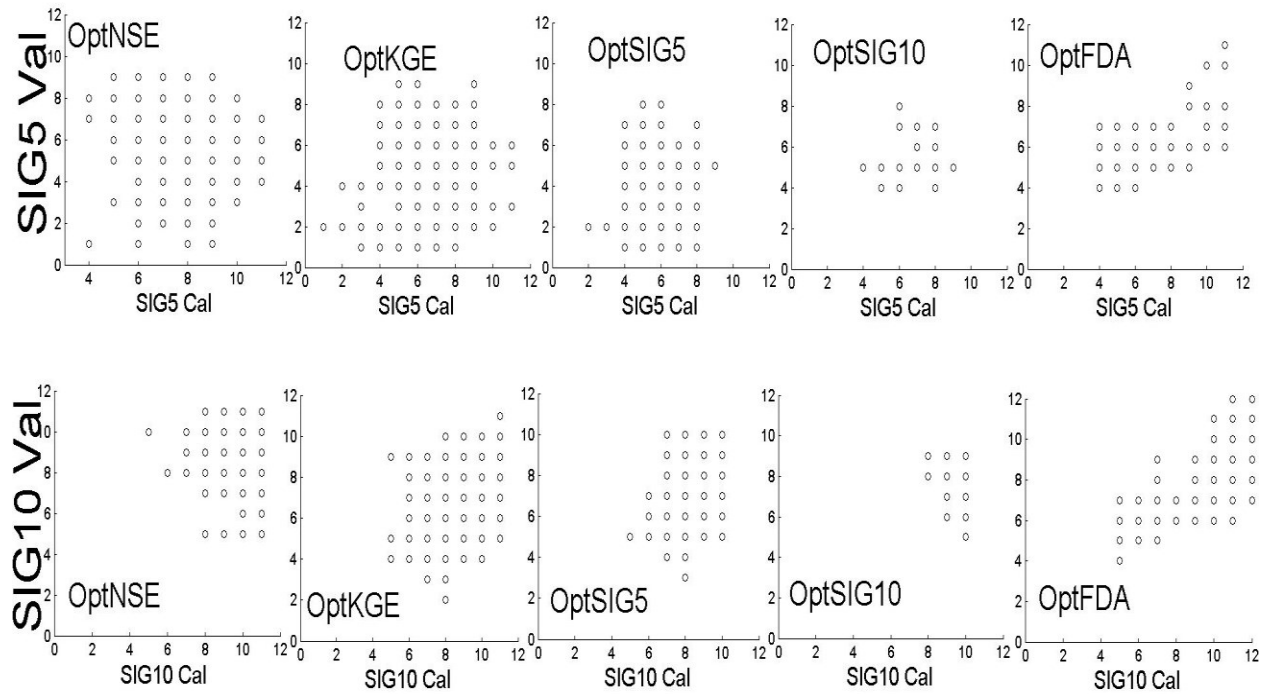


Figure 5-6 Model consistency during calibration period vs model consistency during the evaluation period. SIG5 is the number of signatures satisfied at the 5% acceptance threshold and SIG10 at 10% acceptance threshold.

Parameter sets provided by calibration using NSE or KGE are consistent during the calibration period. The consistency of these sets drops drastically during the evaluation period. On the other hand, signature-based calibration leads to performant parameter sets in particular at the 10% acceptance threshold. These findings are consistent with those of Shafii and Tolson (2015) that signature-based calibration leads to performant parameter sets from a statistical measure perspective. The functional based-calibration leads to consistent and performant parameter set. In fact, some parameter vectors are fully consistent and robust. About 3% of OptFDA satisfy the 12 signatures considered in this study during calibration and about 1.7% during validation. None of the other calibrations provided a parameter vector that satisfies all signatures as seen in Figure 5-6.

4.3 Comparison of the parameter distribution and its performance

Parameter distribution for each set is provided in Table 5-3. The set of parameters OptSIG10 is the least dispersed with the smallest standard deviation. Calibration using KGE or NSE yield to similar mean parameter values for almost all parameters. However, parameters of OptKGE are more dispersed as well as their performance (see Figure 5-5). Signature-based calibration at both acceptance threshold 5% and 10% converge to similar parameter mean values except for the rain-snow threshold (STRNE). Overall, this parameter is highly sensitive to the objective function used. Other parameters controlling the dimension of the Lower zone and its contribution to the baseflow (HNAP and CVNH) also show a high sensitivity to the objective function.

Table 5-3 Mean and standard deviation of parameter values obtained by the different calibration approaches. Highlighted values are the lowest standard deviation.

	OptNSE		OptKGE		OptSIG10		OptSIG5		OptFDA	
Parameter	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
STRNE (°C)	-0.50	0.24	-0.65	0.42	-0.76	0.08	0.29	0.52	-0.89	0.31
TFC (mm/°C/day)	3.27	0.30	3.31	0.62	1.54	0.09	1.54	0.22	2.56	0.53
TFD (mm/°C/day)	6.88	0.92	7.00	1.08	6.86	0.21	6.04	0.96	6.21	0.97
TSC (°C)	1.48	0.31	1.44	0.44	1.58	0.10	1.37	0.36	1.30	0.58
TSD (°C)	-0.60	0.43	-0.55	0.52	-0.89	0.08	-0.48	0.35	-0.55	0.52
TTD	0.38	0.11	0.40	0.14	0.33	0.02	0.46	0.08	0.36	0.10
TTS (°C)	-2.08	0.59	-2.37	0.87	-2.45	0.28	-2.16	0.73	-1.71	0.99
CIN	0.41	0.10	0.51	0.17	0.59	0.04	0.48	0.10	0.51	0.14
CVNH	0.16	0.02	0.40	0.17	0.52	0.05	0.51	0.12	0.47	0.17
CVSB	0.03	0.00	0.04	0.01	0.04	0.00	0.05	0.01	0.01	0.00
CVSI	0.40	0.11	0.37	0.15	0.38	0.07	0.42	0.08	0.47	0.16
XINFMA (mm/day)	20.86	4.04	20.09	6.32	19.65	2.27	16.52	3.13	21.74	6.23
HINF (mm)	80.03	3.28	71.63	7.22	58.19	1.85	63.45	5.59	82.40	4.40
HINT (mm)	79.69	4.41	80.04	7.74	61.29	2.55	61.26	6.77	79.44	3.39
HNAP (mm)	105.19	16.75	123.75	27.11	104.89	7.40	105.79	19.44	80.85	18.21

HPOT (mm)	23.90	6.12	21.63	7.87	18.28	0.60	15.00	3.42	42.32	4.74
HSOL (mm)	203.24	24.08	198.31	30.64	199.01	10.56	202.58	21.19	207.07	31.66
EVNAP	0.17	0.02	0.16	0.03	0.16	0.01	0.14	0.02	0.14	0.03

To evaluate the performance of each approach to simulate the entire flow regime, we evaluate the performance of the model parameterized by the mean values. Figure 5-7 presents the observed and simulated mean interannual hydrographs. The parameterization OptNSE, OptKGE and OptFDA are performant in terms of simulating the rising limb, the falling limb and timing of the peak flow. OptNSE shows superiority in terms of the peak amplitude.

The observed and simulated spring flood by OptSIG5 and OptSIG10 are not synchronized. This can be explained by a low snowmelt rate TFC set by these calibrations in comparison to the other approaches. OptFDA provides the most accurate low flow simulations. In fact, the Lower zone upper threshold of OptFDA is lower by almost 24 mm than the value of OptNSE, OptSIG5, OptSIG10 and about 43mm than the value of OptKGE. This means that the contribution from the Lower zone to the baseflow is higher with OptFDA than with the other parameterizations. OptFDA has the highest bias on the peak amplitude. The threshold at which evaporation from the upper to the Lower zone begins (HPOT) of OptFDA is about 2.8 times the value of OptSIG5 and more than 1.7 times the value of OptNSE and OptKGE. This could mean that the model retains more water in the upper Zone with OptFDA than with the other approaches, which might explain in part the overestimation of the spring flood.

With all configurations, the model is unable to capture flows during summer, in particular days from 160 to 180. This model deficiency is due to the fact that parameters of potential evapotranspiration were not adjusted. Potential evapotranspiration in the CEQUEAU is simulated with the Thornthwaite equation, which is a simple formulation based on air temperature. Other physically-based evapotranspiration formulations are currently being tested with CEQUEAU model (St-Hilaire et al., 2015).

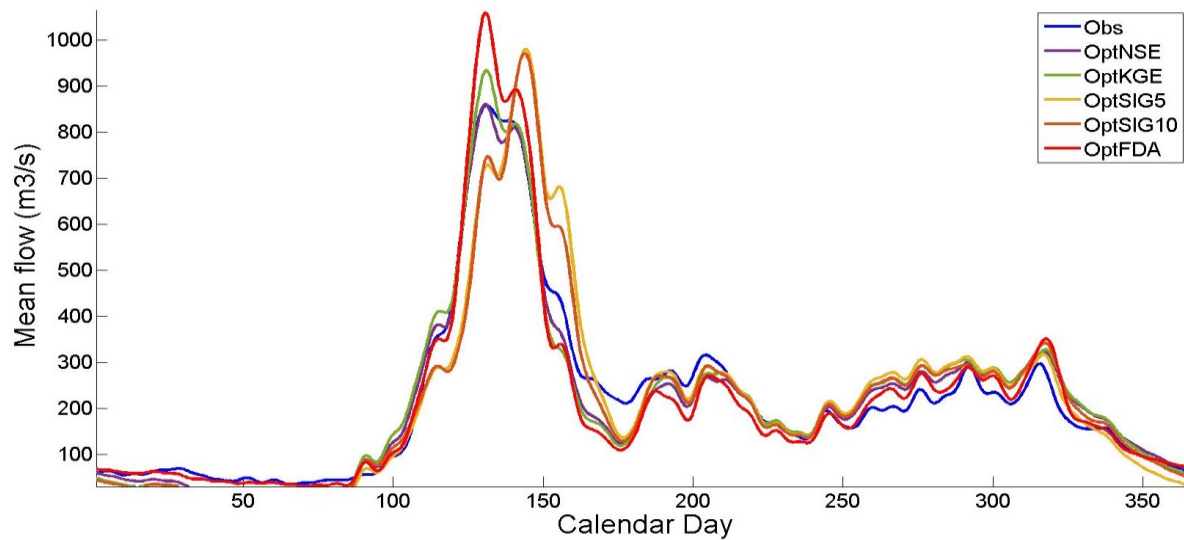


Figure 5-7 The observed and simulated mean hydrographs during the calibration period.

The performance (NSE and KGE) and consistency for calibration and evaluation periods are reported in Table 5-4. The values of the KGE components: the ratio of simulated and observed variance α , linear correlation r and bias β are also reported in Table 5-4.

Table 5-4 Performance of the mean parameter values for each calibration approach. Variability α , bias β and correlation r are the KGE components; OF-SIG10 and OF-SIG5 are the number of signatures satisfied of the 12 signatures considered in this study at the 10% and 5% acceptance threshold resp.

Calibration Period (1991-1999)							
Method/OF	NSE	KGE	r	α	β	OF-SIG10	OF-SIG5
OptNSE	0.910	0.951	0.955	0.988	0.987	10	9
OptKGE	0.886	0.933	0.946	1.039	1.004	10	8
OptSIG10	0.739	0.867	0.877	1.050	1.000	10	9
OptSIG5	0.730	0.850	0.880	1.090	1.001	10	6
OptFDA	0.850	0.848	0.943	1.141	0.986	6	6
Evaluation Period (2000-2013)							
OptNSE	0.898	0.906	0.956	1.078	1.028	10	6
OptKGE	0.863	0.852	0.948	1.132	1.044	6	5
OptSIG10	0.645	0.809	0.845	1.105	1.038	7	5
OptSIG5	0.577	0.753	0.835	1.180	1.042	6	4
OptFDA	0.833	0.797	0.946	1.195	1.021	7	6

OptFDA is the most robust parameterization in terms of both performance and consistency (i.e. similar performance and consistency during calibration and evaluation periods) according to statistics presented in Table 5-4. In contrast, the parameterization of the model with OptSIG5 and OptSIG10 are the least performant and robust parameterization. Moreover, it has the least correlated simulations with observations because of the simulated and observed peak flows are not synchronized as observed in Figure 5-8. The flow variability is adequately presented by OptNSE which has the closest value of α to unity while the OptFDA has the largest deviation of α value from unity during calibration and evaluation period. However, in Figure 5-9 which provides the interannual coefficient of variation of flows, OptFDA is representing the entire flow adequately (high and low flow) variability while the other approaches overestimate the flow variability during winter.

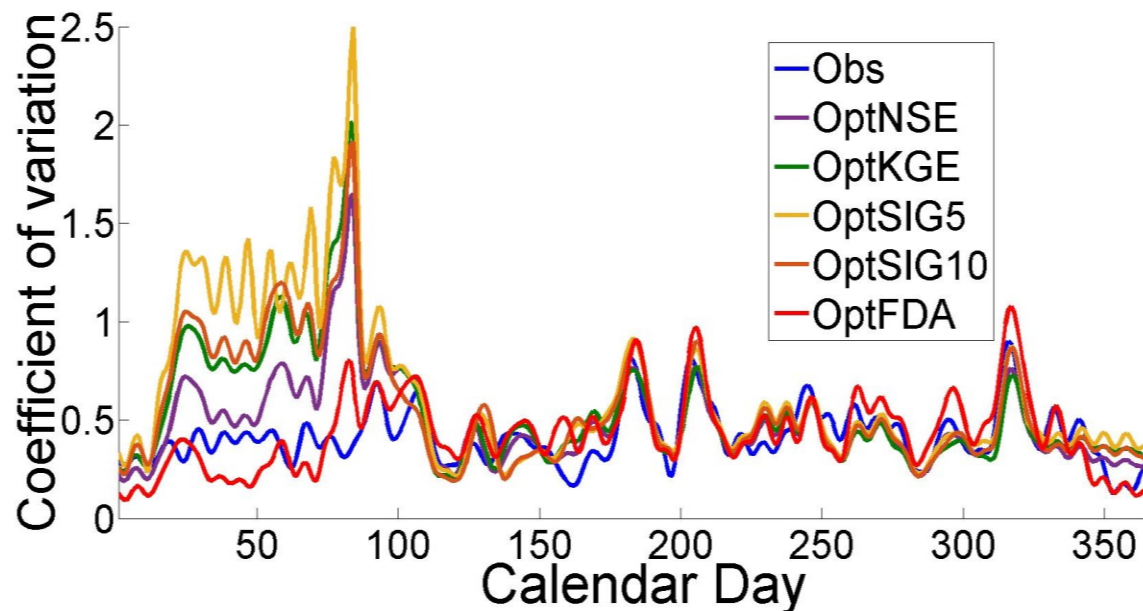


Figure 5-8 Interannual coefficient of variation of flows during the calibration period.

To assess the accuracy of the simulated flows, we applied the Functional Student test for the simulated flows with the five model parameterizations (see Figure 5-9). Winter flows simulated by the calibrated model NSE, KGE and signature-based calibration are significantly different than the observed one as well as flows during the spring flood (days between 150 and 200). The flows simulated by OptFDA can be considered similar to the observations during fall/winter and summer.

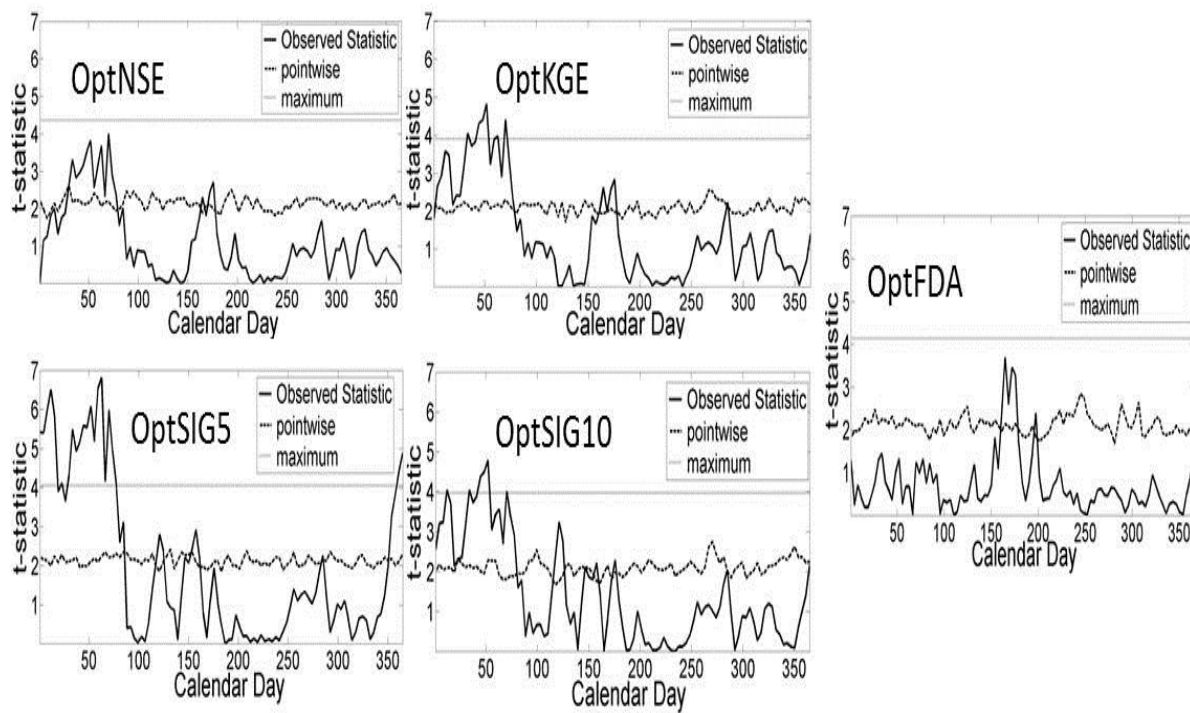


Figure 5-9 The results of the Student Test applied for the parameterized model with parameters mean values obtained with the different approaches. Days when the observed student statistic (black line) is over the p-value (dotted line) are the days when the simulation and observations are significantly different.

4.4 Uncertainty Evaluation

To assess the uncertainty, we compare the observed mean hydrograph to the minimum and maximum simulated mean hydrographs. For each parameter vector of the sets defined by each approach, the simulated flows are smoothed and a mean hydrograph is computed. Then for each day, minimum and maximum interannual mean values of simulated flows are identified and plotted against mean observed values. Figure 5-10 and 5-11 present these uncertainty bands for each calibration approach during calibration and evaluation period. Only the OptFDA minimum and maximum covers the observed value during winter flows and is narrow. During the spring flood, these bounds are larger with all approaches than the rest of the year. However, the bounds of OptNSE are less wide than of the other approaches and the bounds of OptFDA are the widest during the spring flood. Although the bounds of OptSIG5 and OptSIG10 are less wide, the simulated peak flows are not synchronous with the observed peak flows. During fall, the minimum of OptFDA and OptKGE are synchronous with the observations, while with OptNSE, OptSIG5, and OptSIG10 both bounds overestimate the fall flows. Note that the uncertainty bounds of OptFDA are closest to the fall observations during calibration and evaluation periods.

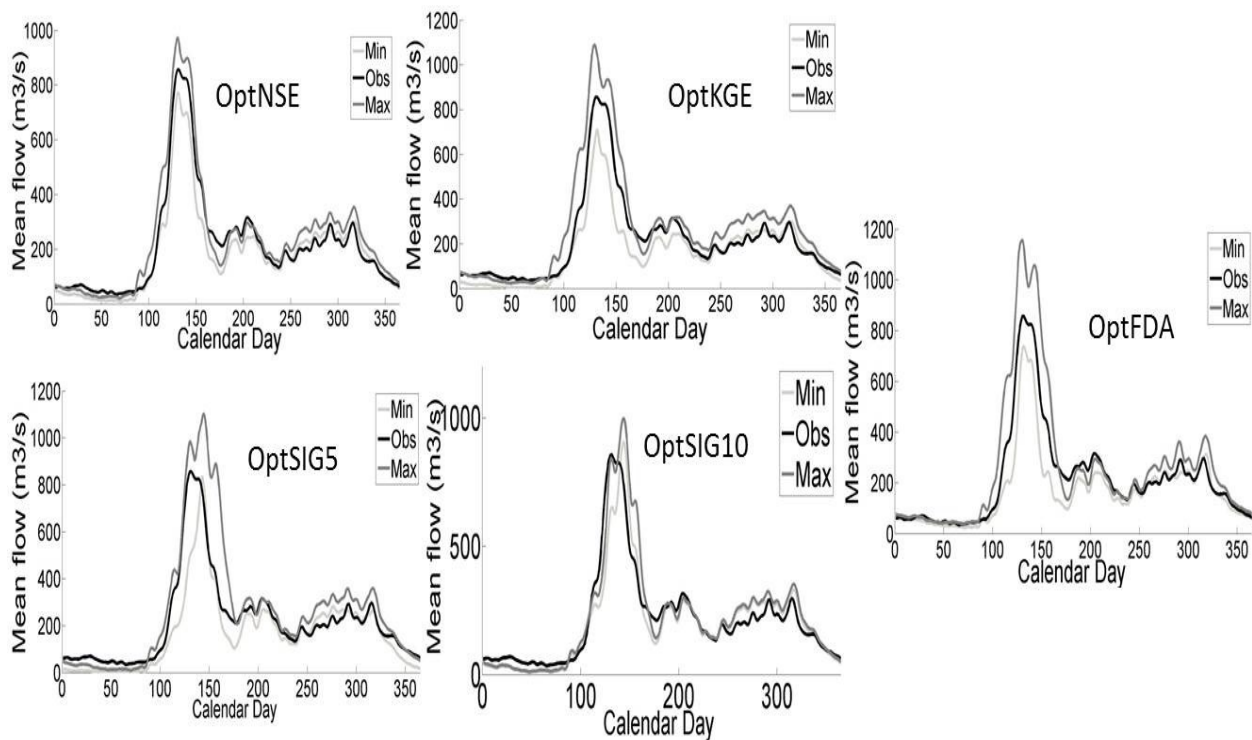


Figure 5-10 Interannual mean hydrograph with uncertainty bounds during the calibration period.

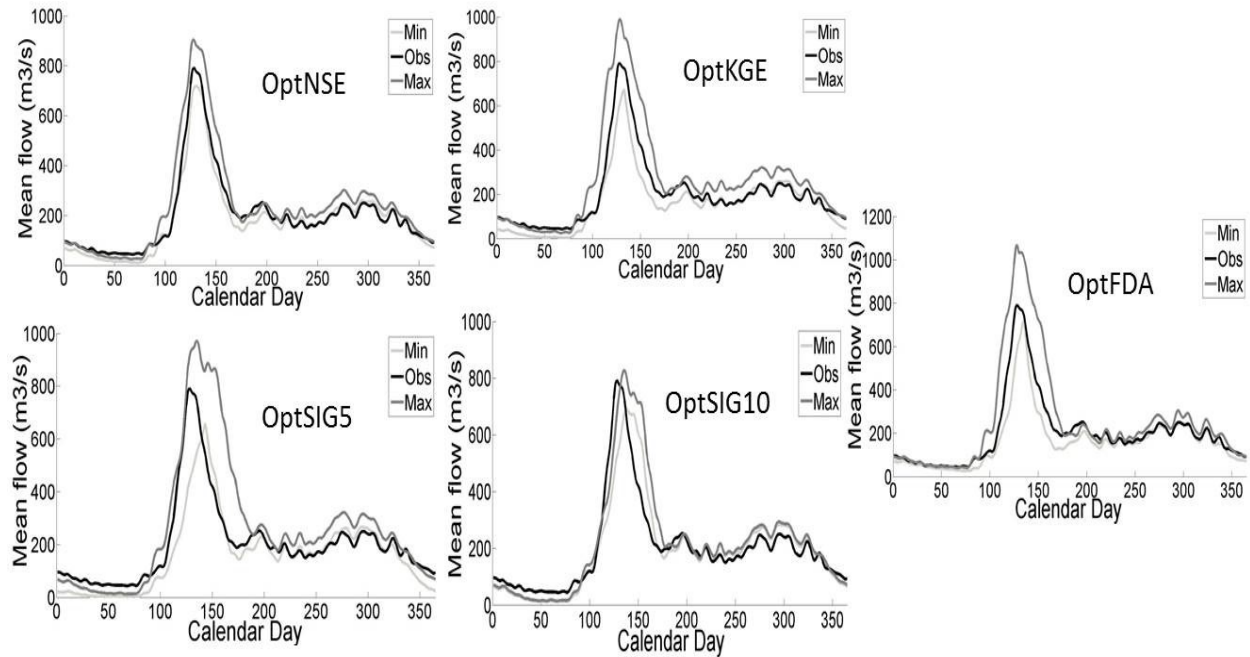


Figure 5-11 Interannual mean hydrograph with uncertainty bounds during the evaluation period.

5 Discussion

All calibration approaches show difficulty to adequately simulate the entire shape of the annual hydrograph. The study area is a snow-dominated watershed where the spring snowmelt is the important hydrological event. Consequently, the three conventional calibration approaches emphasize the high flows at the expense of low flows. In fact, the common signatures that were not satisfied with these approaches are the low flow volume and the FDC mid-segment slope. Calibration using NSE provides the most accurate spring flood simulations with less pronounced uncertainty. Meanwhile, functional-based calibration showed superiority to simulate adequately the dynamic of the entire hydrograph.

Although the consistency and performance of the signature-based calibration, this approach provided the least accurate flows in particular during spring. In this paper, a performance measure was not used with hydrological signatures which might explain the time lag between simulated and observed hydrographs. It also shows the weakness of using solely signatures in a

calibration process in terms of catching the timing. Therefore, as recommended by Van Werkhoven et al. (2009), using a combination of statistical measures and hydrological signatures is necessary to capture the different parts of the hydrograph and their timing. Using other hydrological signatures (e.g. overall runoff ratio, flow percentiles, flow frequency indices or base-flow index) could lead to different results. However, we tried to use the most widely used signatures for model calibration such as those derived from the FDC, flow distribution (mean flow and standard deviation) and flow autocorrelation (Yilmaz et al., 2008; Westerberg et al., 2011; Shafii and Tolson, 2015, Westerberg and McMillan, 2015).

The signature-based calibration provides a large number of highly consistent parameters. This confirms the findings of Shafii and Tolson (2015) that considering explicitly the number of signatures being satisfied lead to a better identification of the consistent parameters. However, these parameter sets do not transfer well to other periods. Overall, a highly consistent parameter set providing a better representation of real world system processes during calibration does not necessarily behave similarly in other periods.

Comparison between NSE and KGE performance lead to relatively different findings of Gupta et al. (2009). For this particular case study where snowmelt is dominant, NSE leads to better performance and consistency. KGE leads only to an improvement of the bias during the calibration period, and the variability is better represented by OptNSE with α much closer to the ideal value which is unity.

Snowmelt parameters, parameters controlling the vertical flow from the Upper to the Lower reservoir of the model and parameters controlling the dimension of the Lower reservoir show high sensitivity to the objective function used. Whereas, the other parameters values are slightly different depending on the objective function used. This is desirable for the calibration process for better parameter identification; but it raises issues for parameter regionalization because a small error in a parameter value causes poor overall model performance (Gupta et al., 2009).

The signature-based calibration at the 5% acceptance threshold provided the least accurate simulated flows. In this paper, the multi-objective calibration approach used is an aggregation of the different criteria. A Pareto-based multi-objective calibration approach might yield to a higher level of hydrological consistency especially in stricter levels of acceptability (Shafii and Tolson, 2015). Given the computational burden of this approach when using a large number of criteria,

calibration with a limited set of signatures that is complementary and representative of all watershed response behaviour remains challenging (Pokhrel et al., 2012).

The functional-based approach is complex compared to the conventional approaches that do not require any model or system output transformation. However, this approach is robust and not prone to the subjectivity of choosing different patterns to be studied, unlike the multi-objective approaches. It aims at conserving the entire shape of the hydrograph and provides information on when the simulations and observations do not overlap in formal mathematical settings. Moreover, this approach shows potential to achieve a high level of hydrological consistency (i.e. satisfying all hydrological signatures). The consistency and the performance can be increased by increasing the computational complexity. The sample size (number of parameter sets per iteration) is fixed here at 1000 for a reasonable computational time because of the complexity of the model. The extra step required by the functional-based calibration (i.e. smoothing the hydrographs) is less significant than the complexity of the model. This method can be used for an event or seasonal based calibration by smoothing a portion of the hydrograph only (e.g. the spring or fall floods) and performing the functional calibration on these seasonal flow series.

There are several reasons other than calibration, why model performance should be evaluated :

- 1) to provide means for evaluating improvements to the modelling approach through model structural modifications, the inclusion of additional observational information, and representation of important spatial and temporal characteristics of the watershed;
- 2) to compare current modelling efforts with previous studies; or
- 3) to evaluate the suitability of a model to a specific watershed (Krause et al., 2005).

In addition to the tools presented in this paper (mean function, Student test, the interannual coefficient of variation), the derivatives of hydrographs can be used for model performance evaluation. In fact, the smoothing required by FDA implies that the constructed curves are differentiable to a certain degree. Thereby, a number of derivatives can be computed in particular the first two, velocity and acceleration, that often have interpretation relevant to study and can be further analyzed (Levitin et al., 2007). For model evaluation, first derivatives of smoothed hydrographs would allow a better comparison of the timing of the peak flows and changing rates of flows (changing rates of the rise and falls of the hydrograph). Thereby, combining functional statistics with derivatives would provide a complete model diagnostic approach as proposed by Gupta et al. (2008) to evaluate model performance to reproduce the entire flow regime and dynamic of a watershed. Moreover, it would allow a better

understanding and distinction between the potential uncertainty sources (model form and structure, input and output uncertainty) by analyzing the change in model error under different hydrologic conditions.

6 Conclusions

This paper tackles issues around the difficulty of defining an objective function for model calibration or evaluation. Performance measures or hydrological signatures widely used are not able to provide a complete evaluation of the system output. Thus, a set of performance measures and hydrological signatures that are complementary are used in a multi-objective framework. Given the complexity and the computational burden of the multi-objective approach, the number of objective functions to be used is limited. This paper introduces a functional approach for model evaluation and calibration that allows evaluating all hydrograph features, hence a complete evaluation of the system output. In fact, it allows evaluating, in one single step, the variability of the amplitude, related to the shape of the hydrograph reflected by the peaks, the rising and the falls of the hydrograph, and the phase variability reflecting the shifts in the timings of these features.

The efficiency of this approach is compared to three calibration approaches: 1) traditional calibration with NSE, 2) a calibration using KGE as a multi-objective calibration 3) a signature-based calibration considering the number of signatures satisfied according to an acceptance threshold. The main results of this study are:

- Calibration with NSE leads to a more consistent and performant model than a calibration with KGE for a snow-dominated watershed.
- The signature-based calibration leads to performant model from the perspective of the classical statistics.
- Calibration based solely on hydrological signatures does not enable an evaluation of model adequacy to capture the timing of the hydrograph features. Signatures should be used along with a performance measure.
- The functional-based calibration provides performant sets of parameters and can achieve high consistency level (i.e. satisfying all hydrological signatures).

- The functional statistics provide full information on when the simulated and observed flows do not overlap. Hence, these statistics allow a better identification of uncertainty sources and pinpoint toward the model component (s) that needs improvement.
- For a snow-dominated watershed, the uncertainty is more pronounced during the spring flood with the four calibration approaches.
- The calibration with NSE provided the most accurate flows during the spring snowmelt but functional-based calibration showed superiority in terms of simulating adequately the entire shape of the hydrograph.
- A highly consistent model does not necessarily transfer well to conditions outside of calibration.
- Equally performant and consistent parameters are not necessarily equally transferable.

Acknowledgment

The authors acknowledge the financial contribution of Rio Tinto, NSERC and MITACS.

References

- Arsenault, R., Poulin, A., Côté, P. and Brissette, F., 2013. A comparison of stochastic optimization algorithms in hydrological model calibration. *J. hydrol. Eng.*, 1934-5584.
- Bardossy, A. and Singh, S., 2008. Robust estimation of hydrological model parameters. *hydro. Earth Syst. Sci.*, 12, 1273-1283.
- Charbonneau, R., Fortin, J. and Morin, G., 1977. The CEQUAU model: Description and examples of its use in problems related to water resource management. *Hydrological Sciences bulletin*, 22(1), p. 193-202.
- Chebana, F., Dabo-Niang, S. and Ouarda, T. B.M.J., 2012. Exploratory functional flood frequency analysis and outlier detection. *Water Resour. Res.* doi: 10.1029/2011WR011040.
- Clausen, B. and Biggs, B., 2000. Flow variables for ecological studies in temperate streams: groupings based on covariance. *J. Hydrol.*, 237(3-4), 184-197.
- Efstratiadis, A. and Koutsoyiannis, D., 2010. One decade of multi-objective calibration approaches in hydrological modelling: a review. *Hydrol. Sci. J.*, 55(1), 58–78.
- Euser, T., Winsemius, H.C., Hrachowitz, M., Fencia, F., Uhlenbrook, S. and Savenije, H.H.G., 2013. A framework to assess the realism of model structures using hydrological signatures. *Hydrol. Earth Syst. Sci.*, 17, 1893-1912.

- Gupta, H., Sorooshian, S. and Yapo, P., 1998. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resour. Res.*, 34(4), 751-763.
- Gupta, H., Wagener, T. and Liu, Y., 2008. Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrol. Process.*, 22(18), 3802-3813.
- Gupta, H., Kling, H., Yilmaz, K. K. and Martinez, G. F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.*, 377(1-2), 80-91.
- Hartmann, A., Wagener, T., Rimmer, A., Lange, J., Brielmann, H. and Weiler, M., 2013. Testing the realism of model structures to identify karst system processes using water quality and quantity signatures. *Water Resour. Res.*, 49, 3345-3358. doi:10.1002/wrcr.20229.
- Kamali, B., Mousavi, S.J. and Abbaspour, K.C., 2013. Automatic calibration of HEC-HMS using single-objective and multi-objective PSO algorithms. *Hydrol. Process.*, 27, 4028-4042.
- Krause, P., Boyle, D. and B  se, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 5, 89-97.
- Levitin, D., Nuzzo, R., Vines, B. and Ramsay, J., 2007. Introduction to Functional Data Analysis. *Canadian Psychology*, 48(3), p. 135-155.
- Liu, R.Y., Serfling, R. and Souvaine, D.L. Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications, DIMACS Series in Discrete Mathematics and Theoretical Computer Science. Volume 72 (eds. R. Liu, R. Serfling, D.L. Souvaine), American Mathematical Society, p. 159-169.
- Madsen, H., 2000. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *J. Hydrol.*, 235(3-4), 276-288.
- Masselot, P., Dabo-Niang, S., Chebana, F. and Ouarda, T. B.M.J., 2016. Streamflow forecasting using functional regression. *J. Hydrol.*, 538, 754-766.
- Moradkhani, H. and Sorooshian, S., 2009. General review of Rainfall-Runoff Modeling: Model calibration, Data Assimilation, and Uncertainty Analysis. In *Hydrological Modelling and the Water Cycle: Coupling the Atmospheric and hydrological Models*, Sorooshian, S., Hsu, K.-L., Coppola, E., Tomassetti, B., Verdecchia, M., Visconti, G. (Eds.) Springer, p. 1-24.
- Morin, G. and Paquet, P., 2007. *Morin, G., Paquet, P. (2007). Mod  le hydrologique CEQUEAU, INRS-ETE, rapport de recherche no R000926, 458p.*
- Nash, J. & Sutcliffe, J., 1970. River flow forecasting through conceptual models. *J. Hydrol.* 10 282-290.
- Pechlivanidis, I.G., Jackson, B. and McMillan, H., 2010. "The use of Entropy as a Model Diagnostic in Rainfall-Runoff Modelling." 2010 International Congress on Environmental Modelling and Software Modelling for Environment's Sake, Fifth Biennial Meeting, Ottawa, Canada: iEMSs (International Environmental Modelling and Software Society).
- Pechlivanidis, I. G., Jackson, B. M., McIntyre, N.R. and Wheeler, H. S., 2011. Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analysis

- methods in the context of recent developments in technology and applications. *Global NEST Journal*, 13(3), 193-214.
- Pfannerstill, M., Guse, B. and Fohrer, N., 2014. Smart low flow signature metrics for an improved overall performance evaluation of hydrological models. *J. Hydrol.*, 510, 447-458.
- Pokhrel, P., Yilmaz, K., and Gupta, H. V., 2012. Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures. *J. hydrol.*, 418-419, 49-60.
- Ramsay, J. and Silverman, B., 2002. *Applied Functional Data analysis: methods and Case Studies*. Second Edition éd. New York: Springer.
- Ramsay, J. and Silverman, B., 2005. *Functional Data Analysis*. 2nd ed. New York: Springer.
- Schaefli, B. and Gupta, H. V., 2007. Do Nash values have value? *Hydrol. Process.*, 21(5), 2075–2080.
- Sguera, C., Galeano, P. and Lillo, R., 2016. Functional outlier detection by a local depth with application to NOX levels. *Stoch. Environ. Res. Risk. Assess.* 30, 1115-1130.
- Shafii, M. and Tolson, B. A., 2015. Optimizing hydrological consistency by incorporating hydrological signatures into model calibration objectives. *Water Resour. Res.*, doi:10.1002/2014WR016520.
- Singh, V. P. and Woolhiser, D. A., 2002. Mathematical modeling of Watershed Hydrology. *J. hydrol. Eng.*, 7(4), 270-292.
- St-hilaire, A., Morin, G., El-Jabi, N. and Caissie, D., 2000. Water temperature modelling in a small forested stream : implication of forest canopy and soil temperature. *Can. J. Civ. Eng.*, 27(6), 1095-1108.
- Ternynck, C., Ben Alaya, M. L., Chebana, F., Dabo-Niang, S. and Ouarda, T. B.M.J., 2015. Streamflow hydrograph classification using functional data analysis. *Journal of hydrometeorology*. doi: 10.1175/JHM-D-14-0200.1.
- Tolson, B.A. and Shoemaker, C.A. 2007. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resour. Res.* doi:10.1029/2005WR004723.
- Tukey, J., 1975. Mathematics and picturing data. *Proceedings of the international Congress of Mathematics*, Vancouver, 523-531.
- Van Werkhoven, K., Wagener, T., Reed, P. and Tang, Y., 2009. Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models. *Advances in Water Resources*, 32, 1154-1169.
- Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W. and Sorooshian S., 2003. Effective and efficient algorithm for multiobjective optimization of hydrology models. *Water Resour. Res.* 39(8). doi:10.1029/2002WR001746.
- Vrugt, J.A. and Sadegh, M., 2013. Toward diagnostic model calibration and evaluation: Approximate Bayesian computation. *Water Resour. Res.*, 4335-4345. doi:10.1002/wrcr.20354.

- Wagener, T., Wheater, H. and Gupta, H., 2004. *Rainfall-Runoff Modelling in Gauged and Ungauged Catchments*. London: Imperial College Press.
- Westerberg, I. and McMillan, H., 2015. Uncertainty in hydrological signatures. *Hydrol. Earth Syst. Sci.*, 19, 3951-3968.
- Westerberg, I.K, Guerrero, J.-L., Younger, P. M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E. and Xu, C.-Y., 2011. Calibration of hydrological models using flow-duration curves. *Hydrol. Earth Syst. Sci.*, 2205-2227. doi:10.5194/hess-15-2205-2011
- Yapo, P. O., Gupta, H. V. and Sorooshian, S., 1998. Multi-objective global optimization for hydrologic models. *J. Hydrol.*, 204(1-4),83-97.
- Yilmaz, K. K., Gupta, H. V. and Wagener, T., 2008. A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. *Water Resour. Res.*, W09417, doi:10.1029/2007WR006716
- Zhang, H., Huang, G.H., Dunling, W. and Zhang, X., 2011. Multi-period calibration of a semi-distributed hydrological model based on hydroclimatic clustering. *Advances in Water Resources*, 34, 1292-1303.
- Zhang, Y., Vaze, J., Chiew, F. H.S., Teng, J. and Li, L., 2014. Predicting hydrological signatures in ungauged catchments using spatial interpolation, index model, and rainfall-runoff modelling. *J. hydrol.*, 936-948.

CHAPITRE 6 CALIBRATION MULTI-CRITERE ET PAR PROCESSUS POUR UN MODÈLE HYDROLOGIQUE CONCEPTUEL

Les travaux réalisés dans ce chapitre sont présentés sous formes d'article scientifique rédigé en anglais et dont le titre est :

A multi-criteria process-based calibration using functional data analysis to improve hydrological model realism

Samah Larabi *¹, André St-Hilaire ¹, Fateh Chebana ¹ and Marco Latraverse ²

¹INRS-ETE, Québec, Québec, Canada

²Rio Tinto Alcan, Saguenay, Québec, Canada

Cet article est accepté pour publication dans Water Resources Management Journal le 23 Août 2017. DOI : 10.1007/s11269-017-1803-6.

Les objectifs de cet article ont été discutés entre les différents auteurs. Le code du modèle hydrologique a été fourni par M. Latraverse. La méthodologie a été proposée et codé par l'auteur principal. L'analyse des résultats et la rédaction ont été effectuées par l'auteur principal en collaboration avec le professeur St-Hilaire. Une révision du contenu de l'article a été effectuée par le professeur Chebana.

Abstract

It has been argued that rainfall-runoff model calibration based solely on streamflow is not sufficient to evaluate the realism of the model to represent the internal fluxes. Therefore, model calibration has evolved to evaluating model performance using a number of hydrological signatures that link the model to the underlying processes. However, this approach uses goodness-of-fit measures, unable to describe the entire dynamic of time series, to evaluate model consistency and to simulate hydrological signatures. The present paper develops a stepwise multicriteria calibration using hydrograph partitioning and calibration criteria defined on the basis of Functional Data Analysis (FDA), a statistical tool that conserves all important features of the hydrograph by approximating times series as a single function. The aim of this approach is to improve model realism by scrutinizing model components and by evaluating its ability to reproduce the entire flow dynamic. The proposed approach is compared to a calibration against daily streamflow only. The stepwise calibration improved the estimation of the flood curve, the annual peak volume as well as the performance of the model at sites other than the calibration station.

Key words: Baseflow separation, CEQUEAU, Functional data Analysis, Multicriteria calibration, Process-based calibration.

1. Introduction

Conceptual rainfall-runoff models often, use storage components to represent fluxes such as infiltration, percolation, runoff or drainage (Wagner et al., 2004). Typically, these models are calibrated against streamflow observations using a goodness-of-fit metric (e.g. Nash-Sutcliffe Efficiency (NSE), Root Mean Squared Error (RMSE)).

Arguably, such metrics cannot provide information about the model physical realism (Krause et al., 2005; Shafii and Gupta; 2007, Gupta et al., 2008, 2009). Gupta et al (2008) advocate the diagnostic approach to evaluate model performance by using a number of hydrological signatures that link the model to the underlying processes. Different studies showed the efficiency of this approach and emphasized that model evaluation solely on its output time series is not sufficient to evaluate model robustness or consistency (Yilmaz et al., 2008; Hingray et al., 2010; Clark et al., 2011; Pokhrel et al., 2012; Euser et al., 2013; Shafii and Tolson, 2015; Rakovec et al., 2016). Other strategies are used to improve calibration quality by adjusting parameters on internal variables such as measurement of evapotranspiration, baseflow, soil moisture, or snow cover or hydrograph partitioning (Cao et al., 2006; Finger et al., 2011; He et al., 2015; Shafii et al., 2017).

Streamflow at any point in time is composed of water going through different pathways and the partitioning between surface runoff and baseflow affects the streamflow dynamics (Clack et al., 2011; Shafii et al., 2017). Recently, Shafii et al. (2017) argue that standard hydrological signatures do not necessarily manifest hydrological partitioning and the relation between different processes. Moreover, the majority of the signature-based calibration approaches use goodness-of-fit metrics against hydrological signatures or internal variables (e.g. Yilmaz et al., 2008; Pokhrel et al., 2012; Pfannerstill, et al., 2014). These metrics are associated with a loss of valuable information, in particular temporal variability (Gupta et al., 2008). Therefore, the use of multiple objective functions is a valuable alternative that allows to account for low and high values separately.

For instance, Hay et al (2006) propose a stepwise calibration by calibrating model components against solar radiation (SR), potential evapotranspiration (PET), water balance and daily runoff components. To compare daily runoff components, the authors used three different objective functions based on a normalized root mean square error to be able to evaluate mean daily flows,

low flows and peak flows; and used a sum of the absolute residuals of observed and simulated logarithms of monthly SR and PET to give a proportional measure of error for low and high values. Labarthe et al. (2014) used a pilot point approach for a stepwise calibration applied to a coupled hydrological-hydrogeological model. This method uses a weighted sum of the square errors to compare between the observed and simulated discharge. These types of metrics, even with a logarithm transformation of data, do not consider the changes in bias, variance or skew in model errors under different hydrological conditions (Beven, 2006).

To consider these changes, there is a need for metrics such as Functional Data Analysis (FDA) that conserve the shape and variability of time series to represent these changes. The advantage of FDA is the ability to describe both sources of variability in time series amplitude (rise and fall) and phase (timing and occurrence) (Levitin et al., 2007).

So far, there are few applications of FDA in hydrology. It was used for hydrographs exploratory analysis and outlier detection by Chebana et al., (2012), for streamflow hydrographs classification by Ternynck et al. (2015) and for streamflow forecasting by Masselot et al., (2016). This paper is the first attempt to use FDA to construct hydrological model calibration and evaluation criteria.

The objective of this paper is to develop a multi-criteria process-based calibration using criteria based on FDA, with a focus on certain model components. The proposed method aims at improving model consistency by using hydrograph decomposition into spring flood curve, surface runoff and baseflow based on the same concept of He et al. (2015) and Shafii et al., (2017). The routing process from the upper reaches to the outlet is also evaluated. FDA-based calibration criteria can capture the entire dynamic and variability of these hydrological features (i.e. during high and low flow periods). Our hypothesis is that scrutinizing model components using FDA will improve model consistency and performance during high and low flows periods as well as at a basin scale. To verify this hypothesis the model is evaluated on a number of sites not used for calibration.

2. Hydrological CEQUEAU model and study area

2.1 Hydrological CEQUEAU model

CEQUEAU is a deterministic, semi-distributed and conceptual hydrological model (Charbonneau et al., 1977; Morin and Paquet, 2007; St-Hilaire et al., 2000). The model takes as input physiographic and land use data, daily total precipitation, minimum and maximum daily temperature. CEQUEAU represents the soil as interconnected reservoirs. The flow is distributed vertically with the production function (Figure 6-1) and simulates the components of the hydrological water-balance. Downstream routing of water is calculated with a transfer function. Details about the model conceptualization can be found in Morin and Paquet, (2007) and St-Hilaire et al., (2000). The key model parameters are summarized in Table 6-1.

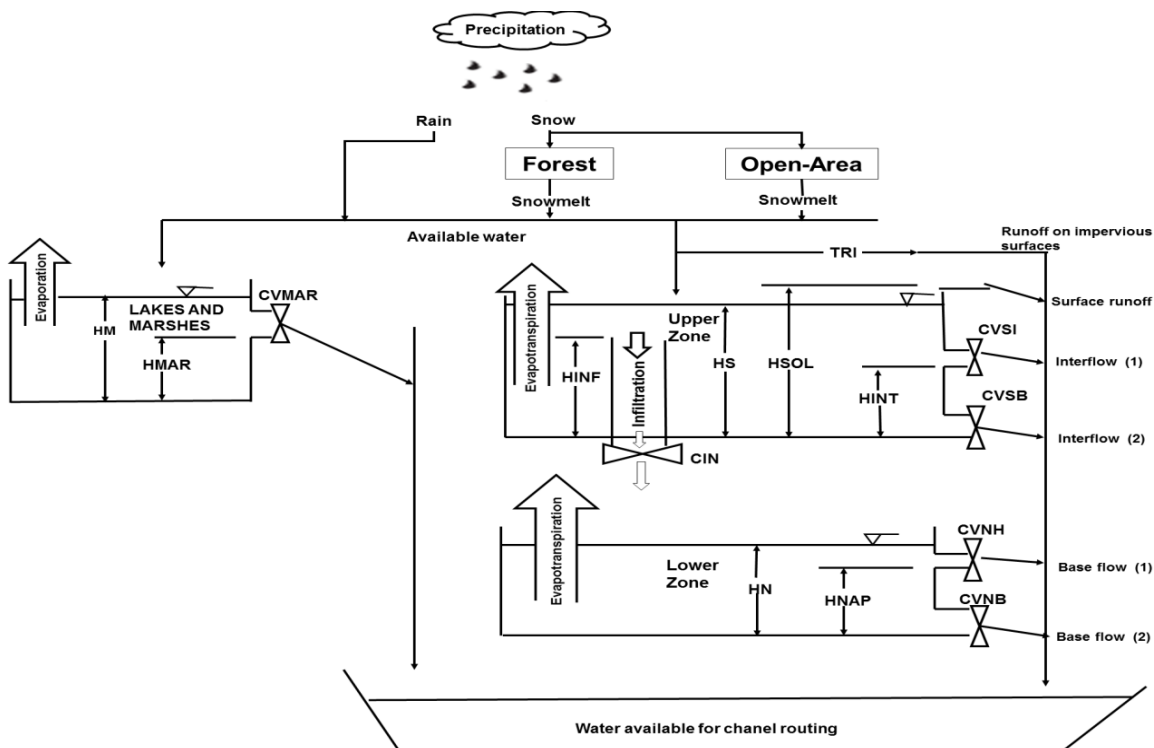


Figure 6-1 Flow chart of the production function of CEQUEAU model (adapted from Morin et Paquet 1995, St-Hilaire et al., 2000).

Table 6-1 Key parameters of the CEQUEAU model with calibrated values against overall performance and using the proposed approach.

Parameter	Description	Calibrated values against streamflow	Calibrated values with the proposed approach
STRNE	Rain/ Snow transformation threshold (°C)	-1.41	-1.57
TFC	Melting rate in forest cover (mm/°C/day)	3.21	3.24
TFD	Melting rate in open area (mm/°C/day)	6.95	8.40
TSC	Melting threshold in forest cover (°C)	1.23	1.42
TSD	Melting threshold in open area (°C)	-0.77	-1.25
TTD	Cold content coefficient for the snowpack	1.84	1.43
TTS	Priming threshold of the snowpack(°C)	-2.57	-3.51
CIN	Percolation coefficient from Upper Zone to Lower Zone	0.44	0.52
CVMAR	Lakes and Marches drainage coefficient	0.56	0.30
CVNH	Lower-zone upper drainage coefficient	0.52	0.41
CVSB	Upper-zone lower drainage coefficient	0.02	0.02
CVSI	Upper-zone intermediate drainage coefficient	0.51	0.37
XINFMA	Daily maximal infiltration (mm/day)	14	24
HINF	Percolation threshold from the upper to the Lower Zone (mm)	82	82
HINT	Percolation threshold from the upper to the Lower Zone (mm)	82	79
HMAR	Lakes and Marshes threshold (mm)	354	335
HNAP	Lower-zone upper threshold (mm)	113	130
HPOT	Threshold over which water is evaporated at the potential rate (mm)	41	38
HSOL	Upper Zone capacity (mm)	173	206
HRIMP	Precipitation threshold for direct runoff on impervious surface surfaces (mm)	5.03	3.37
EVNAP	Percentage of total daily evapotranspiration taken from Lower Zone	0.23	0.23
Zn	Concentration time for the basin	5	2
EXXKT	Fitting parameter for the routing parameter	0.80	0.02

2.2 Study area

The model is calibrated on the spring flood-dominated lac St-Jean drainage basin located in the Province of Québec, Canada. This basin extends over an area of 73 800 km² with 90% of forest cover. The meteorological input data are available from 1985 to 2013. In this basin, flow observations are available for 9 stations (Table 6-2; Figure 6-2). Statistics given in Table 6-2 are computed for the model calibration period between 2002 and 2009. Two hydrometric stations were used for calibration: one in the outlet denoted (LSJ) and Ashuapmushuan (ASHA) upstream. Because of the missing data elsewhere, only these two stations were used to calibrate the model and the others to evaluate the robustness of the calibrated model within the watershed.

Table 6-2 Hydrometric stations in the Lac St-Jean drainage basin.

Station	Notation	Annual Mean Precipitation(mm)	Annual mean Maximum flow(m3/s)	Annual Mean flow (m3/s)
Lac St-Jean (outlet)	LSJ	1006	4139	914
Ashapshuam-Aval	ASHA	757	1353	307
Ashapshuam- Amont	ASHAM	1046	1022	304
Mestabetchouane	METAB	1022	288	52
Mistassibi 3	MISBI3	988	382	80
Mistassini	MISNI	931	1140	222
Mistassini 2	MISNI2	976	546	60
Petite Peribonka	PPER	837	738	22
Mistassibi	MISBI	960	1000	203

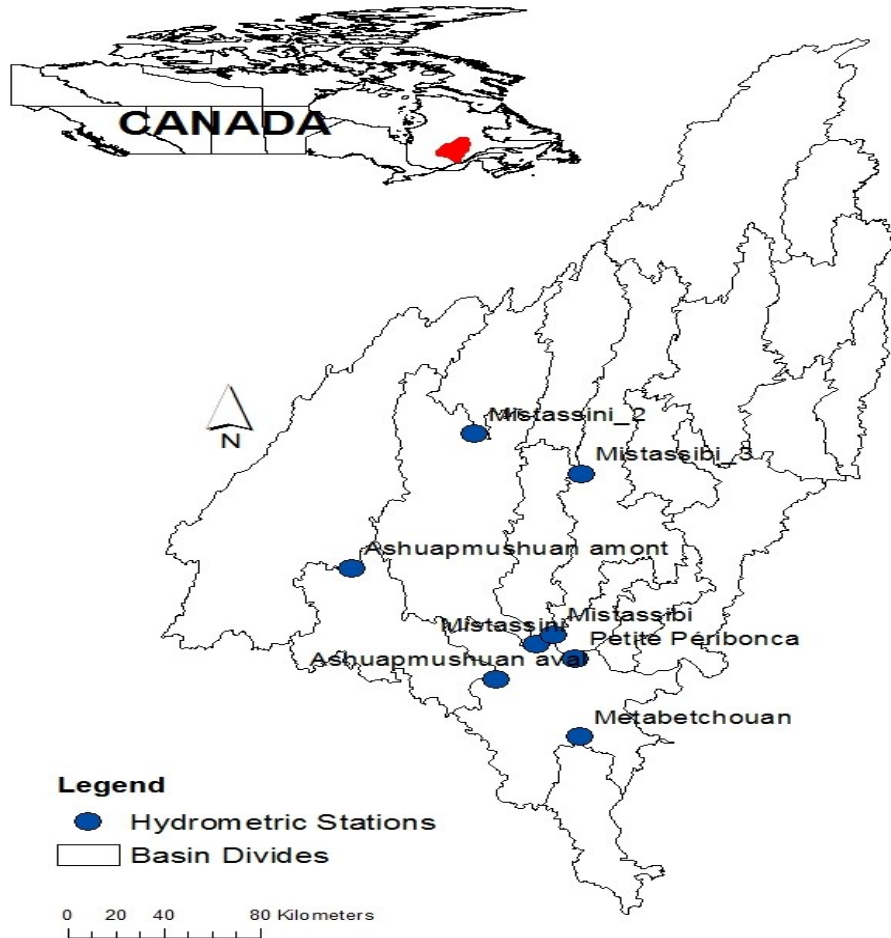


Figure 6-2 The lac St-Jean drainage basin.

3. Methodology

Each component of the hydrological model conceptualizes water storage/flow associated with specific hydrograph components. Therefore, some parts of the hydrograph are not useful to constrain some model parameters (e.g. summer low flows do not contain any information related to snowmelt parameters). When using the entire hydrograph for calibration, the model tends to compensate between the water volumes simulated for the conceptualized reservoirs representing the unsaturated and saturated zones. Consequently, the resulting calibrated model does not necessarily simulate adequately the hydrological processes that are conceptualized in the model. For these reasons, each component of the hydrological model is adjusted separately using an

adapted criterion. All components are adjusted until the global performance of the model ceases to improve. The parameters are adjusted in the following order: 1) snow-related parameters are adjusted on the spring flood curve; 2) Parameters controlling the vertical routing of water (henceforth called physical parameters) through the process of infiltration, percolation and evapotranspiration are adjusted based on a hydrograph separation into direct runoff and baseflow; 3) the water routing parameters are adjusted based on the joint performance of the model at two stations (upstream and downstream). Snow parameters and the physical parameters are adjusted on the station in the outlet (LSJ) and the routing parameters on both stations LSJ and ASHA. The accuracy of model simulations elsewhere in the watershed is checked during the validation of the calibrated model at other stations. Times series of streamflows, direct runoff and baseflow are approximated by temporal functions using FDA. The model is calibrated on these new functional observations (smoothed curves) and the calibration criteria are defined accordingly. Subsection 3.1 presents briefly how the functional data analysis is used in the context of the current study. Subsection 3.2 provides the global performance criterion. Subsections 3.3, 3.4 and 3.5 give the methods to calibrate each model component.

3.1 Smoothing data with functional data analysis

Functional data analysis (FDA) is a mathematical tool to study observed data within a space or temporal framework as a single entity represented by a function. The key concept of FDA is that observations y_j at discrete points in time $\{t_j\}$ are observed on a continuous process $x(t_j)$ (Ramsay and Ramsey, 2002). Therefore, the first step in FDA is smoothing data through a basis expansion where each continuous function is generally defined as (Ramsay & Silverman, 2002):

$$x(t) = \sum_{k=1}^K c_k \phi_k(t), \quad (6.1)$$

where ϕ is the basis function and c_k are the coefficients of expansion estimated from data.

In the current study, all data (flow, direct runoff and baseflow) show some periodicity. Therefore a Fourier basis functions is used:

$$x(t) = c_1 + c_2 \sin \omega t + c_3 \cos \omega t + c_4 \sin 2\omega t + \dots + c_K \cos K\omega t \quad (6.2)$$

where $\omega = \frac{2\pi}{T}$, T is the period in our case $T=365$, K is the number of the basis functions.

Each annual data series $\mathbf{Y}_i \{y_i(t_j), j=1, \dots, T\}$ with $i=1, \dots, N$ where N is the number of years of observation, is converted by Eq. 6.3 into one temporal function on a continuous period $[1, T]$ where $T=365$:

$$y_i(t) = \hat{x}_i(t) = \sum_{k=1}^K \hat{c}_{ik} \phi_k(t) = \hat{\mathbf{c}}_i^T \boldsymbol{\phi}(t) \quad (6.3)$$

where : $\hat{\mathbf{c}}_i$ is the estimated basis coefficient vector of length K and $\boldsymbol{\phi}$ is a K -vector of basis functions.

The basis coefficients can be estimated with traditional least squared error or a penalized least square error (for more details, see Ramsay and Silverman (2002) or Chebana et al. (2012) in hydrological context). The number of functions of the basis K is fixed depending on the observed data so that the curves are smoothed to a certain degree with a low bias. The same number K is used to transform the model simulations at each iteration to curves with equation 6.2. The model is calibrated on the transformed and is evaluated on its performance to reproduce the observed smoothed curves at each year. The curves are fitted on the entire annual hydrograph, and on a portion of the hydrograph associated with a specific hydrological process: baseflow or direct runoff.

3.2 Global performance of the model

The global performance of the calibrated model is evaluated according to its ability to reproduce the annual hydrographs properly in the basin. To this end, the objective function is defined as the sum of the distances between the simulated and observed annual hydrographs as well as their respective first derivatives (see criterion noted Glob_OF in Table 6-3).

Table 6-3 Calibration criteria to evaluate model components considering data (flows, runoff and baseflow) as functional data.

Criterion	Equation	Variables
Overall Performance (Glob_OF)	$S * \sqrt{\sum_{i=1}^N \left\{ \int_{t=1}^{T=365} x_{i,sim}(t) - x_{i,obs}(t) ^2 + \int_{t=1}^{T=365} x'_{i,sim}(t) - x'_{i,obs}(t) ^2 \right\}}$	$x_{i,sim}$ and $x_{i,obs}$: simulated and observed smoothed hydrographs at the i^{th} year, $x'_{i,sim}$ and $x'_{i,obs}$: are their derivatives; N is the number of years of observations; S is the average of the Student statistic.
Snowmelt (Snow_OF)	$S * \sqrt{\sum_{i=1}^N \left\{ \int_{t=81}^{T=181} x_{i,sim}(t) - x_{i,obs}(t) ^2 + \int_{t=1}^{T=365} x'_{i,sim}(t) - x'_{i,obs}(t) ^2 \right\}}$	$x_{i,sim}$ and $x_{i,obs}$: simulated and observed flood curves of the shifted hydrographs at the i^{th} year, $x_{i,sim}$ and $x_{i,obs}$: are the derivatives of the non-shifted hydrographs of the i^{th} year; N is the number of years of observations, S is the average of the Student statistic
Runoff (Runoff_OF)	$\sqrt{\sum_{i=1}^N \int_{t=1}^{T=365} f_{sim,i}(t) - f_{obs,i}(t) ^2 dt + \frac{\sum_{t=1}^{T=365*N} (Q_{isim}(t) - Q_{iobs}(t))}{\sum_{t=1}^{T=365*N} P_i}}$	$f_{sim,i}$ and $f_{obs,i}$ are the simulated and observed smoothed runoff curves at the i^{th} year, Q_{sim} and Q_{obs} are the simulated and observed time series, P is precipitation
Baseflow (Baseflow_OF)	$\sqrt{\sum_{i=1}^N \int_{t=1}^{T=365} b_{sim,i}(t) - b_{obs,i}(t) ^2 dt + \frac{\sum_{t=1}^{T=365*N} (Q_{isim}(t) - Q_{iobs}(t))}{\sum_{t=1}^{T=365*N} P_i}}$	$b_{sim,i}$ and $b_{obs,i}$ are the simulated and observed smoothed baseflow curves at the i^{th} year, Q_{sim} and Q_{obs} are the simulated and observed time series, P is precipitation
Routing (Routing_OF)	$ \varepsilon_{downstream} - \varepsilon_{upstream} \text{ where } \varepsilon = \sum_{i=1}^N 1 - \frac{\int_{t=1}^{T=365} x_{i,sim}}{\int_{t=1}^{T=365} x_{i,obs}}$	$x_{i,sim}$ and $x_{i,obs}$: simulated and observed smoothed hydrographs at the i^{th} year.

Using the first derivatives allows extracting information about the timing of the hydrological events, as illustrated in Figure 6-3 (e.g. a positive first derivative indicates increasing flow, the onset of positive derivative values indicates the onset of the rising limb of the hydrograph, etc.) Drought or low flow spells are represented by the periods when the first derivative is almost flat.

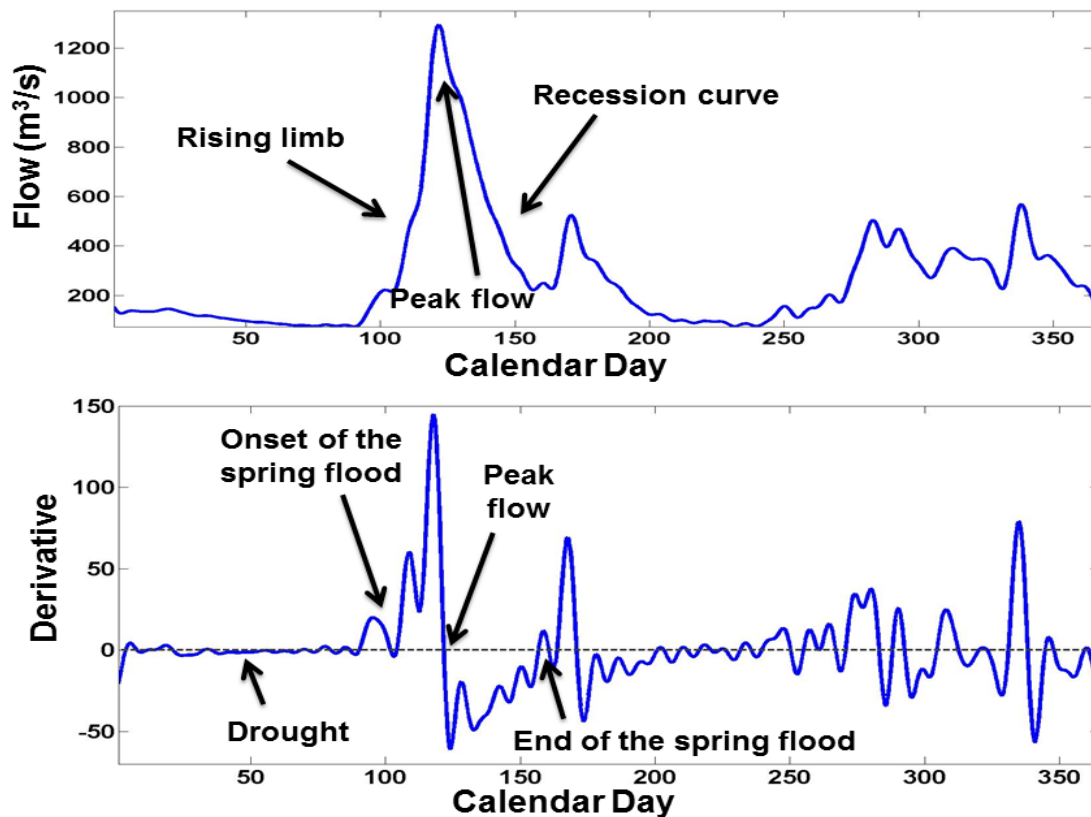


Figure 6-3 Interpretation of the first derivative of a smoothed hydrograph. Upper figure presents the hydrograph features. Bottom figure provide explanation of extracting timing of these features through a change of slope of the first derivative of the smoothed hydrograph.

GLOB_OF criterion is weighted by the average of the functional Student test statistic (Ramsay and Silverman, 2005). This test verifies the hypothesis that the model-simulated hydrographs are identical to the observed hydrographs for each time step. Weighting the objective function, to be minimized by this criterion will penalize parameter sets unable to provide accurate flows for each time step.

3.3 Adjusting snow parameters

Snow parameters control the timing to the peak flow, the peak flow magnitude and the flood volume. Therefore, they are adjusted essentially on the portion of the hydrograph representing the spring flood. Annual flow series are converted into smoothed hydrographs as illustrated in subsection 3.1. Spring flood does not occur each year at the same time (Figure 6-4). To extract volume of the spring flood automatically, smoothed hydrographs are temporally shifted to a pre-determined, common date so that the timing of spring flood is fixed beforehand. Shifting functions consists of a temporally translation into a reference interval using a time-warping function (Ramsay & Silverman, 2002). In our current study, the reference interval is the time interval of the interannual average of spring floods. In other words, annual hydrographs are shifted according to the average annual hydrograph as illustrated in Figure 6-4. Thus the volume is extracted from 21 March ($t=81$) to 30 June ($t=181$).

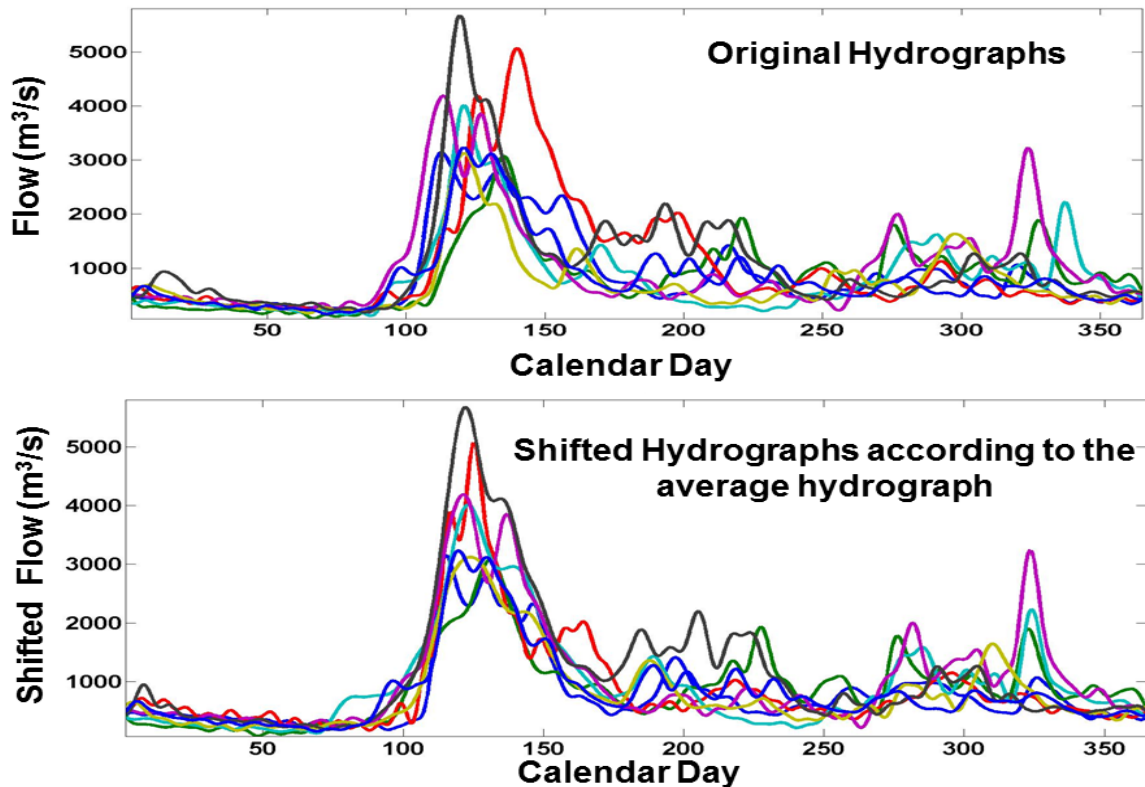


Figure 6-4 Example of smoothed hydrographs. In the upper figure are the original smoothed hydrographs. In the bottom figure are the shifted hydrographs according to the average hydrograph.

To compare the timing of the onset of the spring flood as well as the timing of the peak flow, the first derivative of the non-shifted hydrograph is used as illustrated in Figure 6-3. The criterion defined to adjust snow parameters (noted Snow_OF in Table 6-3) consists of minimizing the distance between the functions (simulated and observed) representing the extracted flood curves and the distance between the first derivatives of the non-shifted hydrographs (see Table 6-3). The criterion is weighted by the average Student statistic as per overall performance criterion.

3.4 Adjusting physical parameters

Physical parameters of the production function are adjusted on the basis of hydrograph separation into baseflow and direct runoff. In the literature, two approaches for hydrograph separation exist: methods based on hydrological events that are graphically based, or recursive methods. In our case, a recursive method is best suited for calibration purposes. Eckhardt (2005) proposed a recursive digital filter with two parameters. The baseflow is defined as follows:

$$b_{t+1} = \frac{(1-BFI_{max})\alpha b_t + (1-\alpha)BFI_{max}y_{t+1}}{1-\alpha BFI_{max}} \quad (6.3)$$

Where b is the baseflow at time step t , y is the flow at time step t , BFI_{max} is the baseflow index, α recession constant, if $b_{t+1} < y_{t+1}$; else: $b_{t+1} = y_{t+1}$.

The filter parameters BFI_{max} and α are fixed according to the default values recommended by Eckhardt (2005) because measurements of baseflow and direct runoff are not available. When these measurements are available, it is strongly recommended to adjust the filter parameters prior to calibration (Lim, et al., 2010). In this paper, observed runoff and baseflow refer to runoff and baseflow obtained through separation of the observed hydrograph. To evaluate water balance, runoff ratio is added to the objective function. Indeed, runoff ratio is widely used as a hydrological signature to evaluate the water balance between the input (precipitation) and output (flow) (Yilmaz et al, 2008; Shafii and Tolson, 2015; Westerberg and Mcmillan, 2015). In our case this criterion would reflect adequacy of the evapotranspiration parameters, as no measurement of evapotranspiration is available. Calibration of physical parameters of the production function is done in two steps. Parameters of the upper reservoir level, related to direct runoff, are adjusted first. Thus, the objective function (noted Runoff_OF in Table 6-3) compares between the simulated and observed runoff curves.

When the level in the upper reservoir reaches a threshold, it supplies the lower reservoir through infiltration. The parameters of the lower reservoir are adjusted in the second step by comparing observed and simulated baseflow smoothed functions by the criterion noted Baseflow_OF in Table 6-3.

3.5 Adjusting routing parameters

To ensure that the calibrated model simulates adequately the flow routing, an additional station located upstream of the outlet hydrometric station is used to adjust the two routing parameters. It is hypothesized that if simulated flows are synchronized with observations upstream as well as downstream, the routing parameters are likely adequately calibrated. Under the assumption that the parameters at the outlet are perfectly calibrated, poor performance of the model upstream is mainly due to the poor performance of the routing transfer function. Therefore, the objective function minimizes the error between the performance of the model at the outlet and its performance upstream as described by the criterion noted Routing_OF in Table 6-3

4. Results and discussion

4.1 Performance at the outlet station used for calibration

The calibrated model with the proposed methodology (noted OptProcess) is compared to a calibration based solely on the overall performance of the model (noted OptGlob) at the outlet station using. NSE values of the calibrated model at hydrometric stations available in the Lac St-Jean drainage basin (Table 6-4). The process-based calibration leads to a better estimate of the high flows with NSE of 0.86 against 0.71 with OptGlob at the outlet station (LSJ) used for calibration. OptGlob overestimates high flows with an annual peak bias of 43% against 11% for OptProcess.

Table 6-4 Performance of the calibrated model (%).

Station	NSE		Bias	
	OptGlob	OptProcess	OptGlob	OptProcess
LSJ	71	86	1	2
ASHA	44	81	2	2
ASHAM	-8	52	23	23
METAB	10	39	20	20
MISBI3	9	60	1	0
MISNI	65	79	14	14
MISNI2	12	63	3	3
PPER	27	63	10	10
MISBI	45	72	0.3	1

The results of the functional Student test are presented in Figure 6-5. Days when the Student statistic (represented by the blue line) is over the p-value (the dotted red line) are the days when the simulated and observed flows are significantly different. The calibrated model with both approaches fails to reproduce accurate flows during winter (days 30 to 80). However, the stepwise process-based calibration improves the simulation of low winter flows with a bias of 16% against 21% for the overall performance calibration.

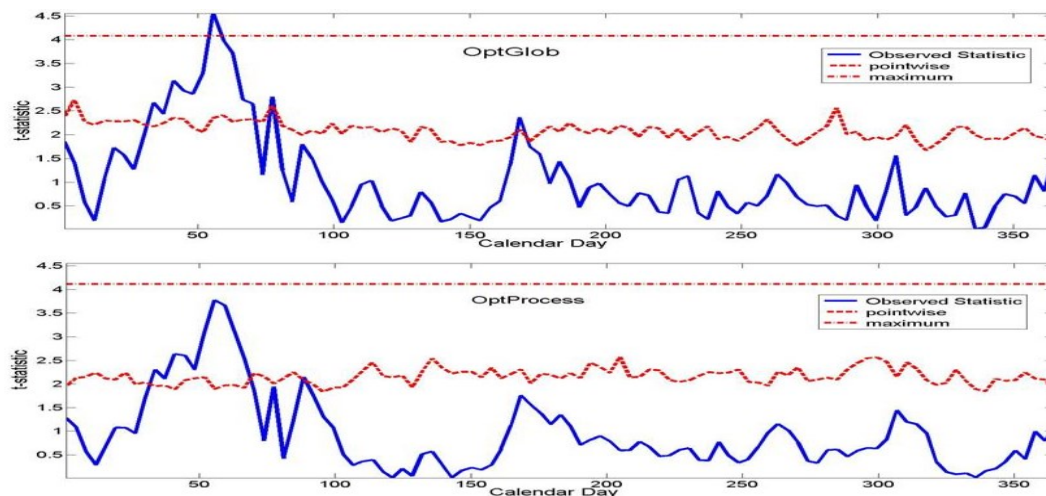


Figure 6-5 Evaluation of model accuracy using the functional Student test. Upper figure compares the observed hydrographs with simulated hydrographs with OptGlob and the bottom figure with OptProcess. Days with Student statistic (blue line) over p-value (dotted red lines) are the days when simulations and observations are significantly different.

During winter, the observed flows at the outlet show a high variability that the model is unable to reproduce, as illustrated in Figure 6-6 presenting daily flow series for selected years. This shortcoming might be due to the fact the baseflow is modelled by a linear reservoir in CEQUEAU. Clark et al. (2011) showed that baseflow models based on either linear reservoirs or single non-linear reservoir are unable to represent baseflow dynamics and recession characteristics and points toward the fact that baseflow representation needs further attention.

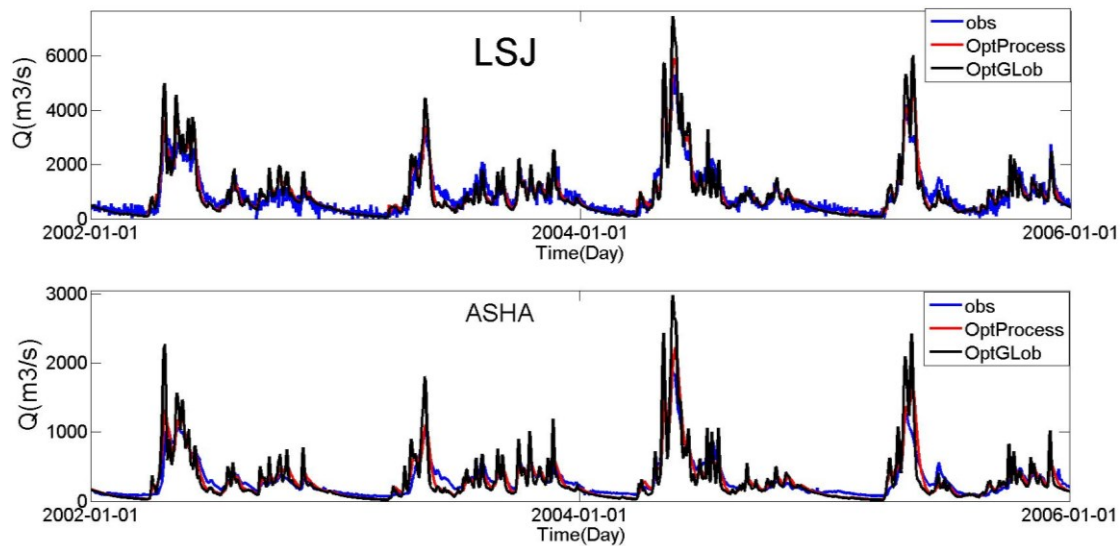


Figure 6-6 Observed and simulated flow series in the outlet (upper figure) and at the ASHA station (bottom figure) for selected years.

4.2 Basin-wide performance

The calibrated model with the process-based calibration has similar performance at the second station ASHA as at the outlet station LSJ. The NSE is over 0.80 with relatively low annual peak bias (18%) compared to OptGlob (73%) and winter low flows are underestimated (see Figure 6-6). Performance of the model parameterized with OptGlob not only drops at this station with NSE value of 0.44 but at all stations of the watershed with mean NSE value around 0.23 against 0.61 with OptProcess. The poor performance of OptGlob at ASHA is due to the overestimation of the high flows as illustrated in Figure 6-6. At the other stations, it can also be due to the time

lag between simulated and observed flow. OptProcess shows a good basin-wide performance except for two stations METAB and ASHAM. This illustrates that the stepwise process-based calibration is more physically based and achieves higher performance than the overall calibration (He et al., 2015). It also emphasizes that including criteria for model calibration other than streamflow improves the realism and representation of processes by the model (e.g. McMillan et al., 2011; Euser et al., 2013; Rakoveck et al., 2016; Shafii et al., 2017). In this study, the evaluation of the model on different stations was intended to show that scrutinizing model components on the underlying processes improves model performance at a basin scale.

The low performance of the model at the METAB and ASHAM stations can be explained by the different flow regime at these drainage areas, which are characterized by a higher slope than the other stations. In fact, the outlet station is usually not representative of the flow regime of large watershed because of spatial variability and heterogeneity. Thus, model calibration only at this station results in poor performance of the model at other stations. To remedy this issue, other studies recommend simultaneous multisite calibration (Zhang et al., 2008; Niraula et al., 2012; Wang et al., 2012; Leta et al., 2017). In this study, at stations (ASHA, MISNI, and MISBI), where the annual mean maximum is $> 1000\text{m}^3/\text{s}$, the model has an acceptable performance in terms of catching the timing and amplitude of the peak flow ($\text{NSE} > 0.72$).

4.3 The realism of the simulated internal processes

The model parameterized with OptGlob and OptPocess provide similar snowpack simulations with a realistic periodicity and melting faster in open area than in forested area. The water volumes in the Lower Zone simulated by OptProcess are higher than the simulated by OptGlob. Water volumes simulated in this lower zone are supposed to represent sustained groundwater inflow and be the main source of discharge during dryer periods. This explains the higher bias on low flows with OptGlob. Figures presenting the snowpack and water volume in the Upper and Lower Zone at the outlet station and ASHA station are given in supplementary materials.

The seasonality of evapotranspiration (high in summer and very low in winter) simulated by both calibrations is realistic, with peak daily evaporation rates around 6 mm and annual evapotranspiration varying from 429 to 495mm/year at the LSJ station and 402 to 473mm/year at the ASHA station during calibration period (2002-2009). Similar values were obtained by

Tremblay et al. (2014) in Eastmain River watershed (located in the province of Québec, Canada), with annual evapotranspiration rates varied for forest area from 397 to 528 mm/year between 2008 and 2010.

In Table 6-1, significant differences are noticed for snow parameters, threshold parameters controlling the vertical distribution of the water volume within the conceptualized reservoirs and routing parameters. This emphasizes the sensitivity of model parameters toward the objective function (Gupta et al., 2009). The sensitivity of model parameters toward the objective is due to the fact that each objective function penalizes heavily on certain hydrograph features and each parameter is sensitive to these features.

Overestimation of high flows with OptGlob is due to the low performance of the model to adequately drain the flow downstream. Indeed, the drainage with OptGlob is slow with a concentration time for the basin (ZN) set to 5 days, which explains the poor basin-wide performance. While with OptProcess ZN is only 2 days which the mean time observed in the watershed. In fact, the routing process in CEQUEAU is related to the longest path (NPJO) between the outlet and the upstream of the watershed and concentration time. The flow is routed downstream by considering a grid covering the watershed. Given that NPJO is 49 squares in the lac St-Jean, fixing ZN to 4.77 days means that the simulated water upstream will be in 10 squares downstream at the end of the day. Whereas fixing ZN to 1.83 means that the water will be in 26 squares downstream. Therefore, with OptGlob more water volume is retained upstream in contrast with OptProcess. Calibration of the model against only the observed flow and under imposed constraints on certain parameters to limit parameter compensation is unable to capture the watershed response. In fact, the Lower-Zone lower drainage coefficient CVNB (see Figure 6.1) is set to 0 and the upper bound of EVNAP (fraction of daily evapotranspiration taken from the Lower Zone) is set to 0.25. The aim of these constraining conditions is to avoid unrealistic behaviours such as emptying the Lower Zone completely through these parameters. This highlights issues around parameters taking unrealistic values to compensate for errors in other model components or errors in model structures (Beven, 2006; Clark et al., 2011).

In this study, model components were evaluated using hydrograph decomposition into spring flood curve, runoff and baseflow. Using information on snow cover, baseflow, soil moisture or surface runoff measurements would provide more reliable information to constrain model

parameter estimation and structure (e.g. Clark et al., 2011, McMillan et al., 2011; Rakovec et al., 2016). However, the main challenge of such approach is that the state variables within the model may not be directly commensurable with their corresponding observations (Clark et al., 2011; shafii et al., 2017). These data are also prone to uncertainty which can produce conflicting predictions errors where improvement in a model component results in deterioration in another (Beven, 2006).

5. Conclusions

The present paper proposes a step-wise calibration based on hydrological processes where each component of the model is adjusted separately. The aim of evaluating model components separately is to reduce parameter compensation and constrain model parameters more effectively according to the underlying processes. The underlying processes are evaluated using a hydrograph partition into spring flood curve, runoff and baseflow.

Using the same concepts, in this study parameters controlling drainage in the unsaturated zone are adjusted on the surface runoff, parameters controlling drainage in the saturated zone are adjusted on the baseflow and snow melt parameters on the spring flood curve in an iterative procedure. Our approach differs from the approach of He et al. (2015), where hydrograph partitioning is discontinuous and based on the source of flow (groundwater, snowmelt water, glacier melt). In this study, time series of flows, runoff and baseflow were approximated by temporal functions using functional data analysis. This statistical tool allows conserving the entire shape of time series which enables evaluating model performance during high and low flow periods simultaneously unlike the widely used goodness-of-fit measures. In fact these measures do not count for the change in bias, variance, skew and correlation of actual series of model errors under the different hydrological conditions (Beven, 2006); whereas functional statistics provide information about these changes in time series as illustrated with the functional Student test.

The proposed approach was compared to an overall performance calibration. This comparison shows that the stepwise process-based calibration can achieve high performance and consistency more effectively on site and at a basin scale. Although there is improvement in model performance with the proposed approach, the baseflow representation needs further

improvement. The proposed approach and the defined criteria can be adapted in order to include field measurement of baseflow, information on snow cover or soil moisture. However, using different data, that are prone to uncertainty, may result in producing conflicting predictions in model state variables.

This approach can also be adapted to calibrate any model on various watersheds by approximating data into temporal functions. Subsequently the FDA techniques suggested in this paper can be applied to define performance indicators for model calibration or for model suitability evaluation. For instance, the functional Student test provides objective evaluation of model ability to simulate on average accurate flows for each time step. As illustrated in this study, the first derivative provides information about the rate of change in the rises and falls of the hydrograph. Other FDA descriptive statistics (e.g. mean, covariance) can be used to evaluate model performance to represent the observed hydrographs distribution.

Supplementary materials

Evaluation of model performance with the calibration criteria given in Table 6-3 in the manuscript and evaluation of the realism of internal state variables.

Evaluation of model performance with calibration criteria

Table 6-5 Results of the calculation of the calibration criteria.

Criterion	OptGlob	OptProcess
Glob_OF (m3/s)	22853	13817
Snowmelt_OF (m3/s)	18132	8602
Runoff_OF (m3/s)	15670	11908
Baseflow_OF (m3/s)	12695	11460
Routing_OF	0.0047	0.0065

The lower the value of the criterion, the better is the model performance. The criteria Glob-OF, Snowmelt_OF, Runoff_OF and Baseflow_OF are a sum of surface between the simulated and observed smoothed hydrographs, runoff and baseflow extracted from the hydrographs. These criteria must be evaluated along with the watershed flow regime characteristics.

Evaluation of snowpack, water levels in the Upper and Lower reservoirs of CEQUEAU at the stations LSJ (outlet) and ASHA (upstream station) during the calibration period (Jan 2002 to Dec 2009).

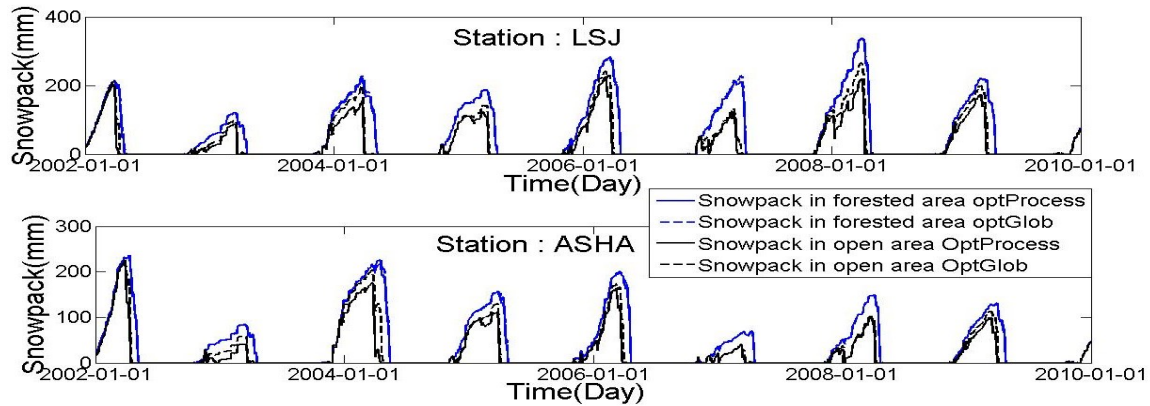


Figure 6-7 Simulated snowpack by the calibrated model with the proposed process-based calibration (OptProcess) and against streamflow solely (OptGlob) at the stations LSJ (outlet) and ASHA (upstream).

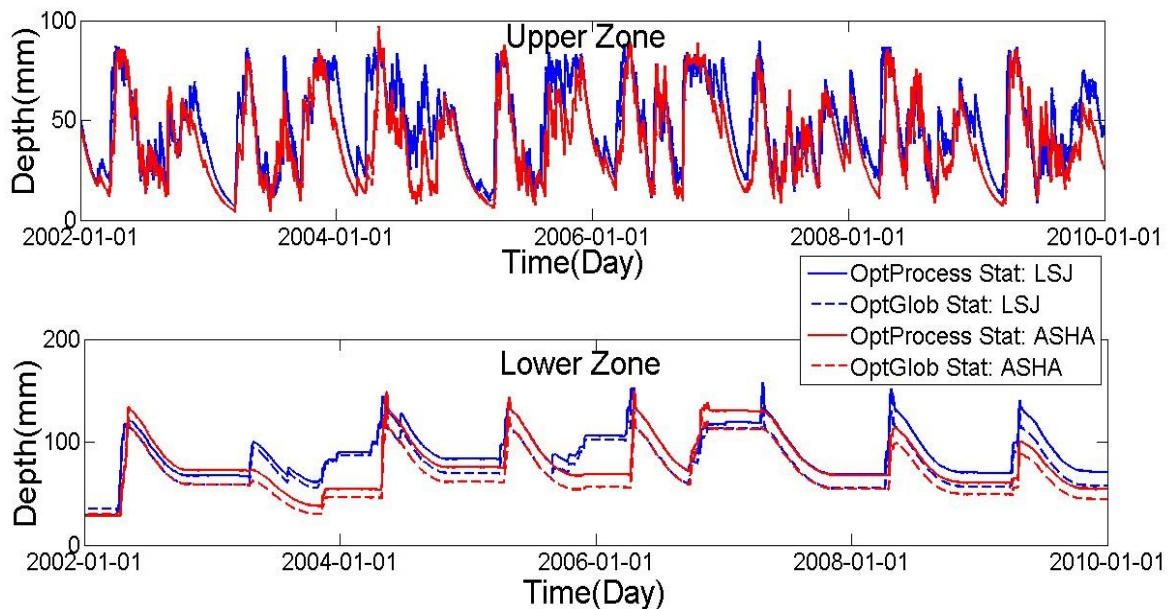


Figure 6-8 Simulated water volume in the saturated (Upper) and unsaturated (Lower) zones by the calibrated model with the proposed process-based calibration (OptProcess) and against streamflow solely (OptGlob) at stations LSJ (outlet) and ASHA (upstream).

Acknowledgements

This project was funded in part by Rio Tinto and NSERC. The authors also acknowledged the in kind contribution of Rio Tinto, especially the assistance of M. Latraverse and B. Larouche.

References

- Beven, K (2006) A manifesto for the equifinality thesis. *J Hydrol* 320:18-36
- Cao W, Bowden W, Davie T, Fenemor A (2006) Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrol Process*: 20:1057-1073
- Charbonneau R, Fortin J, Morin G (1977) The CEQUAU model: Description and examples of its use in problems related to water resource management. *Hydrological Sciences bulletin* 22(1): 193-202
- Clark M. P, McMillan H K , Collins D N , Kavetski D, Woods R (2011) Hydrological field data from a modeller's perspective: Part 2: process-based evaluation of model hypotheses. *Hydrol Process* 25:523–543
- Eckhardt K (2005) How to construct recursive digital filters for baseflow separation. *Hydrol Process* 19:507-515
- Euser T, Winsemius H C , Hrachowitz M, Fencia F, Uhlenbrook S, Savenije H H G (2013) A framework to assess the realism of model structures using hydrological signatures. *Hydrol Earth Syst Sci* 17:1893-1912
- Finger D, Pellicciotti F, Konz M, Rimkus S, Burlando P (2011) The value of glacier mass balance, satellite snow cover images, and hourly discharge for improving the performance of a physically based distributed hydrological model. *Water Resour Res.* doi:10.1029/2010WR009824
- Gupta H, Wagener T, Liu Y (2008) Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrol Proces* 2:3802-3813
- Gupta H, Kling H, Yilmaz K, Martinez G F (2009) Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J Hydrol* 377:80-91
- He Z H, Tian F Q, Gupta H, Hu H C, Hu H P (2015) Diagnostic calibration of a hydrological model in a mountain area by hydrograph partitioning. *Hydrol Earth Syst Sci* 19:1807–1826
- Hay L E, Leavesley G H, Clark M P, Markstrom S L, Viger R J, Umemoto M (2006) Step wise, multiple objective calibration of a hydrologic model for a snowmelt dominated basin. *JAWRA* 42: 877–890. doi:10.1111/j.1752-1688.2006.tb04501.x

- Hingray B, Schaefli B, Mezghani A, Hamdi Y (2010) Signature-based model calibration for hydrological prediction in mesoscale Alpine catchments. *Hydrol Sci J* 55(6):1002-1016. doi:10.1080/02626667.2010.505572
- Krause P, Boyle D, B se F (2005) Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences* 5:89-97
- Labarthe B, Abasq L, de Fouquet C, Flipo N (2014) Stepwise calibration procedure for regional coupled hydrological-hydrogeological models. Geophysical Research Abstracts Vol. 16, EGU2014-11672-1, 2014
- Leta O T, Van Griensven A, Bauwens W (2017) Effect of Single and Multisite Calibration Techniques on the Parameter Estimation, Performance, and Output of a SWAT Model of a Spatially Heterogeneous Catchment. *J Hydrol Eng*. doi: 10.1061/(ASCE)HE.1943-5584.0001471
- Levitin D, Nuzzo R, Vines B, Ramsay J (2007) Introduction to Functional Data Analysis. *Canadian Psychology* 48(3):135-155
- Lim K J, Park Y S, Kim J et al (2010) Development of genetic algorithm-based optimization module in WHAT system for hydrograph analysis and model application. *Computers & Geosciences* 36: 936-944
- McMillan H K, Clark M P, Bowden W B, Duncan M, Woods R A (2011) Hydrological field data from a modeller's perspective: Part 1. Diagnostic tests for model structure. *Hydrol Process* 25: 511–522
- Morin G, Paquet P (2007) Mod le hydrologique CEQUEAU, INRS-ETE, rapport de recherche no R000926, 458p
- Niraula R, Norman L M, Meixner T, Callegary B J (2012) Multi-gauge Calibration for modeling the Semi-Arid Santa Cruz Watershed in Arizona-Mexico Border Area Using SWAT. *Air, Soil and Water Research* 5:41–57
- Pokhrel P, Yilmaz, K, Gupta H (2012) Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures. *J Hydrol* 418-419:49-60
- Pfannerstill M, Guse B, Fohrer N (2014) Smart low flow signature metrics for an improved overall performance evaluation of hydrological models. *J Hydrol* 510:447-458
- Rakovec O, Kumar R, Attinger S, Samaniego L (2016) Improving the realism of hydrologic model functioning through multivariate parameter estimation. *Water Resour Res* 52:7779–7792. doi:10.1002/2016WR019430
- Ramsay J O, Ramsey J B (2002) Functional data analysis of the dynamics of the monthly index of nondurable goods production. *Journal of Econometrics* 107:327-344
- Ramsay J, Silverman B (2002) Applied Functional Data analysis: methods and Case Studies. Springer, New York

- Ramsay J, Silverman B (2005) *Applied Functional Data Analysis: Methods and Case Studies*. Springer, New York
- Schaefli B, Gupta H (2007) Do Nash values have value? *Hydrol Process* 21:2075–2080
- Shafii M, Tolson B (2015) Optimizing hydrological consistency by incorporating hydrological signatures into model calibration objectives. *Water Resour Res.* doi:10.1002/2014WR016520
- Shafii M, Basu N, Craig J R, Schiff S L, Van Cappellen P (2017) A diagnostic approach to constraining flow partitioning in hydrologic models using a multiobjective optimization framework. *Water Resour Res.* doi:10.1002/2016WR019736
- St-Hilaire A, Morin G, El-Jabi N, Caissie D (2000) Water temperature modelling in a small forested stream: implication of forest canopy and soil temperature. *Can J Civ Eng* 27:1095-1108
- St-Hilaire A, Boucher M-A, Chebana F, Ouellet-Proulx S, Zhou Q-X, Larabi S, Dugdale S (2015) Breathing a new life to an older model: The CEQUEAU tool for flow and water temperature simulations and forecasting. In *Proceedings of the 22nd Canadian Hydrotechnical Conference*, Montreal, Qc, Canada, 29 April-2May 2015.
- Tremblay A, Tardif S, Strachan I B, Turpin C (2014) Studying net Evaporation from the Eastmain-1 reservoir. *Hydro Review*. <http://www.hydroquebec.com/sustainable-development/documentation-center/pdf/studying-net-evaporation-eastmain-1-reservoir.pdf>
- Wagener T, Wheater H, Gupta H (2004) *Rainfall-Runoff Modelling in Gauged and Ungauged Catchments*. Imperial College Press, London
- Wang S, Zhang Z, Sun G, Strauss P, Guo J, Tang Y, Yao A (2012) Multi-site calibration, validation, and sensitivity analysis of the MIKE SHE Model for a large watershed in northern China. *Hydrol Earth Syst Sci* 16 :4621–4632. doi:10.5194/hess-16-4621-2012
- Westerberg I, McMillan H (2015) Uncertainty in hydrological signatures. *Hydrol Earth Syst Sci* 12: 4233-4270
- Yilmaz K, Gupta H, Wagener T (2008) A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. *Water Resour Res.* doi:10.1029/2007WR006716
- Zhang X, Srinivasan R, Van Liew M (2008) Multi-site calibration of the SWAT model for hydrologic modeling. *Transactions of the ASABE* 51(6):2039-2049

CHAPITRE 7 CONCLUSION GÉNÉRALE ET PERSPECTIVES

Cette thèse présente nos contributions dans le domaine de la calibration automatique de modèles hydrologiques. Nous avons tâché de répondre à certaines problématiques, à savoir convergence des algorithmes d'optimisation vers des solutions irréalistes bien que performantes, la perte d'information concernant la variabilité temporelle du débit avec les mesures de performances traditionnelles, ainsi que la performance discutable des modèles conceptuels à capter les changements des états internes et processus hydrologiques. Le processus global de la calibration proposé est présenté dans la figure suivante. Ce processus est général et peut être appliqué à n'importe quel modèle.

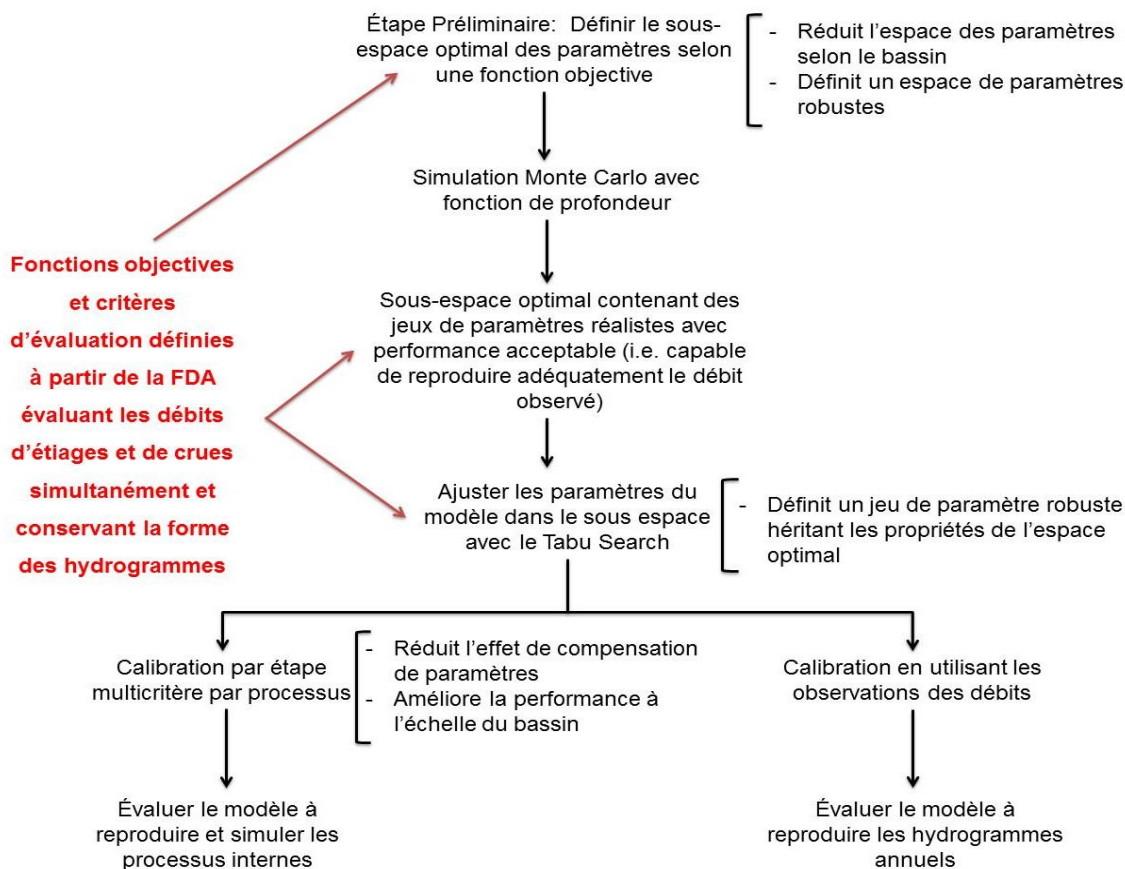


Figure 7-1 Processus général de la calibration proposé dans cette thèse.

➤ Procédure de calibration robuste

En premier lieu, on a proposé une procédure de calibration qui vise d'abord à trouver la région optimale de l'espace des paramètres défini préalablement par le modélisateur et ensuite à calibrer le modèle dans cet espace optimisé. Cette procédure définit l'espace des paramètres selon le bassin en question, ce qui augmente le degré de réalisme des paramètres obtenus. Cette procédure peut être utile et robuste pour la régionalisation des paramètres. En effet, à travers le quatrième objectif utilisant cette procédure (sous-section 5.4), on note une bonne performance des paramètres définis avec cette procédure sur les sites similaires au site utilisé pour la calibration. Ceci démontre le potentiel de cette procédure pour définir une approche plus élaborée pour la régionalisation des paramètres à grande échelle.

La procédure d'optimisation de l'espace des paramètres est basée sur des simulations Monte Carlo avec une fonction de profondeur pour restreindre l'échantillonnage dans la zone contenant des paramètres définissant un modèle performant. Cette procédure pourrait être alors utilisée pour définir la distribution des paramètres en fonction de leur profondeur et ainsi être adaptée pour une analyse d'incertitude. La méthode développée consiste à étudier l'évolution de la valeur moyenne de la fonction objectif. Une autre alternative est d'étudier aussi l'évolution de la distribution de cette fonction objectif.

Dans la présente thèse, on a exploré une version élémentaire de l'algorithme Tabu Search pour l'ajustement des paramètres. L'efficacité de cet algorithme a été comparée au CMAES. À noter que d'autres algorithmes pourraient être utilisés lors de la deuxième étape de la procédure de calibration proposée au lieu du Tabou Search. D'autres études sont nécessaires pour évaluer la performance des algorithmes d'optimisation avec cette procédure.

➤ Fonction objectif conservant la variabilité temporelle du débit

On a aussi proposé une méthode de calibration et évaluation d'un modèle hydrologique basée sur l'analyse de données fonctionnelles. Cette méthode consiste à conserver la forme et la variabilité de l'hydrogramme comme illustrer dans la Figure 7-2. Cette méthode s'avère efficace pour évaluer le modèle à reproduire l'hydrogramme observé en entier. Elle remplace les approches classiques qui consistent à de multiples analyses univariées ou une analyse multivariée pour analyser les différentes composantes de l'hydrogramme.

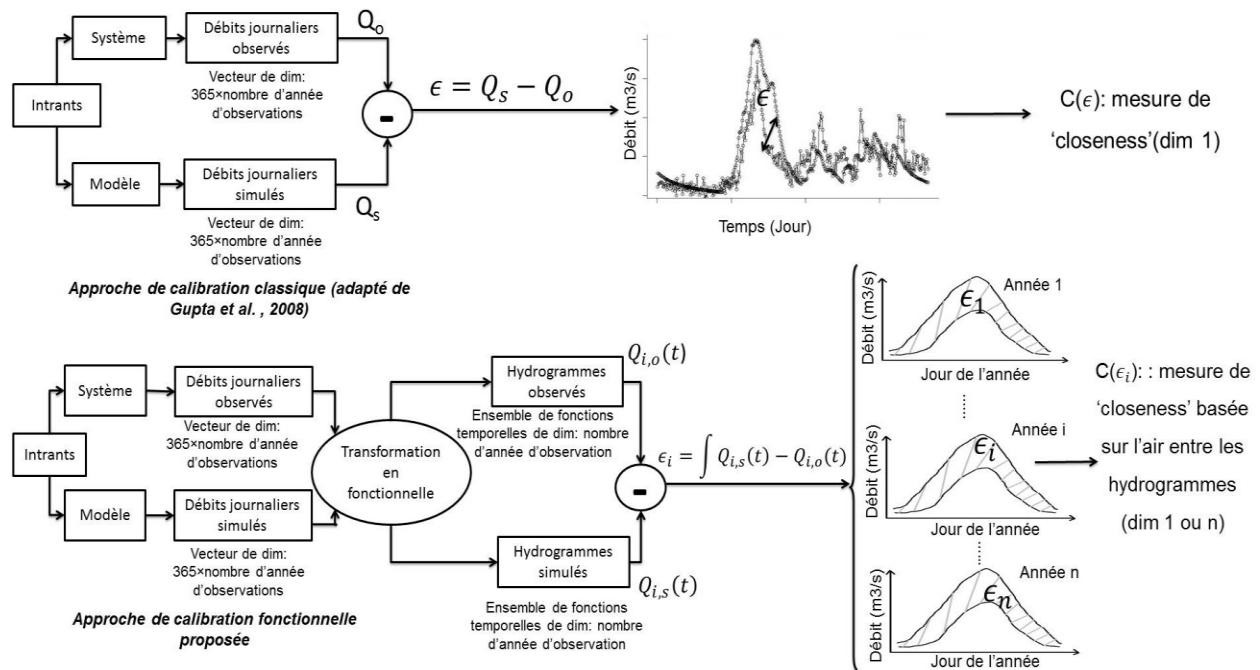


Figure 7-2 Principe de la calibration avec la FDA proposée dans cette thèse ainsi que le principe classique utilisé.

On explore deux formulations de mesure $C(\epsilon_i)$. La première inspirée des tests d'adéquation comparant deux distributions est définie comme la somme des distances entre les hydrogrammes simulés et observés. La deuxième est un ratio entre les surfaces sous chaque courbe (hydrogrammes) ce qui représente le volume annuel. Puisque la procédure développée lors du

premier objectif est destinée à optimiser une seule fonction objectif, dans les deux cas on utilise une agrégation pour calculer un critère global sur la période de données d'observations. Le principe de calibration proposé peut être traité comme un problème multiobjectif en utilisant un algorithme d'optimisation explorant le Front de Pareto et où l'on cherche un compromis entre les différentes formes d'hydrogrammes résultants de la variabilité du climat. D'autres études sont nécessaires pour explorer et améliorer la définition de la fonction objectif définie à partir de la FDA. Cette méthode peut être adaptée pour une calibration par événement ou par saison selon l'objectif et la nature du modèle.

Dans cette thèse, les statistiques fonctionnelles utilisées sont la moyenne et le test de Student fonctionnelle. La dérivée de l'hydrogramme est utilisée comme un indicateur d'occurrence des événements hydrologiques. D'autres statistiques fonctionnelles peuvent être utilisées pour comparer la centralité et la dispersion comme la fonction médiane, mode, covariance (présenté par Chebana et al., 2012).

➤ **Procédure de calibration pour capter la variabilité des processus internes**

On a introduit une stratégie de calibration par processus par étape pour augmenter le degré de réalisme d'un modèle conceptuel et réduire l'effet de compensation des paramètres. On a montré que cette stratégie augmente la performance du modèle, et ce à l'échelle du bassin comparativement à une stratégie classique basée sur l'évaluation du débit uniquement. Comme les observations de débits sont disponibles dans le cadre de cette thèse, les critères décrivant l'adéquation des processus sont définis à partir des éléments extraits des observations de débit. Selon la disponibilité des données (évapotranspiration, débit de base, stock de neige, etc...), cette méthode peut être adaptée pour calibrer le modèle sur d'autres variables mesurées (ex. évapotranspiration, équivalent de neige en eau).

➤ **Procédure de calibration à l'échelle d'un bassin**

Dans cet objectif, on a comparé une calibration multisite contre une calibration par site. On a montré que la calibration multisite peut aboutir à des performances similaires à une calibration par site. La calibration multisite est alors favorisée pour gagner en termes de temps de calcul. Cet objectif de la thèse a permis aussi de faire une première analyse de la variabilité spatiale et temporelle des paramètres du modèle CEQUEAU au bassin versant du lac St-Jean. On a aussi montré que les paramètres résultants sont transférables dans d'autres sites sur ce grand bassin versant.

Les sites sélectionnés dans cette thèse ont été choisis pour couvrir différentes régions du bassin. Toutefois, le choix d'autres sites peut entraîner des résultats différents. Une étude plus avancée considérant les caractéristiques physiographiques de chaque site est nécessaire pour mieux expliquer la faible performance du modèle sur certains sites et aussi pour couvrir une plus grande échelle.

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

- Andréassian, V., Le Moine, N., Perrin, C., Oudin, L., Mathevet, T., Lerat, J. and Berthet, L., 2012. All that glitters is not gold : the case of calibrating hydrological models. *Hydrol. Process.*, 26, 2206-2210.
- Arnold, J. G., Srinivasan, R., Muttiah, R. S., and Williams, J. R., 1998. Large area hydrologic modeling and assessment. Part I: Model development. *Journal of the American Water Resources Association*, 34(1), 73–89.
- Arsenault, R., Latraverse, M. and Duschene, T. (2016). An efficient method to correct underdispersion in ensemble streamflow prediction for seasonal volumetric forecasting. *Water Resources Management*, 30(12), 4363-4380. DOI:10.1007/s11269-016-1425-4
- Arsenault, R., Poulin, A., Côté, P. and Brissette, F., 2013. A comparison of stochastic optimization algorithms in hydrological model calibration. *J. hydrol. Eng.*, 1934-5584.
- Asadzadeh, M., Tolson, B. and Burn, D. H. (2014). A new selection metric for multiobjective hydrologic model calibration. *Water Resour. Res.*, 50, doi:10.1002/2013WR014970.
- Ayadi, M. and Bargaoui Z. (1998). Modélisation des écoulements de l’oued Miliane par le modèle CEQUEAU. *Hydrol. Sci. J.*, 43:5,741-758. DOI:10.1080/02626669809492170
- Bardossy, A. and Singh, S., 2008. Robust estimation of hydrological model parameters. *hydro. Earth Syst. Sci*, 12, 1273-1283.
- Bekele, E. G. and Nicklow, J. W., 2007. Multi-objective automatic calibration of SWAT using NSGA-II. *J. Hydrol.*, 341(3-4), 165-176.
- Bergstrom, S., 1976. Development and application of a conceptual runoff model for Scandinavian countries. SMHI Rep. No. 7, Norrkoping, Sweden.
- Beven, K. J., 1995. Chapter 18: TOPMODEL. Computer models of watershed hydrology, V. P. Singh, ed., Water Resources Publications, Littleton, Colo.
- Beven, K. J., 2001. *Rainfall-Runoff modelling*. The Primer. John Wiley and Sons, Chichester, Uk, 1-360pp.
- Beven, K., 2006. A manifesto for the equifinality thesis. *J. Hydrol.*, 320, 18-36.
- Beven, K. and Binley, A., 1992. The future of distributed models: Model calibration and uncertainty prediction. *Hydrol. Process.*, 6, 279-298.
- Beven, K. J., Calver, A., and Morris, E., 1987. The Institute of Hydrology distributed model. Institute of Hydrology Rep. No. 98, Wallingford, U.K.
- Beven, K. J. and Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.*, 249(1-4), 11-29.

- Boisvert, J., El-Jabi, N., St-Hilaire, A. and El-Adlouni, S., 2016. Parameter estimation of a Distributed hydrological Model using a Genetic Algorithm. *Open Journal of Modern Hydrology*, 151-167.
- Boyle, D. P., Gupta, H. V. and Sorooshian, S., 2000. Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods. *Water Resour. Res.*, 36(12), 3663-3674.
- Brigode, P., Oudin, I. and Perrin, C., 2013. Hydrological model parameter instability: A source of additional uncertainty in estimating the hydrological impacts of climate change? *J. Hydrol.*, 476, 410-425.
- Caissie, D., El-Jabi, N. and St-hilaire, A., 1998. Modelling of stream water temperatures in a small stream using different air to water relations. *Can J. Civ. Eng.*, 25(2), 250-260.
- Cao, W., Bowden, W., Davie, T. and Fenemor, A., 2006. Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrol. Process.*, 1057-1073.
- Casper, M., Grigoryan, G., Gronz, O., Heinemann, G., Ley, R. and Rock, A., 2012. Analysis of projected hydrological behaviour of catchments based on signature indices. *Hydrol. Earth Syst. Sci.*, 409-421.
- Chakraborty, B. and Chaudhuri, P. 2006. Computation of half-space using simulated annealing. *Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science. Volume 72 (eds. R. Liu, R. Serfling, D.L. Souvaine), American Mathematical Society, 159-169.
- Charbonneau, R., Fortin, J-P. and Morin, G., 1977. The CEQUAU model: Description and examples of its use in problems related to water resource management/ Le modèle CEQUEAU: description et exemples d'utilisation dans le cadre de problèmes reliés à l'aménagement. *Hydrological Sciences bulletin*, 22(1), 193-202.
- Chebana, F., Dabo-Niang, S. and Ouarda, T. B.M.J., 2012. Exploratory functional flood frequency analysis and outlier detection. *Water Resour. Res.* doi: 10.1029/2011WR011040.
- Chebana, F. and Ouarda, T. B.M.J., 2008. Depth and homogeneity in regional flood frequency analysis. *Water Resour. Res.* doi: 10.1029/2007WR006771.
- Choi, H. T. and Beven, K., 2007. Multi-period and multi-criteria model conditioning to reduce prediction uncertainty in an application of TOPMODEL within the GLUE framework. *J. Hydrol.*, 332(3-4), 316-336.
- Clark M. P., McMillan, H. K. , Collins, D. N., Kavetski, D. and Woods, R., 2011. Hydrological field data from a modeller's perspective: Part 2: process-based evaluation of model hypotheses. *Hydrol. Process.*, 25, 523–543.

- Clausen, B. and Biggs, B., 2000. Flow variables for ecological studies in temperate streams: groupings based on covariance. *J. Hydrol.*, 237(3-4), 184-197.
- Coron, L., Andéassian, V., Perrin, C., Lerat, J., Vaze, j., Bourqui, M. and Hendrickx, F., 2012. Crash testing hydrological models in contrasted climatic conditions: An experiment on 216 Australian catchments. *Water Resour. Res.* doi:10.1029/2011WR011721.
- Diaz Mercado, V., Bâ, K.M., Quentin, E., Ortiz Madrid, F.H. and Gama, L. (2015). Hydrological Model to Simulate Daily Flow in a Basin with the help of a GIS. *Open Journal of Modern Hydrology*, 5, 58-67
- Duan, Q., Sorooshian, S. and Gupta, V., 1992. Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models. *Water Resour. Res.*, 4(28), 1015-1031.
- Dugdale, S.J., St-Hilaire, A. and Curry R.A. (2017). Automating drainage direction and physiographic inputs to the CEQUEAU hydrological model: Sensitivity testing on the lower Saint John river watershed, Canada. *Journal of Hydroinformatics*. DOI:10.2166/hydro.2017.051.
- Eckhardt, K., 2005. How to construct recursive digital filters for baseflow separation. *Hydrol. Process* 19, 507-515.
- Eleuch, S., Carsteanu, A., Bâ, K., Magagi, R., Goïta, K., and Diaz, C. (2010). Validation and use of rainfall radar data to simulate water flows in the Rio Escondido basin. *Stoch. Environ. Res. Risk. Assess.*, 24:559-565. DOI:10.1007/s00477-009-0336-9.
- Euser, T., Winsemius, H.C., Hrachowitz, M., Fencia, F., Uhlenbrook, S. and Savenije, H.H.G., 2013. A framework to assess the realism of model structures using hydrological signatures. *Hydrol. Earth Syst. Sci.*, 17, 1893-1912.
- Fencia, F., Savenije, H., Matgen, P. and Pfister, L., 2007. A comparison of alternative multiobjective calibration strategies for hydrological modeling. *Water Resour. Res.*, (43), W03434, doi:10.1029/2006WR005098.
- Finger, D., Pellicciotti, F., Konz, M., Rimkus, S. and Burlando, P., 2011. The value of glacier mass balance, satellite snow cover images, and hourly discharge for improving the performance of a physically based distributed hydrological model. *Water. Resour. Res.* doi:10.1029/2010WR009824.
- Fortin, J. P., Turcotte, R., Massicotte, S., Moussa, R., Fitzback, J., and Villeneuve, J. P., 2001a. A distributed watershed model compatible with remote sensing and GIS data. I: Description of model. *J. Hydrol. Eng.*, 6 (2), 91-99.
- Fortin, J. P., Turcotte, R., Massicotte, S., Moussa, R., Fitzback, J., and Villeneuve, J. P., 2001b. A distributed watershed model compatible with remote sensing and GIS data II: Application to Chaudiere watershed. *J. Hydrol. Eng.* 6 (2), 100-108.
- Francés, F., Vélez, J.I. and Vélez, J. J., 2007. Split-parameter structure for the automatic calibration of distributed hydrological models. *J. Hydrol.*, 332(1), 226-240.

- Franchini, M., Galeati, G. and Berra, S., 1998. Global optimization techniques for the calibration of conceptual rainfall-runoff models. *Hydrol. Sci J.*, 43(3), 443-458.
- Gharari, S., Hrachowitz, M., Fenicia, F., and Savenije, H. H.G., 2013. An approach to identify time consistent model parameters: sub-period calibration. *Hydrol. Earth Syst. Sci.*, 17, 149-161. doi:10.5194/hess-17-149-2013.
- Gill, M. K., Kaheil, Y. H. , Khalil, A. , McKee, M. and Bastidas, L., 2006. Multiobjective particle swarm optimization for parameter estimation in hydrology. *Water Resour. Res.*, 42, W07417, doi:10.1029/2005WR004528.
- Glover, F. 1990. Tabu Search a tutorial. *Interfaces* 20, pp. 74-94.
- Glover, F., and Laguna, M. 1997. Tabu search. Kluwer Academic. Publisher, springer.
- Glover, F., Taillard, E. and De Werra, D., 1993. a user's guide to tabu search. *Annals of Operations Research* , 41, 3-28.
- Griensven, A.V, Meixner, T., Grinwald, S., Bishop, T., Diluzio, M. and Srinivasan, R. 2006. A global sensitivity analysis tool for parameters of multi-variable catchment models. *J. Hydrol.*, 3 (24), 10-23.
- Gunkel, A., Shadeed, S., Hartmann, A., Wagener, T. and Lange, J., 2015. Model signatures and aridity indices enhance the accuracy of water balance estimations in a data-scarce Eastern Mediterranean catchment. *J. Hydrol.*, 487–501. doi:10.1016/j.ejrh.2015.08.002
- Gupta, H., Sorooshian, S. and Yapo, P., 1998. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resour. Res.*, 34(4), 751-763.
- Gupta, H., Sorooshian, S. and Yapo, P. O., 1999. Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. *J. Hydrol. Eng.*, 4(2), 135-143.
- Gupta, H., Wagener, T. and Liu, Y., 2008. Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrol. Process.*, 22 (18), 3802-3813.
- Gupta, H., Kling, H., Yilmaz, K. K. and Martinez, G. F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.*, 377(1-2), 80-91.
- Hansen, N. and Ostermeier A. 1996. Adapting Arbitrary Normal Mutation Distributions in Evolution Strategies : The Covariance Matrix Adaptation. In Proceedings of the 1996 IEEE Intern. Conf. on Evolutionary Computation (ICEC '96), 312-317.
- Hansen, N. and Ostermeier A. 1997. Convergence Properties of Evolution strategies with the Derandomized covariance Matrix Adaptation: The $(\mu/\mu_1, \lambda)$ -CMA-ES. EUFIIT'97, 5th Europ. Cong. On Intelligent Techniques and Soft Computing, Proceedings, 650-654.

- Hansen, N. and Ostermeier A. 2001. Completely Derandomized Self-adaptation in Evolution strategies. By the Massachusetts institute of Technology Evolutionary Computation 9(2), 159-195.
- Hartmann, A., Wagener, T., Rimmer, A., Lange, J., Brielmann, H. and Weiler, M., 2013. Testing the realism of model structures to identify karst system processes using water quality and quantity signatures. *Water Resour. Res.*, 49,3345-3358. doi:10.1002/wrcr.20229.
- He, Z. H., Tian, F. Q., Gupta, H., Hu, H. C. and Hu, H. P., 2015. Diagnostic calibration of a hydrological model in a mountain area by hydrograph partitioning. *Hydrol. Earth. Syst. Sci.*, 19,1807–1826.
- Hingray, B., Schaefli, B., Mezghani, A. and Hamdi, Y., 2010. Signature-based model calibration for hydrological prediction in mesoscale Alpine catchments. *Hydrol. Sci. J.*, 1002-1016.
- Holland, J., 1975. *Adaptation in Natural and Artificial Systems*. Michigan USA, Ann Arbor.
- Kavetski, D., Kuczera, G. and Franks, S. W., 2006. Calibration of conceptual hydrological models revisited: 1. Overcoming numerical artefacts. *J. Hydrol.*, 320(1-2), 173-186.
- Kirchner, J. W., 2006. Getting the right the answers for the right reasons: Linking measurments, analyses, and models to advance the science of hydrology. *Water Resour. Res.* W03S04, doi:10.1029/2005WR004362
- Klemeš, V., 1986. Operational testing of hydrological simulation models.. *Hydrol. Sci. J.*, 13-24.
- Kollat, J. B. , Reed, P. M. and Wagener, T., 2012. When are multiobjective calibration trade-offs in hydrologic models meaningful? *Water Resour. Res.* doi:10.1029/2011WR011534.
- Krause, P., Boyle, D. and Båse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*,5, 89-97.
- Leta, O. T., Van Griensven, A., and Bauwens, W., 2017. Effect of Single and Multisite Calibration Techniques on the Parameter Estimation, Performance, and Output of a SWAT Model of a Spatially Heterogeneous Catchment. *J. Hydrol. Eng.* doi: 10.1061/(ASCE)HE.1943-5584.0001471
- Levitin, D., Nuzzo, R., Vines, B. and Ramsay, J., 2007. Introduction to Functional Data Analysis. *Canadian Psychology*, 48(3), 135-155.
- Liew, M. and Garbrecht, J., 2003. Hydrologic simulation of the little washita river experimental watershed using SWAT. *Journal of the American Water Resources Association*, 39(2), 413-426.
- Lim, K., Park, Y. S., Kim, J., Shin, Y.-C., Kim, N. W., Kim, S. J., Jeon, J.-H. and Engel, B. A. 2010. Development of genetic algorithm-based optimization module in WHAT system for hydrograph analysis and model application. *Computers & Geosciences*, 36, 936-944.

- Liu, R., 1990. On a notion of data depth based on random simplices. *The annals of Statistics*, 18(1), 405-414.
- Li, X., Weller, D. E. and Jordan, T. E., 2010. Watershed model calibration using multi-objective optimization and multi-site averaging. *J. Hydrol.*, 380(3-4),277-288.
- Macdonald, J.S., Morrison, J., Patterson, D.A., Heinonen, J. and Foreman, M. (2007). Examination of factors influencing Nechako River Discharge, temperature, and aquatic habitat. *Can. Tech. Rep. Fish. Aquat. Sci.* 2773:vii+32p.
- Macdonald, J., Morrison, J. and Patterson, D., 2011. The efficacy of Reservoir Flow Regulation for Cooling Migration Temperature for Sockeye Salmon in the Nechako Watershed.
- Madsen, H., 2000. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *J. Hydrol.*, 235(3-4),276-288.
- Madsen, H., Wilson, G. and Ammentorp, H. C., 2002. Comparaison of different automated strategies for calibration of rainfall-runoff models. *J. Hydrol.*, 261(1-4),48-59.
- Mahalanobis, P., 1936. on the Generalized Distance in statistics. *Proceedings of the National Academy of India*, 49-55.
- Masselot, P., Dabo-Niang, S., Chebana, F. and Ouarda, T. B.M.J., 2016. Streamflow forecasting using functional regression. *J. Hydrol.*, 538,754-766.
- McMillan, H. K., Clark, M. P., Bowden, W. B., Duncan, M. and Woods, R. A., 2011. Hydrological field data from a modeller's perspective: Part 1. Diagnostic tests for model structure. *Hydrol. Process.*, 25, 511–522.
- Moradkhani, H. and Sorooshian, S., 2009. General review of Rainfall-Runoff Modeling: Model calibration, Data Assimilation, and Uncertainty Analysis. In *Hydrological Modelling and the Water Cycle: Coupling the Atmospheric and hydrological Models*, Sorooshian, S., Hsu, K.-L., Coppola, E., Tomassetti, B., Verdecchia, M., Visconti, G. (Eds.) Springer, 1-24.
- Morin, G. and Paquet, P., 2007. *Morin, G., Paquet, P. (2007). Modèle hydrologique CEQUEAU, INRS-ETE, rapport de recherche no R000926, 458p.*
- Moussa, R. and Chahinian, N., 2009. Comparison of different multi-objective calibration criteria using a conceptual rainfall-runoff model of flood events. *Hydrol. Earth Syst. Sci.*,519-535.
- Nash, J. and Sutcliffe, J., 1970. River flow forecasting through conceptual models. *J. Hydrol.*, 10, 282-290.
- Neitsch, S., Arnold, J., Kiniry, J. and Williams, J., 2005. Soil and Water Assessment Tool Theoretical Documentation. Version 2005.
- Niraula, R., Norman, L. M., Meixner, T. and Callegary, B. J., 2012..Multi-gauge Calibration for modeling the Semi-Arid Santa Cruz Watershed in Arizona-Mexico Border Area Using SWAT. *Air, Soil and Water Research*, 5,41–57.

- Oja, H., 1983. Descriptive Statistics for multivariate distributions. *Statistics & Probability Letters*, 1(6),327-332.
- Pechlivanidis, I. G., Jackson, B. M., McIntyre, N.R. and Wheater, H. S., 2011. Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Global NEST Journal*, 13(3), 193-214.
- Pechlivanidis, I. G., Jackson, B. M., Mcmillan, H. K. and Gupta, H. V., 2012. Using an informational entropy-based metric as a diagnostic of flow duration to drive model parameter identification. *Global NEST Journal*, 14(3), 325-334.
- Pechlivanidis, I., Jackson, B., McMillan, H. and Gupta, H., 2014. Use of an entropy-based metric in multiobjective calibration to improve model performance. *Water Resour. Res.* doi:10.1002/2013WR014537.
- Pfannerstill, M., Guse, B. and Fohrer, N., 2014. Smart low flow signature metrics for an improved overall performance evaluation of hydrological models. *J. Hydrol.*, 510, 447-458.
- Pokhrel, P., Yilmaz, K. and Gupta, H., 2012. Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures. *J. Hydrol.*, 418-419, 49-60.
- Quick, M. C., and Pipes, A., 1977. UBC watershed model. *Hydrol. Sci. Bull.*, XXI(1/3), 285–295.
- Rakovec, O., Kumar, R., Attinger, S. and Samaniego, L., 2016. Improving the realism of hydrologic model functioning through multivariate parameter estimation. *Water Resour Res*, 52, 7779–7792. doi:10.1002/2016WR019430.
- Ramsay, J. and Ramsey, J., 2002. Functional data analysis of the dynamics of the monthly index of nondurable goods production. *Journal of econometrics*, 107,327-344.
- Ramsay, J. and Silverman, B., 2002. *Applied Functional Data analysis: methods and Case Studies..* Second Edition éd. New York: Springer.
- Ramsay, J. and Silverman, B., 2005. *Functional Data Analysis*. 2nd ed. New York: Springer.
- Refsgaard, J. C., and Storm, B., 1995. Chapter 23: MIKE SHE. Computer models of watershed hydrology, V. P. Singh, ed., Water Resources Publications, Littleton, Colo.
- Sawicz, K., Wagener, T., Sivapalan, M., Troch, P.A. and Carillo, G., 2011. Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrol. Earth Syst. Sci.*, 9(15), 2895–2911.
- Schaeffli, B. and Gupta, H. V., 2007. Do Nash values have value? *Hydrol. Process.*, 21(15), 2075–2080.

- Schaefli, B. and Zehe, E., 2009. Hydrological model performance and parameter estimation in the wavelet-domain. *Hydrol. Earth Syst. Sci.*, 13, 1921-1936.
- Seibert, J., 2000. Multi-criteria calibration of conceptual runoff using a genetic algorithm. *Hydrol. Earth Syst. Sci.*, 4, 215-224.
- Seibert, J. 2003. Reliability of Model Predictions Outside Calibration conditions. *Nordic Hydrology*, 34 (5), 477-492.
- Sguera, C., Galeano, P. and Lillo, R., 2016. Functional outlier detection by a local depth with application to NOX levels. *Stoch. Environ. Res. Risk. Assess.*, 30, 1115-1130.
- Shafii, M. and De Smedt, F., 2009. Multi-objective calibration of a distributed hydrological model (WetSpa) using a genetic algorithm. *Hydrol. Earth Syst. Sci.*, 13, 2137-2149.
- Shafii, M. and Tolson, B. A., 2015. Optimizing hydrological consistency by incorporating hydrological signatures into model calibration objectives. *Water Resour. Res.*, doi:10.1002/2014 WR016520.
- Shafii, M., Basu, N., Craig, J. R., Schiff, S. L. and Van Cappellen, P., 2017. A diagnostic approach to constraining flow partitioning in hydrologic models using a multiobjective optimization framework. *Water Resour Res.* doi:10.1002/2016WR019736.
- Singh, V. P. and Woolhiser, D. A., 2002. Mathematical modeling of Watershed Hydrology. *J. hydrol. Eng*, 7(4),270-292.
- Sorooshian, S., Duan, Q. and Gupta, V. K., 1993. Calibration of Rainfall-Runoff models: Application of Global Optimization to the Sacramento Soil Moisture Accounting Model. *Water Resour. Res.*, 1185-1194. doi: 10.1029/92WR02617
- St-Hilaire A. (2000) Modélisation de la température de l'eau et des solides dissous en rivière sur un petit bassin versant forestier. Thèse de doctorat. INRS-ETE.
- St-hilaire, A., Morin, G., El-Jabi, N. and Caissie, D., 2000. Water temperature modelling in a small forested stream : implication of forest canopy and soil temperature. *Can. J. Civ. Eng.*, 27(6), 1095-1108.
- St-Hilaire, A., El-Jabi, N., Caissie D. and Morin, G., 2003. Sensitivity analysis of a deterministic water temperature model forest canopy and soil temperature in Catamaran Brook (New Brunswick, Canada). *Hydrol. Proces.*, 17, 2033-2047.
- St-Hilaire, A., Boucher, M-A. Chebana, F., Ouellet-Proulx, S., Zhou, Q-X, Larabi, S., Dugdale, S. (2015). Breathing a new life to an older model : The CEQUEAU tool for flow and water temperature simulations and forecasting. In Proceedings of the 22nd Canadian Hydrotechnical Conference, Montreal, QC, Canada, 29 April- 2 May 2015.
- Tang, Y., Reed, P. and Wagener, T., 2006. How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration? *Hydrol. Earth Syst. Sci.*, 10, 289-307.

- Ternynck, C., Ben Alaya, M. L., Chebana, F., Dabo-Niang, S. and Ouarda, T. B.M.J., 2015. Streamflow hydrograph classification using functional data analysis. *Journal of hydrometeorology*. doi: 10.1175/JHM-D-14-0200.1
- Tolson, B. A. and Shoemaker, C., 2007. dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resour. Res.*, 40, W01413. doi:10.1029/2005WR004723.
- Tremblay, A., Tardif, S., Strachan, I.B. and Turpin, C., 2014. Studying net Evaporation from the Eastmain-1 reservoir. *Hydro Review*. www.hydroquebec.com/sustainable-development/documentation-center/pdf/studying-net-evaporation-eastmain-1-reservoir.pdf
- Tukey, J., 1975. Mathematics and picturing data. *Proceedings of the international Congress of Mathematics*, Vancouver, 523-531.
- Tung, C.P. and Chou, C.A. 2004. Pattern classification using tabu search to identify the spatial distribution of groundwater pumping. *Hydrogeology Journal*, 12, 488-496.
- Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W. and Sorooshian S., 2003. Effective and efficient algorithm for multiobjective optimization of hydrology models. *Water Resour. Res.* 39(8), 1214. doi:10.1029/2002WR001746.
- Vrugt, J.A and Robinson, B.A. 2007. Improved evolutionary optimization from genetically adaptive multimethod search. *Proc. Natl. Acad. Sci. U.S.A.*, 104, 708-711.
- Wagener, T., Wheeler, H. and Gupta, H., 2004. *Rainfall-Runoff Modelling in Gauged and Ungauged Catchments*. London: Imperial College Press.
- Wagner, T., Sivaoalan, M., Troch, P. and Wodds, R., 2007. Catchment classification and hydrologic similarity. *Geography compass*, 4, 901-931.
- Wang, S., Zhang, Z., Sun, G., Strauss, P., Guo, J., Tang, Y., Yao, A., 2012. Multi-site calibration, validation, and sensitivity analysis of the MIKE SHE Model for a large watershed in northern China. *Hydrol. Earth. Syst. Sci.*, 16, 4621–4632. doi:10.5194/hess-16-4621-2012
- Wazneh, H., Chebana, F. and Ouarda, T., 2013. Optimal depth-based regional frequency analysis. *Hydrol. Earth Syst. Sci.*, 17, 2281-2296.
- Westerberg, I. and McMillan, H., 2015. Uncertainty in hydrological signatures. *Hydrol. Earth Syst. Sci.*, 19, 3951-3968.
- White, K. and Chaubey, I., 2005. Sensitivity Analysis, Calibration, and Validations for a Multisite and multivariable SWAT Model. *Journal of the American Water Resources Association*, 5(41), 1077-1089.
- Yapo, P. O., Gupta, H. V. and Sorooshian, S., 1998. Multi-objective global optimization for hydrologic models. *J. Hydrol.*, 204(1-4), 83-97.

- Yilmaz, K. K., Gupta, H. V. and Wagener, T., 2008. A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. *Water Resour. Res.*, W09417, doi:10.1029/2007WR006716
- Zhang, X., Srinivasan, R. and Liew, M., 2008. Multi-site calibration of the SWAT model for hydrologic modeling. *TRANSACTIONS OF THE ASABE*, 51(6), 2039-2049.
- Zhang, X., Srinivasan, R., Zhao, K. and Liew, M., 2008. Evaluation of global optimization algorithms for parameter calibration of a computationally intensive hydrologic model. *Hydrol. Process.*, 430-441.
- Zheng, C. and Wang, P. 1996. Parameter structure identification using Tabu search and simulated annealing. *Advances in Water Resources*, 4 (19), 215-224.
- Zuo, Y. and Serfling, R., 2000. General notions of statistics depth functions. *The annals of statistics*, 461-482.