Développement d'approches de calage efficaces pour les modèles hydrologiques coûteux en temps de calcul, au moyen du modèle HYDROTEL

Par

Pierre-Luc HUOT

THÈSE PAR ARTICLES PRÉSENTÉE À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE COMME EXIGENCE PARTIELLE À L'OBTENTION DU DOCTORAT EN GÉNIE
Ph. D.

MONTRÉAL, LE 18 AVRIL 2019

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

Pierre-Luc Huot, 2019

Cette licence Creative Commons signifie qu'il est permis de diffuser, d'imprimer ou de sauvegarder sur un autre support une partie ou la totalité de cette œuvre à condition de mentionner l'auteur, que ces utilisations soient faites à des fins non commerciales et que le contenu de l'œuvre n'ait pas été modifié.

PRÉSENTATION DU JURY

CETTE THÈSE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE :

Mme Annie Poulin, directrice de thèse
Département de génie de la construction à l'École de technologie supérieure

M. Charles Audet, codirecteur de thèse
Département de mathématiques et de génie industriel à l'école de Polytechnique de Montréal

M. Stéphane Alarie, membre du jury
Institut de recherche d'Hydro-Québec

M. Michel Baraër, membre du jury
Département de génie de la construction à l'École de technologie supérieure

M. Louis Dufresne, président du jury
Département de génie mécanique à l'École de technologie supérieure

M. André St-Hilaire, membre du jury externe indépendant
Centre Eau Terre Environnement de l'Institut national de recherche scientifique

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

LE 21 FÉVRIER 2019

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Je tiens à remercier en premier lieu ma directrice de recherche Annie Poulin avec qui j'ai le bonheur de collaborer depuis maintenant plus de six ans. Tu as su m'encadrer, me motiver et me pousser à me dépasser jusqu'à devenir le chercheur que je suis aujourd'hui. Malgré tous mes projets personnels et professionnels qui parfois empiétaient sur mes études, tu as fait preuve d'une grande ouverture d'esprit et d'une écoute attentive et importante. Même si la fin de mes études aux cycles supérieures approche, je suis convaincu que nous continuerons à entretenir cette belle relation de collaboration dans le cadre de multiples projets à venir. Je t'envoie un merci sincère et j'espère que tu as eu le même plaisir que moi à travailler ensemble sur ce sujet de thèse pas comme les autres.

Je remercie également mon codirecteur Charles Audet pour m'avoir fait grandir dans un domaine comme celui l'optimisation. Tu as toujours su me conseiller judicieusement, respecter mes capacités personnelles et valoriser la couleur particulière de ma recherche. Je ne peux passer sous silence ta disponibilité hallucinante et ta rapidité de correction incroyable, voire même inexplicable. Merci d'être dévoué pour tes étudiants et de m'avoir si bien dirigé pendant ces quatre dernières années.

Stéphane Alarie, merci de penser souvent en dehors du carré de sable. Tu as permis d'amener mes travaux à un niveau bien plus élevé que je n'aurais pu l'espérer. Je te remercie pour tes questions complexes qui m'ont permis de développer ma compréhension de l'optimisation et mon esprit critique, j'ai fait bien du chemin grâce à toi. Même si personne ne souhaite t'accorder ton juste titre de co-directeur à cause d'un bout de papier supposément important, tu l'as toujours été pour moi.

Pops, Mom, merci de me croire parfois fou d'être un étudiant à 30 ans. Si j'y suis arrivé, c'est parce qu'à travers l'éducation que vous m'avez apportée, j'ai appris qu'avec un peu de volonté, tout peut nous être à porter de mains. Je sais que vous êtes fiers de moi, que vous m'accordez une confiance immense et que vous m'appuierez toujours dans mes décisions. Je

vous remercie pour votre générosité débordante et d'être des modèles inspirants pour mes enfants et moi. J'espère qu'un jour, mes enfants auront la même reconnaissance que j'ai envers vous. Je vous aime.

Je termine mes remerciements avec ceux qui s'adressent à ma tendre amoureuse, Ma chatonne Marie-Ève, celle qui a permis la concrétisation de tant de mes projets. Quelle partenaire de vie incroyable tu es depuis toutes ces années. Sincèrement, je n'arriverais à accomplir rien de tout cela si tu n'étais pas à mes côtés. Merci d'être à l'écoute quand mon stress se fait sentir, merci d'être compréhensive de ma patience fragile, merci de me valoriser et de me faire sentir exceptionnel, merci d'en faire toujours plus que nécessaire, merci d'être une femme qui est capable de tout, mais surtout merci de m'aimer tendrement pour ce que je suis, sans vouloir me changer. C'est grâce à toi si aujourd'hui nous avons un nid familial, d'extraordinaires enfants (un peu malcommodes), un mode de vie qui nous ressemble et des ambitions toujours de plus en plus inspirantes. Ta personne est celle que je veux pour continuer à vivre heureux, à rêver davantage et à aimer toujours plus. Merci. Je t'aime.

Développement d'approches de calage efficace pour les modèles hydrologiques coûteux en temps de calcul

Pierre-Luc HUOT

RÉSUMÉ

Le calage des modèles hydrologiques peut se formuler comme un problème d'optimisation sans dérivée, également appelé problème d'optimisation de « boîte noire » ; c'est-à-dire qu'il est impossible pour l'optimiseur d'exploiter la structure de la fonction objectif dans le but d'améliorer le processus de calage. Ce genre de problème d'optimisation peut devenir coûteux en temps de calcul lorsque des modèles hydrologiques distribués spatialement sont utilisés. L'exécution d'une seule simulation à l'aide de ce type de modèles peut prendre plusieurs minutes et l'optimisation peut requérir plusieurs milliers de simulations. Le calage peut donc impliquer des temps de calcul importants et une approche d'optimisation efficace se doit d'être établie de manière à rendre ces outils applicables aux contextes opérationnels.

La première phase de cette recherche découle de travaux précédents qui ont permis d'identifier le développement prometteur d'une nouvelle stratégie de calage automatique pour les modèles hydrologiques coûteux en temps de calcul. Cette nouvelle approche d'optimisation adaptée au calage du modèle hydrologique HYDROTEL, distribué, à base physique et coûteux en temps de calcul, est développée en combinant les stratégies d'optimisation efficaces de deux algorithmes existants, soient « Dynamically Dimensioned Search » (DDS) et « Mesh Adaptive Direct Search » (MADS). D'abord, la capacité d'exploration globale de l'espace de solutions (espace paramétrique) de l'algorithme DDS permet l'obtention rapide d'un jeu de paramètres produisant une bonne valeur de la fonction objectif (mesurant l'écart entre les débits observés et simulés à l'exutoire d'un bassin versant). Ensuite, les stratégies de recherche de l'algorithme MADS permettent de fournir un raffinement local où les conditions d'optimalité du jeu de paramètre final sont satisfaites. Les résultats obtenus à l'aide de cette nouvelle méthode démontrent que, pour le calage d'HYDROTEL à 10 paramètres de calage, des économies moyennes de 70 % en temps de calcul sont possibles par rapport aux algorithmes de calage traditionnellement employés. Alors que 40 % d'économies en temps de calcul est plutôt obtenu lors du calage d'HYDROTEL à 19 paramètres, sans oublier que les valeurs finales de la fonction objectif sont comparables à celles obtenues avec des algorithmes d'optimisation existants.

La deuxième phase vise à évaluer le potentiel d'utilisation de différentes avenues de construction de modèles substituts. D'une part, les modèles à fidélité réduite selon trois axes de simplification du modèle original sont examinés : (1) la diminution du nombre de stations météorologiques virtuelles sur le territoire, (2) la diminution de la durée de la période de simulation et (3) la diminution de la discrétisation spatiale du territoire en ajustant le nombre d'Unités Hydrologiques Relativement Homogène (UHRH) qui modélisent le bassin versant. La combinaison de ces trois axes de simplification sera également évaluée. Les critères d'évaluation sont la représentativité du modèle substitut envers le modèle original et la diminution du temps de calcul pour la simulation. Les résultats démontrent que la

combinaison des trois axes de simplification permet d'offrir des modèles substituts ayant des niveaux de représentativité et des temps de calcul intéressants pour l'utilisation dans un contexte de calibration. D'autre part, la représentativité des fonctions de surface telles que les fonctions polynomiales et les modèles de Krigeage sera évaluée. Une analyse selon diverses tailles de l'historique des solutions qui permet de construire la fonction de surface et diverses tailles de l'espace paramétrique permettra de tirer les capacités de représentativité des fonctions de surface. Les résultats démontrent que les deux types de fonctions de surface peuvent très bien représenter des sous-espaces paramétriques de petite taille avec un nombre minimal de 100 solutions.

Puis, une troisième phase vise à intégrer les modèles substituts avantageux en terme de temps de calcul au sein de la nouvelle approche hybride d'optimisation DDS-MADS. Plusieurs expérimentations proposeront différents cadres d'optimisation pour la calibration du modèle hydrologique HYDROTEL. Au final, en fonction des objectifs de calibration et des contraintes dont l'utilisateur fait face, quelques cadres d'optimisation sont retenus. Certains offrent une meilleure réduction du temps de calcul et d'autres performent davantage en terme de valeur finale de la fonction objectif. Le meilleur compromis entre la diminution du temps de calcul et la qualité des valeurs finales de la fonction objectif est obtenu lors de l'utilisation de l'approche hybride DDS-MADS performant uniquement sur les modèles à fidélité réduite, suivi de l'algorithme MADS performant sur les modèles hydrologiques originaux.

Mots clés : modélisation hydrologique, calibration, algorithmes d'optimisation, modèles substituts, efficacité, temps de calcul

Development of Efficient Calibration Frameworks for Computationally-Intensive Hydrological Models

Pierre-Luc HUOT

ABSTRACT

The calibration of hydrological models is formulated as a blackbox optimization problem; i.e. the only information available is the objective function value which can be used by the optimizer to improve the calibration process. Running a single simulation may take several minutes in the case of distributed hydrological models, and the calibration process may require thousands of model evaluations; the computational time can thus easily expand to several hours or days, which can be an issue for many operational contexts. Calibration can therefore involve significant computational time and an effective optimization approach must be chosen.

Based from previous works, this research aims first to propose a new promising automatic calibration approach for computationally-intensive hydrological models. The calibration approach is applied to the distributed and computationally-intensive HYDROTEL model on three different river basins located in Québec (Canada) and is developed by combining the efficient optimization strategies of two existing algorithms: the "Dynamically Dimensioned Search" algorithm (DDS) and the "Mesh Adaptive Direct Search" algorithm (MADS). First, the global exploration ability of the DDS algorithm is able to quickly obtained parameter sets which generate a good-quality value of the objective function (difference between observed and simulated streamflows at the watershed outlet). Then, the search strategies of the MADS algorithm provides a local refinement process based on the satisfaction of optimality conditions. Five transitional features are added to adequately merge both algorithms together. Average time savings of 70% of computational time on HYDROTEL with 10 parameters and 40% on HYDROTEL with 19 parameters were achieved by the new calibration approach in comparison with other algorithms traditionally used. In addition to this important reduction in computational times, final values of the objective function are similar to those obtained with existing optimization algorithms.

The second phase aims to evaluate the potential to use different surrogate models that are low-cost and representative of the calibration problems. Three possibilities to construct reduced-fidelity surrogate models from the HYDROTEL model are examined: (1) the reduction of the number of "pseudo-meteorological" stations located on the territory, (2) the reduction of the calibration time-period and (3) the reduction of the watersheds spatial discretization by decreasing the number of simulation units within the modelling, called Relatively Homogenous Hydrological Units (RHHUs). Representativeness and computational time for each type of surrogate model and for the combination of all of them are evaluated and analysed. Results show that the combination of the three types of reduced-fidelity models provides high-level of representativeness and good ratios of computational time between original and surrogate models. Representativeness of the polynomial functions and the Kriging models are also analysed for two variables: the Design of Experiments (DoE)

size and the parametric domain size. Results demonstrate that both types of response surface functions can very well represent short parametric domain with a minimum of 100 evaluated solutions.

Finally, the last part of this study focuses on the use of the low-cost and representative surrogate models previously developed within the DDS-MADS calibration approach and a range of calibration frameworks are proposed. Calibration frameworks are assessed and compared with one another and results demonstrate that exploiting reduced-fidelity models within the DDS-MADS calibration approach decreases the overall computational time while maintaining the quality of the final solutions. The tested frameworks provide a range of tradeoffs between computational time and objective function value, which offers the users the possibility to select the appropriate framework according to their calibration objectives and optimization constraints.

Keywords: hydrological modeling, calibration, optimization algorithm, surrogate models, efficiency, computational time

TABLE DES MATIÈRES

Page

CHAPITRE 1 INTRODUCTION GÉNÉRALE 1

CHAPITRE 2 REVUE DE LITTÉRATURE 5
2.1 Modélisation hydrologique 5
2.2 Calage des modèles hydrologiques distribués 7
2.2.1 Algorithmes d'optimisation hybrides 9
2.2.2 Utilisation de modèles substituts 11
2.3 Problématique de recherche 13

CHAPITRE 3 Objectifs de la recherche 15
3.1 Objectifs 15
3.1.1 Développement d'une nouvelle technique d'optimisation hybride DDS-MADS 15
3.1.2 Évaluation du potentiel d'utilisation des modèles substituts à même le processus de calagebrides 17
3.1.3 Utilisation efficace des modèles substituts au sein d'un processus de calage coûteux en temps de calcul 19

CHAPITRE 4 A HYBRID OPTIMIZATION APPROACH FOR EFFICIENT CALIBRATION OF COMPUTATIONALLY INTENSIVE HYDROLOGICAL MODELS 21
4.1 Introduction 22
4.1.1 Literature review 23
4.1.2 Research objectives and contributions 26
4.1.3 Paper organization 26
4.2 Calibration of hydrological models 27
4.2.1 Model description 27
4.2.2 Objective function 30
4.2.3 Studied watersheds 30
4.3 Preliminary results 34
4.3.1 Benchmark optimization algorithms 34
4.3.2 Computational time 36
4.3.3 Performance comparison 37
4.4 Proposed hybrid calibration approach 41
4.4.1 A hybrid DDS-MADS calibration approach 41
4.4.2 Numerical results 47
4.5 Validation 54
4.6 Discussion and future work 58
4.7 Acknowledgements 62

CHAPITRE 5 LOW-COST AND REPRESENTATIVE SURROGATE HYDROLOGICAL MODELS, PART I: CONSTRUCTION OF SURROGATES 63
5.1 Introduction 64
5.1.1 Literature review 65
5.1.2 Research objectives and contributions 66
5.1.3 Paper organization 67
5.2 The HYDROTEL problems 67
5.2.1 Model structure 68
5.2.2 Representativeness of surrogates versus computational time 70
5.2.3 Potential surrogate models 73
5.3 Original calibration problems 75
5.3.1 Computational time 76
5.3.2 Meteorological data and calibration time-periods 76
5.3.3 Objective function 77
5.4 Computational experiments 78
5.4.1 Reducing the number of meteorological stations 78
5.4.2 Reducing the calibration time-period 83
5.4.3 Reducing the number of RHHUs 88
5.4.4 Final combined surrogate models 93
5.5 Discussion 97
5.6 Acknowledgements 98

CHAPITRE 6 LOW-COST AND REPRESENTATIVE SURROGATE HYDROLOGICAL MODELS. PART II: USE WITHIN CALIBRATION FRAMEWORKS 99
6.1 Introduction 100
6.1.1 Literature review 101
6.1.2 Paper organization 104
6.2 Benchmark Algorithms and calibration framework 104
6.2.1 The Dynamically Dimensioned Search (DDS) algorithm 104
6.2.2 The hybrid DDS-MADS optimization approach 105
6.3 Use of surrogate models 107
6.3.1 Review of reduced-fidelity models highlights 108
6.3.2 Calibration framework experiments 109
6.3.3 Results from the calibration frameworks on HYDROTEL 10 114
6.3.4 Results from the calibration frameworks on HYDROTEL 19 118
6.4 Discussion and conclusion 124
6.5 Acknowledgements 126

CHAPITRE 7 FONCTIONS DE SURFACE 127
7.1 Formulation des fonctions de surface 127
7.2 Expérimentations 130
7.3 Discussion comparative 134

CHAPITRE 8 DISCUSSION GÉNÉRALE 139
8.1 Approche de calage efficace 139
8.2 Modèles substituts représentatifs et moins coûteux en temps de calcul 142
8.3 Modèles substituts exploités au sein de méthodes d'optimisation efficaces 146
8.4 Limitations de l'étude 148
8.5 Défis de la thèse 149

CONCLUSION 153

ANNEXE I 155

ANNEXE II 157

ANNEXE III 159

ANNEXE IV 161

ANNEXE V 163

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES 171

LISTE DES TABLES

Page

Table 4.1 Summary of selected model parameters and boundaries for HYDROTEL 10 and HYDROTEL 19 29

Table 4.2 Summary of computational times and $1 - NSE$ results from automatic optimization for each "Watershed-Optimization Problem" combination 37

Table 5.1 Submodels used in each computing steps of the HYDROTEL routine and impact on computational time consumed when using the proposed surrogate models 69

Table 5.2 Summary features of the six original simulation model problems relative to three studied watersheds modeled on HYDROTEL 10 and HYDROTEL 19 77

Table 5.3 Average CPU time for one single simulation and CPU time ratio of surrogates having a reduced number of meteorological stations for all "Watershed-HYDROTEL 10" combinations 82

Table 5.4 Average CPU time for one single simulation and CPU time ratio of surrogates having a reduced number of meteorological stations for all "Watershed-HYDROTEL 19" combinations 82

Table 5.5 Average CPU time for one single simulation and CPU time ratio of surrogates having a reduced calibration time-period for all "Watershed-HYDROTEL 10" combinations 87

Table 5.6 Average CPU time for one single simulation and CPU time ratio of surrogates having a reduced calibration time-period for all "Watershed-HYDROTEL 19" combinations 87

Table 5.7 Average CPU time for one single simulation and CPU time ratio of surrogates having a reduced number of RHHUs for all "Watershed-HYDROTEL 10" combinations 92

Table 5.8 Average CPU time for one single simulation and CPU time ratio of surrogates having a reduced number of RHHUs for all "Watershed-HYDROTEL 19" combinations 93

Table 5.9 Average CPU time for one single simulation and CPU time ratio of final combined reduced-fidelity models for all "Watershed-HYDROTEL 10" combinations 94

Table 5.10 Average CPU time for one single simulation and CPU time ratio of final combined reduced-fidelity models for all "Watershed-HYDROTEL 19" combinations.....95

Table 6.1 Three calibration frameworks exploiting reduced-fidelity models and response surface functions tested on the original HYDROTEL 10 problems.....110

Table 6.2 Five calibration frameworks exploiting reduced-fidelity models and response surface functions tested on the original HYDROTEL 19 problems.....113

Table 6.3 Minima, Maxima and averages of computational times (hours) and computational time reductions (%) for each calibration framework A, B and C in comparison with DDS-MADS-Bench and DDS-Bench. Ceizur, Cowansville and Toulnustouc watersheds are modelled on HYDROTEL 10.....117

Table 6.4 Minima, Maxima and averages of computational times (hours) and computational time reductions (%) for each calibration framework A to E in comparison with DDS-MADS-Bench and DDS-Bench. Ceizur, Cowansville and Toulnustouc watersheds are modelled on HYDROTEL 19.....122

LISTE DES FIGURES

Page

Figure 3.1 Présentation des stratégies d'optimisation avantageuses (encerclés en vert) des algorithmes DDS et MADS pour la mise en place d'une approche de calage hybride efficace en temps de calcul 16

Figure 3.2 Identification des trois avenues de modèles à fidélité réduite examinées et de leur combinaison (D) : réduction du nombre de pseudo-stations météorologiques (A), réduction de la durée de la période de calage (B) et de réduction du degré de discrétisation spatiale (C) 18

Figure 3.3 Cycle de calage qui intègre l'exploitation des modèles substituts.....20

Figure 4.1 Drainage maps of the three River basins with the studied watersheds (grey area): Ceizur, Cowansville and Toulnustouc.....33

Figure 4.2 Final $1 - NSE$ results from optimization by DDS, MADS and SCE-UA for the calibrations of HYDROTEL 10 and HYDROTEL 19 on the Ceizur watershed at 250, 500 and 1,000 simulations, and at the end of the optimization 39

Figure 4.3 Pseudocode of the hybrid DDS-MADS calibration approach 45

Figure 4.4 Experiments on rounding the objective function value (number of digits) for the Ceizur watershed on HYDROTEL 10 problem.....47

Figure 4.5 Evolutions of the best objective function values obtained ($1 - NSE$) with the hybrid DDS-MADS approach compared to DDS-Bench values. Optimization was performed for the HYDROTEL 10 and 19 problems on the Cowansville watershed.....48

Figure 4.6 Evolutions of the best objective function values obtained ($1 - NSE$) with the hybrid DDS-MADS approach compared to DDS-Bench values. Optimization was performed for the HYDROTEL 10 and 19 problems on the Ceizur watershed.....49

Figure 4.7 Number of model simulations required by the MADS step to improve the endpoints of the DDS step on HYDROTEL 10 and HYDROTEL 19 for the Ceizur and Cowansville watersheds.....51

Figure 4.8 Final $1 - NSE$ optimization results and computational time savings of DDS-MADS (separated into DDS and MADS steps) compared to DDS-Bench results. Optimization was performed for the HYDROTEL 10 and 19 problems on the Cowansville watershed 52

Figure 4.9 Final $1 - NSE$ optimization results and computational time savings of DDS-MADS (separated into DDS and MADS steps) compared to DDS-Bench results. Optimization was performed for the HYDROTEL 10 and 19 problems on the Ceizur watershed.....53

Figure 4.10 Evolutions of the best objective function values obtained ($1 - NSE$) with the hybrid DDS-MADS approach compared to DDS-Bench values. Optimization was performed for the HYDROTEL 10 and 19 problems on the Toulnustouc watershed.....56

Figure 4.11 Final $1 - NSE$ optimization of DDS-MADS compared to DDS-Bench and DDS-600 results in the calibration and validation periods for the HYDROTEL 10 on the Toulnustouc watershed. Final $1 - NSE$ optimization of DDS-MADS compared to DDS-Bench and DDS-1200 results in the calibration and validation periods for the HYDROTEL 10 on the Toulnustouc watershed 58

Figure 5.1 Representativeness of two hypothetical cases illustrating original versus surrogate functions (A & B) with R-Square coefficient (R^2) and Spearman's rank correlation coefficient (Rs) 72

Figure 5.2 $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Toulnustouc watershed comparing the surrogates having a reduced number of meteorological stations with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation.....80

Figure 5.3 $1-NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Toulnustouc watershed comparing the surrogates having a reduced calibration time-period with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation...85

Figure 5.4 $1-NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Toulnustouc watershed comparing the surrogates having a reduced number of RHHUs with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation 90

Figure 5.5 $1-NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Ceizur, Cowansville and Toulnustouc watersheds comparing the final selected surrogates with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation. The reduction levels are also indicated below each case presented in the figure 96

Figure 6.1 Final $1-NSE$ results from calibration frameworks A, B and C (boxplots), and for DDS-MADS-Bench and DDS-Bench. Average computational times are listed for all optimization methods. Calibrations were performed on all three modelled watersheds: Ceizur, Cowansville and Toulnustouc with HYDROTEL 10 115

Figure 6.2 Final $1-NSE$ results from calibration frameworks A to E (boxplots), and for DDS-MADS-Bench and DDS-Bench. Average computational times are listed for all optimization methods. Calibrations were performed on all three modelled watersheds: Ceizur, Cowansville and Toulnustouc with HYDROTEL 19 120

Figure 7.1 Écarts entre les valeurs de la fonction objectif $1 - NSE$ obtenues avec les sept modèles de krigeage (corrélation) et celles obtenues avec le modèle hydrologique GR4J pour les bassins versants Oklahoma et Texas 129

Figure 7.2 Évaluation de la représentativité des fonctions polynomiales selon les valeurs de la fonction objectif $1 - NSE$ obtenues sur HYDROTEL 10 pour le bassin versant Ceizur. Les coefficients R_s et R^2 sont présentés pour chacun des deux scéanrios d'échantillonnage 132

Figure 7.3 Évaluation de la représentativité des fonctions polynomiales selon les valeurs de la fonction objectif $1 - NSE$ obtenues sur HYDROTEL 19 pour le bassin versant Ceizur. Les coefficients R_s et R^2 sont présentés pour chacun des deux scéanrios d'échantillonnage 133

Figure 7.4 Évaluation de la représentativité des modèles de krigeage selon les valeurs de la fonction objectif $1 - NSE$ obtenues sur HYDROTEL 10 pour le bassin versant Ceizur. Les coefficients R_s et R^2 sont présentés pour chacun des deux scéanrios d'échantillonnage 136

Figure 7.5 Évaluation de la représentativité des modèles de krigeage selon les valeurs de la fonction objectif $1 - NSE$ obtenues sur HYDROTEL 19 pour le bassin versant Ceizur. Les coefficients R_s et R^2 sont présentés pour chacun des deux scéanrios d'échantillonnage 137

Figure-A I-1 Final $1 - NSE$ results from optimizations by DDS, MADS and SCE-UA for the calibrations of HYDROTEL 10 and HYDROTEL 19 on the Cowansville watershed at 250, 500 and 1,000 simulations, and at the end of the optimization.....153

Figure-A II-1 $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Ceizur watershed comparing the surrogates having a reduced number of meteorological stations with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation.....157

Figure-A II-2 $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Cowansville watershed comparing the surrogates having a reduced number of meteorological stations with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation.....158

Figure-A III-1 $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Ceizur watershed comparing the surrogates having a reduced calibration time-period with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation 159

Figure-A III-2 $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Cowansville watershed comparing the surrogates having a reduced calibration time-period with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation.....160

Figure-A IV-1 $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Ceizur watershed comparing the surrogates having a reduced number of RHHUs with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation 161

Figure-A IV-2 $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Cowansville watershed comparing the surrogates having a reduced number of RHHUs with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation 162

Figure-A V-1 Évaluation de la représentativité des fonctions polynomiales selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 10 pour le bassin versant Cowansville. Les coefficients R_s et R^2 sont présentés pour chacun des deux scéanrios d'échantillonnage 163

Figure-A V-2 Évaluation de la représentativité des fonctions polynomiales selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 19 pour le bassin versant Cowansville. Les coefficients R_s et R^2 sont présentés pour chacun des deux scéanrios d'échantillonnage 164

Figure-A V-3 Évaluation de la représentativité des modèles de krigeage selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 10 pour le bassin versant Cowansville. Les coefficients R_s et R^2 sont présentés pour chacun des deux scéanrios d'échantillonnage 165

Figure-A V-4 Évaluation de la représentativité des modèles de krigeage selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 19 pour le bassin versant Cowansville. Les coefficients R_s et R^2 sont présentés pour chacun des deux scéanrios d'échantillonnage 166

Figure-A V-5 Évaluation de la représentativité des fonctions polynomiales selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 10 pour le bassin versant Toulnustouc. Les coefficients R_s et R^2 sont présentés pour chacun des deux scéanrios d'échantillonnage 167

Figure-A V-6 Évaluation de la représentativité des fonctions polynomiales selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 19 pour le bassin versant Toulnustouc. Les coefficients R_s et R^2 sont présentés pour chacun des deux scéanrios d'échantillonnage 168

Figure-A V-7 Évaluation de la représentativité des modèles de krigeage selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 10 pour le bassin versant Toulnustouc. Les coefficients R_s et R^2 sont présentés pour chacun des deux scéanrios d'échantillonnage 169

Figure-A V-8 Évaluation de la représentativité des modèles de krigeage selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 19 pour le bassin versant Toulnustouc. Les coefficients R_s et R^2 sont présentés pour chacun des deux scéanrios d'échantillonnage 170

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

ABC	Artificial Bee Colony
ANNs	Artificial Neural Networks
ASA	Adaptive Simulated Annealing
CCE	Competitive Complex Evolution
CMAES	Covariance Matrix Adaptation Evolution Strategy
DACE	Design and Analysis of Computer Experiments
DEH	Direction de l'Expertise Hydrique
DDS	Dynamically Dimensioned Search
DDS-Bench	Dynamically Dimensioned Search used as Benchmark
DDS-MADS	Hybrid optimization approach merging the Dynamically Dimensioned Search with the Mesh Adaptive Direct Search
DDS-MADS-Bench	DDS-MADS used as Benchmark
DFRTT-IM	Dipole Flow and Reactive Tracer Test – Interpretation Model
GIS	Geographic Information System
HDE	Hybrid Differential Evolution
HJ	Hooke and Jeeves Algorithm
KGE	Critère du Kling-Gupta (Kling-Gupta Efficiency)
KNN	K-Nearest Neighboor Classifier
LHS	Latin Hypercube Sampling
MADS	Mesh Adaptive Direct Search
NLWIS	National Land and Water Information Service
NSE	Critère du Nash-Sutcliffe (Nash-Sutcliffe Efficiency)

NSERC	National Sciences and Engineering Research Council
NSGA II	Non-Dominated Sorting Genetic Algorithm II
PADDS	Pareto Archived Dynamically Dimensioned Search
ParaPADDS	Parallel Pareto Archived Dynamically Dimensioned Search
PDDS	Asynchronous Parallel Dynamically Dimensioned Search
PSO	Particle Swarm Optimization
RBF	Radial Basis Function
RHHU	Relatively Homogenous Hydrological Unit
RMSE	Root Mean Squared Error
Rs	Coefficient de corrélation de Spearman
R^2	Coefficient de détermination
SCE-UA	Shuffled Complex Evolution of the University of the Arizona
SIG	Systèmes d'Information Géographique
SVMs	Support Vector Machines
SWAT2000	Soil and Water Assessment Tool, version 2000
VNS	Variable Neighborhood Search

CHAPITRE 1

INTRODUCTION GÉNÉRALE

Le territoire de la province de Québec est tapissé par un vaste réseau hydrographique composé d'un fleuve, de rivières, de lacs et de ruisseaux. La disponibilité abondante de cette ressource en eau et son accessibilité engendrent un lot d'activités, allant de la production d'hydroélectricité à l'approvisionnement des réseaux de distribution d'eau potable et jusqu'à l'étude des zones inondables et des impacts des changements climatiques sur la ressource (MELCC, 2019). De nos jours, toutes ces pratiques passent inévitablement par l'utilisation d'outils informatisés tels que les modèles hydrologiques qui simulent, à l'échelle d'un bassin versant, la représentation des divers processus hydrologiques (Singh et Woolhiser, 2002). Ces modèles sont indispensables puisqu'ils permettent, d'une part, de mieux comprendre la distribution spatio-temporelle et le cheminement de la ressource en eau à l'échelle du bassin versant et, d'autre part, d'apporter une panoplie d'informations pertinentes à la prise de décisions pour la gestion de cette ressource en lien notamment avec l'exploitation d'infrastructures de l'eau (Singh et Woolhiser, 2002).

Chaque fois qu'un modèle hydrologique est appliqué à un nouveau bassin versant, les valeurs de ses paramètres internes doivent être ajustées. Ce processus d'ajustement, communément appelé calage, calibration ou optimisation, permet d'adapter le modèle hydrologique au bassin versant modélisé en reproduisant une série temporelle de conditions historiques observées, le plus souvent les débits en rivière. L'ajustement manuel de ces paramètres est possible via un processus d'essais et d'erreurs effectué par l'utilisateur, mais ceci peut s'avérer complexe et laborieux, surtout lorsque le nombre de paramètres du modèle est important (souvent supérieur à 10). Le recours à des procédures automatisées, reposant sur l'utilisation d'algorithmes d'optimisation, est très souvent la méthode envisagée et exige, en général, peu de travail de la part de l'utilisateur. L'algorithme d'optimisation a alors comme rôle de générer des jeux de paramètres via un processus de recherche réfléchi et structuré au sein de l'espace paramétrique (Pechlivanidis et al., 2011). Chaque nouveau jeu de paramètres est ensuite passé au modèle hydrologique qui effectue une simulation produisant une valeur de la fonction objectif. Cette valeur, qui constitue une indication de la performance de la

combinaison de paramètres testée, est alors retournée à l'algorithme d'optimisation qui poursuit ensuite sa recherche (Duan et al., 1994 ; Moriasi et al., 2007 ; Pushpalatha et al., 2012). Au final, la méthode d'optimisation fournira, au meilleur de son raisonnement et après la réalisation d'une certaine quantité de simulations du modèle, un jeu de paramètres qui permet de reproduire le comportement observé du bassin versant.

La diversification dans l'utilisation des modèles hydrologiques a mené à la création de plusieurs types de modèles qui permettent de répondre à des applications diversifiées. Parmi les différents types de modèles, certains s'appuient sur l'utilisation de bases de données issues de la télédétection à fine résolution ou des systèmes d'informations géographiques (SIG) pour spatialiser la modélisation à l'échelle du bassin versant (Singh et Woolhiser, 2002 ; Das et al., 2008 ; Pechlivanidis et al., 2011). Toutefois, les simulations à l'aide de ce type de modèle impliquent des temps de calcul importants et ceci peut alourdir considérablement le temps de calage. Conséquemment, quelques heures à plusieurs jours peuvent être nécessaire à la réalisation de l'étape de calage (Razavi et al., 2010). Il est donc impératif de choisir une méthode d'optimisation efficace puisque ce choix pourra avoir un impact considérable, notamment dans un contexte opérationnel.

Les trois principaux objectifs de cette étude se divisent comme suit : (1) le développement d'une nouvelle approche de calage efficace pour les modèles hydrologiques coûteux en temps de calcul, (2) l'évaluation et la construction de modèles substituts représentatifs et moins coûteux en temps de calcul et (3) l'analyse et l'utilisation efficace des modèles substituts développés au sein d'un cadre d'optimisation.

Cette thèse par articles est séparée en 8 chapitres. Suivant ce premier chapitre introductif, une revue de littérature scientifique en lien avec la problématique examinée dans cette étude est d'abord présentée au Chapitre 2. Suivra une description des trois objectifs principaux au Chapitre 3. Le Chapitre 4 présente l'article intitulé « A Hybrid Optimization Approach for Efficient Calibration of Computationally-intensive Hydrological Models ». Cette publication a été soumise à la revue scientifique *Hydrological Sciences Journal* en juillet 2018. Elle

présente une nouvelle approche hybride combinant deux algorithmes d'optimisation existants qui calibre efficacement des modèles hydrologiques coûteux en temps de calcul. Les Chapitres 5 et 6 présentent un projet de recherche divisé en deux parties. Les deux articles conjoints s'intitulent « Low-Cost and Representative Surrogate Hydrological Models, Part I: Construction of Surrogates » et « Low-Cost and Representative Surrogate Hydrological Models, Part II: Use Within Calibration Frameworks » et ont été soumis à la revue scientifique *Environmental Modelling & Software* en décembre 2018. Dans un premier temps, ce projet de recherche vise à évaluer la représentativité de diverses avenues pour la construction de modèles à fidélité réduite moins coûteux en temps de calcul. Puis, l'analyse de l'utilisation de ces modèles au sein de méthodes de calage est alors effectuée. Les résultats obtenus permettent de proposer des cadres d'optimisation pleinement adaptés à la problématique de calage de modèles hydrologiques coûteux en temps de calcul. En plus des modèles à fidélité réduite qui sont explorés dans l'Article 2 du Chapitre 5, le Chapitre 7 aborde les fonctions de surface qui ne sont pas discutées dans les trois articles. La formulation de deux types de fonctions de surface : les fonctions polynomiales et les modèles de krigeage sont testés dans le but d'étudier également cette famille de modèles substituts. Finalement, le Chapitre 8 présente une discussion générale survolant l'ensemble des objectifs de cette thèse de doctorat et des recommandations sont par la suite exprimées au sein de la conclusion.

CHAPITRE 2

REVUE DE LITTÉRATURE

Ce deuxième chapitre passe en revue la littérature pertinente liée à ce projet de recherche. Deux thèmes principaux sont abordés : la modélisation hydrologique et le calage des modèles hydrologiques. La problématique de recherche fait également l'objet d'une section distincte.

2.1 Modélisation hydrologique

La modélisation hydrologique est un outil de simulation constitué d'une suite d'équations mathématiques qui représentent les divers processus hydrologiques (évapotranspiration, ruissellement, infiltration, etc.) à l'échelle d'un bassin versant (Singh et Woolhiser, 2002). Dans le but de reproduire la relation qui existe entre les précipitations et les débits en rivière, des séries de données météorologiques (précipitations et températures, le plus souvent) et des caractéristiques du bassin versant sont des éléments indispensables à la constitution d'une telle modélisation (Singh et Woolhiser, 2002 ; Moradkhani et Sorooshian, 2009). La simulation des processus hydrologiques de même que l'utilisation des séries de données météorologiques et des caractéristiques intrinsèques du bassin versant peuvent être intégrées au sein du modèle selon différents degrés de spatialisation. Le territoire est alors découpé en unités spatiales de modélisation et on retrouve différents degrés ou échelles de spatialisation dans la littérature : globale (une seule entité), semi-distribuée (division grossière en sous-bassins versants) et distribuée (division raffinée en mailles ou sous-bassins versants élémentaires à dimension réduite) (Singh et Woolhiser, 2002 ; Das et al., 2008 ; Moradkhani et Sorooshian, 2009 ; Pechlivanidis et al., 2011).

La modélisation distribuée permet quant à elle une meilleure représentation de la variabilité spatiale des caractéristiques physiographiques du bassin versant (types de sols, occupation du territoire, topographie, etc.) et de la distribution spatiale des données météorologiques lorsque plusieurs stations d'observations ou encore de grilles de données sont disponibles, mais aussi des données produites en divers points du bassin lors des simulations. La spatialisation des processus au sein du modèle offre certains avantages comme la possibilité

d'étudier certains processus hydrologiques non perceptibles à une échelle globale. Par exemple, certains contextes d'analyse nécessitent l'étude des débits en rivières en des points autres que l'exutoire (Lei et al., 2011 ; Huang et al., 2014) ou encore l'étude du retrait des couverts glaciaires (Verbunt et al., 2003). Ces contextes peuvent être déterminants pour la prise de décision quant à la gestion de la ressource en eau ou à la construction en milieu hydrique.

Malgré la croissance de la puissance des ressources informatiques, le recours à des modèles hydrologiques distribués engendre des coûts en temps de calcul importants lors des simulations du modèle. À titre d'exemple, l'exécution d'une seule simulation à l'aide d'un modèle distribué peut prendre plusieurs minutes et le processus de calage peut requérir des milliers de simulations. De façon générale, la durée d'une simulation de ces modèles hydrologiques dépend entre autres des équations mathématiques gouvernant la représentation des processus hydrologiques, des stratégies de résolution numériques employées, de l'échelle de discrétisation spatiale utilisée pour la modélisation, de la superficie du territoire modélisé, du nombre de stations météorologiques fournissant les données d'entrées ainsi que du nombre de paramètres à optimiser (Mungunthan et al., 2005 ; Zhang et al., 2009 ; Razavi et al., 2010 ; Huot, 2014). On retrouve des exemples concrets dans la littérature quant au temps de calcul qui peut parfois s'avérer important. Selon l'étude de Razavi et al. (2012b), une simulation à l'aide du modèle hydrologique « Soil and Water Assessment Tool, version 2000 » (SWAT2000 ; Neitsch et al., 2002) sur une période de simulation de 9 années et un bassin versant de taille moyenne (environ 1000 km²) nécessite des temps de calcul moyens de 1,8 minutes lorsqu'effectuée sur un ordinateur à processeur Intel Pentium de 2,8-GHz. De ces mêmes travaux, une simulation du transport des eaux souterraines à l'aide du modèle « Dipole Flow and Reactive Tracer Test – Interpretation Model » (DFRTT-IM ; Roos, 2009) consomme en moyenne 37 minutes, toujours avec le même type d'équipement informatique. Les travaux de Huot (2014) et de Poulin et al. (2011) ont quant à eux employé le modèle hydrologique HYDROTEL (Fortin et al., 2001a-b) qui nécessite environ une minute de temps de simulation pour un bassin versant de taille moyenne (environ 1000 km²) et pour une période de simulation de 4 années sur un ordinateur à processeur Intel Core i7 de 3,40-GHz.

Ces multiples exemples démontrent à quel point l'utilisation de la modélisation hydrologique distribuée peut devenir coûteuse et avoir un impact considérable sur l'utilisation des outils de modélisation dans un contexte opérationnel.

2.2 Calage des modèles hydrologiques distribués

En plus des divers éléments essentiels mentionnés précédemment, la plupart des modèles hydrologiques doivent être adaptés aux bassins versants étudiés par l'ajustement des valeurs de ses paramètres internes qui sont intégrés au sein des équations mathématiques et qui sont appelés à varier en fonction du territoire (bassin versant) à l'étude. Ce processus d'ajustement, appelé calage, calibration ou optimisation, vise à identifier un jeu de paramètres qui minimise l'écart entre la réponse hydrologique simulée par le modèle et des conditions historiques observées sur le bassin versant, le plus souvent des débits à l'exutoire (Duan et al., 1994 ; Singh et Woolhiser, 2002 ; Moriasi et al., 2007 ; Moradkhani et Sorooshian, 2009 ; Pechlivanidis et al., 2011 ; Pushpalatha et al., 2012). Le calage vise donc à évaluer une multitude de jeux de paramètres selon un critère objectif, communément appelé fonction objectif, dans le but de trouver une combinaison qui satisfait le mieux possible les besoins de l'utilisateur du modèle hydrologique. Ces paramètres sont alors essentiels à la performance des simulations hydrologiques octroyant ainsi un rôle important au processus de calage.

Le problème d'optimisation posé par le calage des modèles hydrologiques se classe dans la catégorie des problèmes sans dérivée (« Derivative-Free Optimization ») ; c'est-à-dire qu'il est impossible pour tout algorithme d'optimisation d'exploiter la structure de la fonction objectif (dérivées ou leurs approximations) pour trouver un jeu de paramètres optimal (Conn et al, 2009 ; Audet, 2014). L'optimisation sans dérivée appartient à la famille plus générale appelée l'optimisation de boîte noire, qui utilise pour seule information la valeur de la fonction objectif obtenue lors des simulations du modèle hydrologique. Plusieurs simulations du modèle hydrologique sont alors nécessaires, tel que mentionné précédemment, et l'algorithme devra baser ses décisions uniquement sur les réponses obtenues (Audet et al.,

2010 ; Shan et Wang, 2010 ; Audet et al., 2012 ; Rios et Sahinidis, 2013). Des approches d'optimisation adaptées à ce type de problème doivent donc être utilisées.

La littérature expose une multitude d'approches de calage qui apportent des stratégies d'optimisation efficaces et variées allant des métaheuristiques conventionnelles (Duan et al., 1993 ; Ingber, 1993 ; Madsen, 2003 ; Hansen, 2006 ; Tolson et Shoemaker, 2007) aux méthodes d'optimisation à recherche par coordonnée (Nelder et Mead, 1965 ; Abramson et Audet, 2006 ; Audet et Dennis, 2006 ; Audet et al., 2010 ; Le Digabel, 2011 ; Audet et al., 2012 ; Rios et Sahinidis, 2013 ; Audet, 2014). Bien que cet ensemble de méthodes puisse démontrer certains potentiels d'efficacité, le développement de ces approches n'est pas orienté spécifiquement vers les problèmes d'optimisation coûteux en temps de calcul. Parmi l'ensemble des algorithmes d'optimisation développés, certains ont déjà démontré, dans le cadre de plusieurs études, leur utilité pour le calage de modèles hydrologiques globaux non coûteux en temps de calcul (Duan et al., 1994 ; Madsen, 2003 ; Tolson et Shoemaker, 2007 ; Arsenault et al., 2014 ; Minville et al., 2014). Tandis que d'autres proposent des solutions à la problématique de calage des modèles hydrologiques coûteux en temps de calcul (Tolson et Shoemaker, 2007 ; Shoemaker et al., 2007 ; Razavi et al., 2010 ; Huang et al., 2014 ; Trudel et al., 2014 ; Audet et al., 2014).

Selon Razavi et al. (2010), il existe quatre façons distinctes de classer les approches d'optimisation qui abordent la question de réduction du temps de calcul : (1) l'utilisation ou le développement d'algorithmes d'optimisation plus efficaces, (2) l'exploitation du parallélisme pour l'exécution de plusieurs simulations simultanément, (3) l'emploi de modèles d'évaluations à évasion opportuniste (arrêt prématuré de simulations que l'on identifie comme étant de mauvaise qualité) et (4) l'usage de modèles substituts à même le processus d'optimisation. On retrouve dans la littérature des travaux qui s'attardent indépendamment à l'une ou l'autre de ces quatre façons (Booker et al., 1999 ; Shoemaker et al., 2007 ; Razavi et al., 2010 ; Razavi et al., 2012a ; Regis et Shoemaker, 2013 ; Razavi et Tolson, 2013 ; Huang et al., 2014), mais aussi de manière combinée (Razavi et al., 2010 ; Regis et Shoemaker, 2013 ; Audet, 2014).

Les travaux de Huot (2014), études préliminaires à la présente proposition de recherche, ont démontré l'efficacité des stratégies de certains algorithmes d'optimisation lorsqu'ils étaient employés pour le calage du modèle hydrologique HYDROTEL (distribué, à base plus physique et coûteux en temps de calcul). Plusieurs avenues prometteuses ont été identifiées, dont celle de combiner la capacité de recherche globale de l'algorithme « Dynamically Dimensioned Search » (DDS ; Tolson et Shoemaker, 2007) avec les stratégies de raffinement local de l'algorithme « Mesh Adaptive Direct Search » (MADS ; Audet et Dennis, 2006). L'utilisation de modèles substituts moins coûteux en temps de calcul à même le processus de calage a aussi été l'une des avenues prometteuses identifiées dans ces travaux. Ces deux avenues intéressantes seront étudiées dans le cadre de ce projet de recherche et seront plus amplement détaillées au Chapitre 3 qui traite des objectifs de recherche. En référence aux types d'approches abordant la question de réduction de temps de calcul de Razavi et al. (2010), ces deux avenues prometteuses s'intègrent aux types (1) et (4), soit par le développement d'algorithmes d'optimisation plus efficaces et par l'utilisation des modèles substituts, sans toutefois s'intéresser aux types (2) et (3) : l'emploi du parallélisme et des modèles d'évaluations à évasion opportuniste qui ne seront pas exploités dans les présents travaux.

2.2.1 Algorithmes d'optimisation hybrides

On retrouve dans la littérature plusieurs travaux qui exposent le couplage de différents algorithmes d'optimisation dans le but de résoudre certaines problématiques issues de l'ingénierie. Communément appelés algorithmes d'optimisation couplés ou hybrides, ces méthodes permettent généralement un processus d'optimisation mieux adapté à ces problèmes particuliers (Liu et al., 2004 ; Wetter et Wright, 2004 ; Hasan et al., 2008 ; Audet et al., 2008 ; Kampf et Robinson, 2009 ; Juan et al., 2010 ; Li et al., 2015 ; Junghans et Darde, 2015 ; Garg, 2016). L'utilisation d'une seule méthode d'optimisation existante n'apporte pas toujours les résultats souhaités quant à la valeur de la fonction objectif finale ou encore au temps de calcul, qui s'avère trop important. Par exemple, dans le cadre d'un

problème d'optimisation lié à la thermodynamique du bâtiment, Junghans et Darde (2015) ont démontré que l'utilisation d'un algorithme à recuit simulé couplé à un algorithme génétique pouvait apporter de meilleurs résultats du point de vue de la valeur de la fonction objectif finale comparativement à l'emploi de l'algorithme génétique uniquement. Kampf et Robinson (2009) ont également développé un algorithme hybride efficace combinant le « Covariance Matrix Adaptation Evolution Strategy (CMAES ; Hansen et Ostermeier, 2001) et le « Hybrid Differential Evolution » (HDE ; Chang et al., 2007) pour la résolution de problèmes d'optimisation associés à l'utilisation potentielle de l'énergie solaire captée sur les bâtiments. Plus près du sujet des présents travaux, Liu et al. (2004) ont démontré qu'il était possible de réduire les temps de calcul de 70% tout en ayant des valeurs de fonction objectif finales comparables lorsque le « K-Nearest Neighboor Classifier » (KNN ; Dasarathy, 1990) était jumelé à l'algorithme « Non-Dominated Sorting Genetic Algorithm II » (NSGA II ; Deb et al., 2002) pour le calage de modèle hydrologique. Aussi, le populaire « Particle Swarm Optimization » (PSO) de Eberhart et Kennedy (1995) a été couplé à d'autres algorithmes d'optimisation existants comme le « Hooke and Jeeves Algorithm » (HJ ; Hooke et Jeeves, 1961) ou encore avec le « Artificial Bee Colony » (ABC ; Karaboga, 2005) dans le but de fournir des processus d'optimisation plus adapté à certains problèmes issus de la thermodynamique du bâtiment et de systèmes d'alimentation électrique (Wetter et Wright, 2004 ; Hasan et al., 2008 ; Li et al., 2015). Bien que l'algorithme PSO soit efficace pour explorer l'espace de solutions et identifier des jeux de paramètres de bonne qualité, cette méthode populaire peut quelques fois rester captive dans des zones d'optima locaux de moins bonne qualité. Il est donc souvent employé pour ses capacités de recherche globale et l'objectif de couplage à un autre algorithme d'optimisation comme le HJ ou le ABC permet de diminuer ce genre d'inconvénient (Hasan et al., 2008 ; Li et al., 2015). Tous ces travaux démontrent alors que l'hybridation d'algorithmes existants présente un intérêt pour les utilisateurs lorsqu'il s'agit de rendre les processus d'optimisation plus efficaces.

2.2.2 Utilisation de modèles substituts

Razavi et al. (2012a) ont défini un modèle substitut ou « surrogate » comme étant une approximation simplifiée d'un modèle de simulation et ayant la caractéristique d'être moins coûteux en temps de calcul. Le Digabel (2011), pour sa part, définit un « surrogate » comme une fonction qui partage des similarités avec la fonction objectif d'un modèle de simulation, est moins coûteuse en temps de calcul et est utile à exploiter dans le but de réduire le nombre d'évaluations du modèle de simulation. On retrouve principalement deux façons de concevoir des modèles substituts : l'utilisation de fonctions de surface conçues à partir de l'évaluation d'un historique de points de l'espace des solutions connus (jeux de paramètres) et la création a priori de modèles à fidélité réduite (Regis et Shoemaker, 2007 ; Razavi et al., 2012a). Le Digabel (2011) distingue aussi ces deux types de substituts : les fonctions adaptatives (ou fonctions de surface), et les fonctions non-adaptatives (ou modèles à fidélité réduite). D'une part, les fonctions de surface exploitent l'historique des jeux de paramètres déjà évalués en appliquant des techniques d'interpolation empirique dans le but de reproduire le plus fidèlement possible la fonction objectif (ou fonction à optimiser) originale. Il est également possible d'améliorer la qualité de la fonction de surface en effectuant des mises à jour de celle-ci en ajoutant ou en retirant (rarement) des jeux de paramètres évalués en cours d'optimisation, d'où le qualificatif de fonctions adaptatives. Le principe de mise à jour est particulièrement intéressant dans le cadre des fonctions adaptatives puisqu'il permet, soit de construire préalablement au processus d'optimisation un substitut et ensuite de procéder à un ajustement en continu, soit de construire entièrement le substitut au fur et à mesure que le processus d'optimisation progresse (Razavi et al., 2012a ; Razavi et al., 2012b). Parmi les techniques de fonctions de surface, on retrouve entre autres les techniques de krigeage (Krige, 1951 ; Lophaven et al., 2002), les fonctions polynomiales (Hussain et al., 2002 ; Fen et al., 2009), les « Radial Basis Functions » (RBF ; Hussain et al., 2002 ; Mugunthan et al., 2005 ; Regis et Shoemaker, 2007), les réseaux de neurones (*artificial neural networks* – ANNs ; Papadrakakis et al., 1998) et les « Support Vector Machines » (SVMs ; Zhang et al., 2009). D'autre part, les modèles à fidélité réduite sont indépendants de la connaissance d'un historique d'évaluations et représentent plutôt une simplification de la structure du modèle

original dans le but de conserver une représentativité (Regis et Shoemaker, 2007 ; Razavi et al., 2012a ; Razavi et Tolson, 2013). Il n'est donc pas possible de les mettre à jour durant le processus d'évaluation de la même façon que les fonctions adaptatives. La création d'un modèle à fidélité réduite est donc propre à chaque problème d'optimisation et une connaissance accrue de ce dernier est nécessaire pour être en mesure de construire un substitut qui est le moins coûteux possible en temps de calcul, mais qui conserve une bonne représentativité du modèle original. Certaines études utilisent des pistes de solution intéressantes pour la création de modèles à fidélité réduite qui sont ensuite exploités au sein des processus de calage. Razavi et Tolson (2013) ont présenté une façon originale de concevoir un substitut en utilisant une copie du modèle hydrologique original ayant toutefois une période de calage largement écourtée, mais représentative de la période de calage du modèle original. Haghnegahdar et al. (2015) ont quant à eux étudié divers niveaux de discrétisations spatiales (échelles de modélisation) dans le but d'évaluer l'impact sur la valeur de la fonction objectif finale par rapport aux temps de calage engendrer. Bien qu'ils n'aient pas explicitement démontré l'utilisation de modèles à fidélité réduite au sein d'un processus d'optimisation, leurs travaux apportent une avenue intéressante quant à la création de substitut représentatif.

On retrouve dans la littérature plusieurs utilisations des modèles substituts qui traitent des problèmes coûteux en temps de calcul dans des applications issues de l'ingénierie et même, plus spécifiquement, issues de la modélisation hydrologique ou environnementale. Par exemple, les travaux de Dennis et Torczon (1995) ainsi que ceux de Booker et al. (1999) ont proposé des cadres d'exploitation des substituts à même les processus d'optimisation forts pertinents quant à la résolution de problèmes concernant la conception des pales de rotor d'hélicoptère. Plus près du domaine de cette étude, Regis et Shoemaker (2004 ; 2007 ; 2009 ; 2013) ont grandement mis à contribution l'exploitation de fonctions de surface telles que les RBF au sein de processus de calage de modèles hydrologiques et de restauration des eaux souterraines. Ils ont ainsi démontré qu'il était possible d'améliorer les valeurs de la fonction objectif finale tout en diminuant les temps de calcul onéreux. Khu et al. (2004) ont quant à eux démontré que l'utilisation des ANNs à même le processus de calage d'un modèle

hydrologique permettait de réduire jusqu'à 40% le temps de calcul en comparaison avec un algorithme d'optimisation métaheuristique traditionnel. Razavi et al. (2012b) ont également publié une étude comparative de trois différentes fonctions de surface (RBF, ANNs, Krigeage) qui indique que l'emploi de certaines fonctions de surface peut amener une meilleure performance comparativement à des méthodes de calage existantes dans un contexte où l'on souhaite diminuer le temps de calcul du processus de calage. Dans un dernier exemple, Razavi et Tolson (2013) ont démontré dans leurs travaux que l'utilisation d'un modèle hydrologique reposant sur un modèle à fidélité réduite caractérisé par une période de simulation plus courte, mais soigneusement choisie, permet une potentielle diminution du temps de calcul.

Également, certains algorithmes existants emploient d'emblée les fonctions substituts à même les stratégies d'optimisation choisies dans le but de fournir une meilleure efficacité. L'algorithme d'optimisation MADS (Audet et Dennis, 2006) emploie d'ailleurs une fonction quadratique pour représenter localement l'espace paramétrique dans le but d'ordonnancer le classement des points qui seront par la suite évalués au sein du problème d'optimisation (Conn et Le Digabel, 2013). Tous ces travaux ici énumérés ont démontré l'avantage de l'utilisation des substituts dans une perspective d'amélioration de l'efficacité des méthodes d'optimisation et c'est pourquoi ce projet de recherche sera consacré en grande partie au domaine des fonctions substituts.

2.3 Problématique de recherche

Suite à la revue de la littérature présentée précédemment, il est clair que le calage des modèles hydrologiques distribués coûteux en temps de calcul constitue une problématique réelle à laquelle il est pertinent de s'intéresser. Bien que plusieurs études s'attaquent à la problématique de consommation importante de temps de calcul lors de processus de calage, il est difficile d'identifier des méthodes d'optimisation qui offrent à la fois : (1) une robustesse d'optimisation garantissant le plus souvent possible l'obtention d'un jeu de paramètres final de bonne qualité, (2) une réduction significative du temps de calcul (diminution accrue du

nombre de simulations effectuées par l'algorithme) et (3) une approche d'optimisation qui combine les avantages de plusieurs techniques de réduction du temps de calcul tels qu'identifié par Razavi et al. (2010). Les objectifs de cette étude sont définis dans le but de répondre à ces trois critères et d'ainsi développer et fournir de nouvelles approches d'optimisation répondant à cette problématique.

De manière à bien comprendre les diverses analyses et évaluations qui seront effectuées dans ces travaux, il est important de définir correctement ce que représente l'efficacité au sein de la problématique du calage des modèles hydrologiques coûteux en temps de calcul. Lors de la mesure de l'efficacité d'un processus d'optimisation, il va de soi que l'on s'intéresse, d'une part, à la qualité de la fonction objectif finale trouvée comme critère de performance puisque c'est en fonction de ce critère que l'on permet à l'algorithme d'optimisation d'aboutir à une réponse finale de bonne qualité. D'autre part, dans un contexte où le temps de simulation du modèle hydrologique a un impact direct sur l'utilisation pratique de cet outil, il devient inévitable de concevoir le temps de calcul comme un second indicateur de l'efficacité. Dans un contexte d'optimisation comme le calage de modèles hydrologiques coûteux en temps de calcul, l'efficacité sera définie comme étant la capacité à effectuer le moins grand nombre de simulations du modèle hydrologique possible, de manière à diminuer le temps de calage, tout en assurant l'obtention d'un jeu de paramètres final de bonne qualité. Ainsi, le calage implique le sacrifice d'une part de qualité de la fonction objectif finale pour obtenir un gain en temps de calcul, de la même manière qu'une diminution du temps de calcul aura un impact de dépréciation sur la qualité de la fonction objectif finale. C'est donc le compromis entre le temps de calcul et la qualité de la solution finale qui permettra d'évaluer l'efficacité des méthodes d'optimisation développées dans le cadre de ces travaux.

CHAPITRE 3

OBJECTIFS DE LA RECHERCHE

Ce chapitre expose les objectifs principaux de cette thèse déterminés en fonction de la problématique de recherche définie précédemment.

3.1 Objectifs

Cette étude a pour but de répondre à trois objectifs principaux en lien avec la problématique du calage efficace des modèles hydrologiques coûteux en temps de calcul. Ces objectifs ont été identifiés suite à la définition de la problématique et à la revue de la littérature qui ont été précédemment exposées. Le premier article fait état du premier objectif de cette étude qui vise à développer une nouvelle approche d'optimisation adaptée au contexte de calage des modèles hydrologiques coûteux en temps de calcul. Les deux articles suivants s'orientent plutôt vers les deux autres objectifs, soient l'évaluation du potentiel d'utilisation de divers modèles substituts et l'utilisation efficace de ceux-ci au sein d'un processus de calage. Ces deux articles ont d'ailleurs été rédigés de pair pour répondre et diviser adéquatement les deux objectifs traités.

3.1.1 Développement d'une nouvelle technique d'optimisation hybride DDS-MADS

Le premier objectif visent le développement d'une nouvelle méthode d'optimisation offrant un compromis avantageux entre la diminution du nombre d'évaluations du modèle hydrologique (indicateur du temps de calcul), que l'on souhaite le plus faible possible, et la livraison d'un jeu de paramètres final de bonne qualité (valeur de la fonction objectif), que l'on souhaite la meilleure possible (définition de l'efficacité mentionnée à la Section 2.2). Cette nouvelle technique d'optimisation vise à fournir une amélioration, en termes d'efficacité, du processus de calage du modèle hydrologique HYDROTEL par rapport à ce que l'on obtient avec des algorithmes d'optimisation existants. La Figure 3.1 illustre l'association des deux algorithmes en une technique d'optimisation hybride qui exploite les forces des algorithmes DDS et MADS tel qu'identifiées dans les travaux de Huot (2014).

D'une part, la capacité de recherche globale de DDS permet l'obtention rapide d'une solution de bonne qualité. D'autre part, le processus de raffinement local de MADS qui amène à une convergence vers un optimum local permet l'arrêt automatique du calage en fonction de la qualité de la fonction objectif.. Dans le cadre de cet objectif, le défi réside plus particulièrement dans la manière de jumeler les deux algorithmes, de façon à ce que l'approche d'optimisation finale soit fluide, simple et réponde bien à la problématique de cette étude. Pour ce faire, cinq fonctionnalités ont été implémentées aux deux algorithmes et sont clairement expliquées au Chapitre 4 (Article 1). Les résultats finaux démontrent bien les gains obtenus en temps de calcul lors de l'utilisation de cette approche hybride tout en conservant une valeur de la fonction objectif finale de bonne qualité.

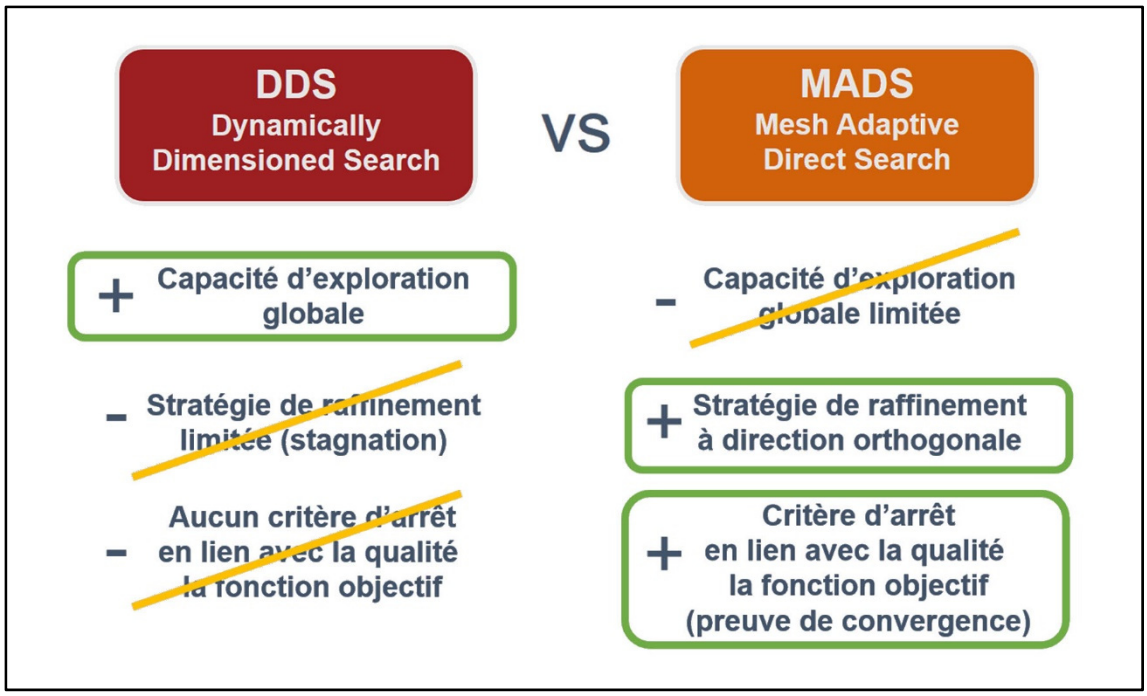


Figure 3.1. Présentation des stratégies d'optimisation avantageuses (encerclées en vert) des algorithmes DDS et MADS pour la mise en place d'une approche de calage hybride efficace en temps de calcul

3.1.2 Évaluation du potentiel d'utilisation des modèles substituts à même le processus de calage

Comme il a été mentionné précédemment, l'emploi de modèles substituts au cours du processus d'optimisation peut apporter une économie intéressante sur le temps de calcul. Cependant, l'intégration de ces modèles substituts à un processus efficace de calage n'est pas chose simple puisqu'elle demande certaines étapes préliminaires pour s'assurer d'un niveau élevé de représentativité des substituts.

Dans le but d'arriver à fournir une ou plusieurs méthodes d'optimisation intégrant ceux-ci, la possibilité de créer des modèles substituts représentatifs sera d'abord examinée. D'une part, les modèles à fidélité réduite seront examinés dans le but de connaître les avenues potentielles qui offrent un substitut représentatif et moins coûteux en temps de calcul. Trois différentes avenues de modèles à fidélité réduite, illustrées à la Figure 3.2, seront explorées, évaluées et combinées à partir du modèle hydrologique HYDROTEL (modèle original), qui est coûteux en temps de calcul. La première avenue est celle de la réduction du nombre de pseudo-stations météorologiques qui se retrouve sur le territoire modélisé (Figure 3.2 A), la seconde fait état de la réduction de la durée de la période de calage (Figure 3.2 B) et la dernière est la réduction du degré de discrétisation spatiale de la modélisation (Figure 3.2 C) ; c'est-à-dire la diminution du nombre d'unités de simulation (sous-bassins versants élémentaires). Tout en conservant l'objectif d'une représentativité élevée, ces trois avenues seront également combinées dans le but d'augmenter le rapport entre le temps de calcul d'une simulation du modèle hydrologique original et celui d'une simulation du modèle substitut (Figure 3.2 D). L'Article 2 présenté au Chapitre 5 se consacre en totalité à l'évaluation et la mise en place des trois diverses avenues de construction des modèles à fidélité réduite ainsi que de de leur combinaison. Les résultats et analyses qui seront obtenus sur ces modèles à fidélité réduite permettront de faire état des caractéristiques avantageuses de chacune. Celles-ci mèneront à la réalisation du troisième objectif.

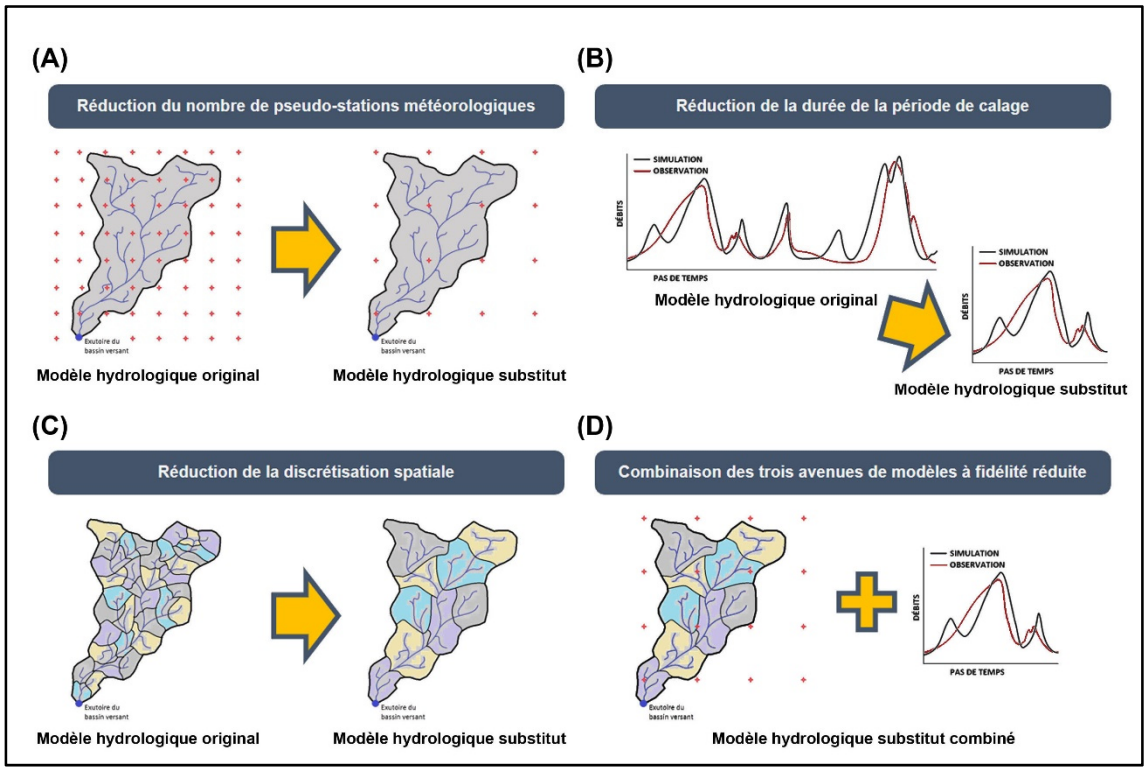


Figure 3.2. Identification des trois avenues de modèles à fidélité réduite examinées et de leur combinaison (D) : réduction du nombre de pseudo-stations météorologiques (A), réduction de la durée de la période de calage (B) et réduction du degré de discrétisation spatiale (C)

D'autre part, les fonctions de surface seront étudiées de manière à identifier le nombre d'évaluations nécessaires (jeux de paramètres à évaluer) qui servira à la confection d'un modèle représentatif, mais également à connaître la taille de l'espace paramétrique qui permet aux modèles substituts de conserver un niveau de représentativité acceptable. Deux fonctions de surfaces seront examinées, soient les fonctions polynomiales (Hussain et al., 2002 ; Fen et al., 2009) et les modèles de krigeage (Krige, 1951 ; Lophaven et al., 2002). Bien que les fonctions de surface ne sont pas couvertes dans l'un ou l'autre des trois articles intégrant cette thèse, le Chapitre 7 traitera des expérimentations des fonctions polynomiales et des modèles de krigeage et une discussion comparative entre les deux fonctions de surface suivra.

3.1.3 Utilisation efficace des modèles substituts au sein d'un processus de calage coûteux en temps de calcul

Une fois l'évaluation et la mise en place des modèles substituts complétées, cet objectif vise à identifier pour les modèles substituts la façon dont ces derniers peuvent être efficacement exploités au sein d'un cycle de calage. La Figure 3.3 présente un cycle de calage qui intègre l'exploitation des modèles substituts. L'algorithme d'optimisation a alors l'option de simuler un nouveau jeu de paramètres (1) au sein du modèle hydrologique original coûteux en temps de calcul qui fournira alors une valeur de fonction objectif juste, ou (2) au sein du modèle substitut qui fournira rapidement une valeur estimée de la fonction objectif. Autrement dit, la procédure d'optimisation doit permettre d'identifier quand et/ou comment les modèles substituts peuvent intervenir au sein du processus d'optimisation de manière à augmenter l'efficacité en temps de calcul sans toutefois diminuer le niveau de qualité de la valeur finale de la fonction objectif. De manière à identifier la réponse au quand et/ou comment, la nouvelle approche d'optimisation hybride DDS-MADS développée dans le cadre du premier objectif sera l'approche qui servira à exploiter les modèles substituts développés. Finalement, plusieurs cadres de calibration combinant l'efficacité des stratégies d'optimisation employées par l'approche hybride DDS-MADS et l'exploitation des modèles à fidélité réduite seront présentés. L'Article 3 au Chapitre 5 proposera un éventail de cadres de calibration adaptés à la problématique de modèles hydrologiques coûteux en temps de calcul. En fonction des objectifs de calage et des contraintes d'optimisation, l'utilisateur pourra alors préconiser le cadre de calibration qui répond le mieux à ses besoins.

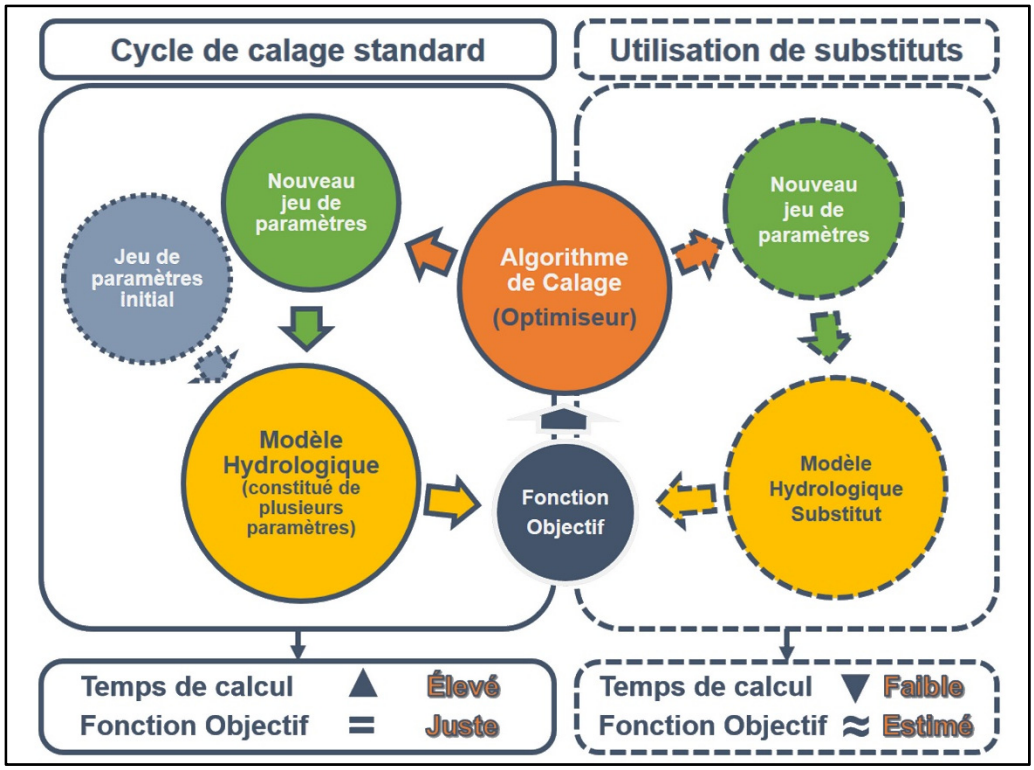


Figure 3.3. Cycle de calage qui intègre l'exploitation des modèles substituts

CHAPITRE 4

A HYBRID OPTIMIZATION APPROACH FOR EFFICIENT CALIBRATION OF COMPUTATIONALLY INTENSIVE HYDROLOGICAL MODELS

Pierre-Luc Huot[1], Annie Poulin[1], Charles Audet[2] et Stéphane Alarie[3]

[1] Département de la Génie de la Construction, École de technologie supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3.
[2] Département de Mathématiques et Génie Industriel, Polytechnique de Montréal et GERAD,
2900 boul. Édouard-Montpetit, Montréal, Québec, Canada H3T 1J4.
[3] Institut de Recherche d'Hydro-Québec,
1800 boul. Lionel Boulet, Varennes, Québec, Canada J3X 1S1.

Article soumis pour publication dans « Hydrological Sciences Journal », juillet 2018.

Abstract

The calibration of hydrological models is formulated as a blackbox optimization problem where the only information available is the objective function value. Distributed hydrological models are generally computationally intensive, and their calibration may require several hours or days which can be an issue for many operational contexts. Different optimization algorithms have been developed over the years and exhibit different strengths when applied to the calibration of computationally intensive hydrological models. This paper shows how the Dynamically Dimensioned Search (DDS) and the Mesh Adaptive Direct Search (MADS) algorithms can be combined to significantly reduce the computational time of distributed hydrological model calibrations while ensuring robustness and stability regarding the final objective function values. Five transitional features are described to adequately merge both algorithms. The hybrid approach is applied to the distributed and computationally intensive HYDROTEL model on three different river basins located in Québec (Canada).

Keywords: Distributed hydrological model, Computationally-intensive simulation model, Optimization algorithm, Efficient calibration, Dynamically Dimensioned Search, Mesh Adaptive Direct Search.

4.1 Introduction

Hydrological catchment models are simulation tools built with a sequence of mathematical equations to represent hydrological processes such as rainfall, snow accumulation and melt, evapotranspiration, infiltration and runoff at a catchment-area scale (Singh and Woolhiser 2002). Hydrological models are used to reproduce the relations between water inflows, such as precipitation, and the discharge at the catchment outlet. They require representative meteorological time series and a good knowledge of catchment characteristics to adequately simulate streamflow (Singh and Woolhiser 2002, Moradkhani and Sorooshian 2008). All of these characteristics, meteorological data time series and hydrological processes can be integrated within the models according to different spatialization scales, corresponding to the level of a territory's subdivisions. Hydrological models can be classified into three main categories according to their level of spatial discretization: lumped models that represent the catchment area as a single spatial entity; semi-distributed models that roughly divide the catchment area into sub-watersheds/catchments; and distributed models where the area is organized into fine mesh elements with reduced size (Singh and Woolhiser 2002, Das et al. 2008, Moradkhani and Sorooshian 2008, Pechlivanidis et al. 2011).

Hydrological models are fundamental tools for many engineering applications: hydroelectricity production, water supply, delineation of flood zones, etc. (Beven 2001, Singh and Woolhiser 2002). For all these application fields, distributed modelling provides a better representation of spatial variability in the catchment than lumped models. The distributed models have the advantage of representing some hydrological processes that are impossible to assess otherwise, such as the spatial distribution of snow accumulation on the catchment, or the transport of pollutants and sediments in the water system (Moradkhani and Sorooshian 2008). They are also particularly useful for glacier coverage analysis over mountainous catchments and for the analysis of hydrological response in ungauged watersheds when a fine mesh scale is required (Verbunt et al. 2003, Moradkhani and Sorooshian 2008, Lei et al. 2011, Huang et al. 2014).

Hydrological models need to be adapted to each watershed by adjusting the internal model parameters of the equations governing the simulation of the hydrological processes, through model calibration. This leads to the identification of a model parameter set minimizing a measure of the difference between the observed and predicted/simulated variables, generally the discharge at the catchment outlet (Duan et al. 1994, Singh and Woolhiser 2002, Moriasi et al. 2007, Moradkhani and Sorooshian 2008, Pechlivanidis et al. 2011, Pushpalatha et al. 2012). Various combinations of model parameter sets are explored during the calibration procedure. The use of an objective criterion allows finding a combination of model parameters that satisfy the user's needs (Moriasi et al. 2007). Good calibration is critical since it is the key to users' confidence in the model. The calibrated model will then be used as a tool to support management decisions on water resources or infrastructures built in/around water environment.

4.1.1 Literature review

Hydrological model calibration belongs to the class of "blackbox optimization" problems (Audet 2014, Audet and Hare 2017) in which the function to be optimized and the function that defines the feasible region of the model parameter space (delimited by user-defined boundaries) are not analytically known. These functions are evaluated by launching a computer simulation, which can be time-consuming, mainly in the case of distributed models. Due to the complexity of simulations, the response surfaces of such models generated by an objective criterion are nonsmooth, discontinuous, and contaminated by numerical noises. The term "blackbox" refers to the fact that the exploitation of the internal structure of the optimization problem is difficult, inefficient or simply impossible. The only information used by the optimization procedure is the output produced by the hydrological model simulation, generally the value of the objective criterion. Consequently, a combination of multiple hydrological simulations and sufficiently efficient optimization strategies is essential to obtain a fine-tuned model (Conn et al. 2009, Shan and Wang 2010, Audet et al. 2014, Boukouvala et al. 2016, Audet and Hare 2017).

In recent decades, important developments have been made in hydrological modelling, including growth in computational power. However, the use of distributed models still generates an intensive computational cost, especially for the calibration procedure. Running a single simulation may take several minutes with distributed models, and the calibration procedure may require more than thousands of simulations. Intensive computational time is subject to the needs for which hydrological modelling is being used and depends on computational power available and utilization context. Some operational contexts, such as flow forecasting at reservoir for hydropower production or flood streamflow monitoring for civil security, need weekly or daily recalibration. These typical cases could lead to many daily complications for decision makers. NOAA (National Oceanic and Atmospheric Administration - www.noaa.gov) reports that there are many reasons why users must recalibrate hydrological models: considerable experience gain on hydrological modelling, identification of situations or problems improperly modelled, significant changes on land use (new agricultural practices, large forest fire or new urbanization project), establishment of new forecast points on the catchment, addition of new historical input data (precipitations or temperature, for instance) for hydrological forecasting (Anderson 2002).

The time consumed by a simulation (thus a calibration procedure) depends on the mathematical equations governing the hydrological processes, the level of spatial discretization, the input data size, the number of internal model parameters to be calibrated and the scale of the territory modelled (Mungunthan et al. 2005, Zhang et al. 2009, Razavi et al. 2010). Examples of time-consuming hydrological models can be found in Razavi et al. (2010): a single simulation of the Dipole Flow and Reactive Tracer Test – Interpretation Model (DFRTT-IM; Roos 2009) takes about 37 minutes on average on a 2.8 GHz Intel Premium processor. The Soil and Water Assessment Tool, version 2000 (SWAT2000; Neitsch et al. 2002) consumes 1 minute on average for a single simulation (5.25 years on a daily time-step) on a 2.8 GHz Pentium processor.

Various optimization algorithms are based on standard heuristic methods, such as the "Shuffled Complex Evolution method of the University of Arizona" (SCE-UA; Duan et al.

1993), the “Adaptive Simulated Annealing” (ASA; Ingber 1993), the “Particle Swarm Optimization” (PSO; Eberhart and Kennedy 1995), the “Covariance Matrix Adaptation Evolution Strategy” (CMAES; Hansen 2006) or the “Dynamically Dimensioned Search” (DDS; Tolson and Shoemaker 2007), to name just a few popular ones. All heuristic based optimization methods can yield promptly good-quality or near-optimal solutions to solve various optimization problems, but do not guarantee convergence to local nor to global optimal solutions (Rardin and Uzsoy 2001). Direct-search methods (Torczon 1997, Audet and Dennis 2006) are not heuristics since they are supported by a convergence analysis, and can be used to optimize blackbox problems. All these algorithms provide a “state-of-the-art” for hydrological model calibration. In the past few decades, hydrologists have studied their optimization strategies, evaluated their efficiency, applied them in many case studies and adopted them in operational contexts. The efficiency of many of these optimizers has been shown when used for the calibration of non-computationally intensive hydrological models (e.g. Duan et al. 1994, Madsen 2003, Tolson and Shoemaker 2007, Arsenault et al. 2014, Minville et al. 2014). However, the development of these optimizers is not always oriented towards a high computational time issue, and so efficient optimizers must be chosen for this kind of problem.

According to Razavi et al. (2010), four types of approaches can be considered for dealing with the issue of computationally intensive problems: (1) the development of more computationally efficient optimization algorithms; (2) the use of parallel computing networks for running several simulations simultaneously; (3) the use of model preemption (evading model simulations that are identified as being of poor-quality); and (4) the use of surrogate models inside the optimization procedure. These approaches are complementary and can either be combined (Razavi et al. 2010, Regis and Shoemaker 2013, Audet et al. 2008) or used separately (Booker et al. 1999, Shoemaker et al. 2007, Razavi et al. 2010, Razavi et al. 2012, Regis and Shoemaker 2013, Razavi and Tolson 2013, Huang et al. 2014, Maier et al. 2014, Tsoukala et al. 2016).

4.1.2 Research objectives and contributions

This study focuses on the first type of approaches identified by Razavi et al. (2010); i.e. the development of more computationally efficient optimization methods. The present study proposes to combine the strengths of two existing optimization methods for the calibration of computationally intensive hydrological models. The objectives of this new optimization approach are: (1) to significantly reduce the number of model simulations to proceed for the calibration procedure, while (2) achieving good-quality calibration results, (3) to show stability and robustness in terms of the localization of local and/or global optima solutions, based on the satisfaction of optimality conditions and (4) to propose a straightforward calibration approach requiring no specific knowledge or sensitivity analysis regarding the optimization problem.

A preliminary step assesses the efficiency of three benchmark optimization algorithms for the calibration of a distributed and computationally intensive hydrological model. The results highlight the optimization strategies that show real potential for reducing the computational time. Two algorithms are then merged to propose a hybrid approach exploiting the previously identified efficient strategies. The main contribution of this paper pertains to the transition features, resulting in an efficient operating optimization framework. The hybrid calibration approach is based on many intermediate experiments completed on two modelled watersheds. A third watershed is introduced to avoid overfit and used for validating the hybrid optimization approach.

4.1.3 Paper organization

This paper is organized as follows. Section 4.2 presents the hydrological model HYDROTEL, followed by the formulation of two optimization problems and the modelled watersheds of this study. Section 4.3 benchmarks three optimization algorithms to identify the optimization strategies that can reduce the computational time. Section 4.4 proposes the detailed operating framework of the hybrid calibration approach, followed by numerical results on two experimental watersheds. Section 4.5 validates the hybrid approach by

calibrating a third modelled watershed, and also validates the final model parameter sets by testing them on a different simulation time period. Discussion and proposed future work conclude the paper.

4.2 Calibration of hydrological models

The calibration of hydrological models is here formulated as a blackbox optimization problem where the only information available is the objective function value. This section describes the three components that define the blackbox calibration problems: the hydrological model, the objective function and the studied watersheds.

4.2.1 Model description

The distributed and physically-based computationally intensive hydrological model HYDROTEL (Fortin et al. 2001a, 2001b, Turcotte et al. 2003, Turcotte et al. 2007, Bouda et al. 2012, Bouda et al. 2014) is used in this study. HYDROTEL has been used operationally by the DEH (Direction de l'Expertise Hydrique, Québec, Canada) for flow forecasting and monitoring for more than 10 years. In this operational context, the model is applied to natural flow watersheds, covering a wide range of surface areas, and it needs to be recalibrated on a regular (up to daily) basis, especially during the snowmelt (spring flood) period (MELCC 2018). The spatial discretization of the territory is represented by several simulation units named RHHUs (Relatively Homogenous Hydrological Units), and all hydrological processes are simulated on each RHHU. This hydrological model is mainly designed to account for the spatial variability of hydrological processes and to exploit Geographic Information System (GIS) and/or remote sensing data providing topography, land cover, land use, natural waterways and soil type, to name just a few examples.

HYDROTEL has 27 internal model parameters which can be all calibrated independently on each RHHU. However, all internal model parameters are grouped at the basin scale to limit the number of calibration parameters while ensuring spatial consistency of distributed parametric information and avoiding spatial discontinuities due to equifinality (Ricard et al.

2013). According to previous studies (Fortin et al. 2001a, Turcotte et al. 2003, Turcotte et al. 2007, Poulin et al. 2011, Ricard et al. 2013), some model parameters can be fixed at first, as they generally have a low sensitivity impact on simulation results or they are purely additive or corrective factors applied to the input data. The impact of model parameters on the objective function value has also been studied by Huot (2014) with the Morris global sensitivity analysis method (Morris 1991). Based on these studies, two optimization problems were designed, one with 10 model parameters and one with 19 model parameters (hereafter referred to as HYDROTEL 10 and HYDROTEL 19, respectively). Both problems allow the assessment of the impacts of different dimensionalities, levels of complexity and levels of equifinality. Table 4.1 shows the selected model parameters and the boundaries for each model parameter included within HYDROTEL 10 and HYDROTEL 19.

Table 4.1. Summary of selected model parameters and boundaries for HYDROTEL 10 and HYDROTEL 19

Submodels Used	N°	HYDROTEL Model Parameters	Units [a]	HYDROTEL 10 [b]	HYDROTEL 19 [b]	Lower Bound	Upper Bound
Thiessen Polygons	1	Lapse Rate of precipitation	mm/day		X	0	2
	2	Lapse Rate of Temperature	°C/100m		X	-0,8	-0,3
	3	Base Refreezing Temperature	°C	X	X	-6	-4
Mixed (degree-day) energy-budget method	4	Melt Rate Coefficient at the Soil-Snow Interface	mm/day		X	0,4	0,6
	5	Maximum Density of Snow Cover	kg/m³		X	300	550
	6	Ratio of Snow Compaction			X	0,005	0,015
	7	Degree-Day-Factor - Conifers Trees	mm/day/°C	X	X	0,01	6,39
	8	Temperature Threshold for Melt - Conifers Trees	°C	X	X	0,01	1,99
	9	Degree-Day-Factor - Deciduous Trees	mm/day/°C		X	0,01	3,99
	10	Temperature Threshold for Melt - Deciduous Trees	°C		X	0,01	1,99
	11	Degree-Day-Factor - Non-Forest Environments	mm/day/°C	X	X	0,01	3,99
	12	Temperature Threshold for Melt - Non-Forest Environments	°C	X	X	-6	1,59
Hydro-Quebec method	13	Potential Evapotranspiration Factor		X	X	0,7	1,2
BV3C	14	Lower limit of soil layer #1	m	X	X	0,05	1
	15	Lower limit of soil layer #2 [c]	m	X	X	0,05	1
	16	Lower limit of soil layer #3 [d]	m	X	X	0,05	4
	17	Recession Coefficient	m/hour	X	X	1E-08	0,0001
Kinematic Wave Equation	18	Manning's Roughness Coefficient - Forest Cover			X	0,04	0,3
	19	Depth of Runoff of the Geomorphological Hydrograph	m		X	0,0002	0,002

[a] Blank spaces mean no unit

[b] Blank spaces mean fixed parameter

[c] Lower Limit of Soil Layer #2 = $0{,}05 \leq x15 \leq 1 + x14$

[d] Lower Limit of Soil Layer #3 = $0{,}05 \leq x16 \leq 4 + x15$

4.2.2 Objective function

To assess the quality of each model simulation, the objective function (to be minimized) is computed between simulated and observed streamflows as 1 minus the Nash-Sutcliffe Efficiency criterion ($1 - NSE$), described as follows:

$$1 - NSE = \frac{\sum_{i=1}^{n}(O_i - S_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \tag{4.1}$$

where O_i is the ith observed streamflow value, S_i is the ith simulated streamflow value, $\bar{O}$ is the mean of observed streamflows and n is the total number of observations. $1 - NSE$ ranges from 0 to ∞, with $1 - NSE = 0$ indicating a perfect fit between observed and simulated values.

There are alternatives to using $1 - NSE$ as objective function. This function has been frequently used in hydrological modelling (Servat and Dezetter 1991, Moriasi et al. 2007) and in model calibration studies (Tolson and Shoemaker 2007, Poulin et al. 2011, Ricard et al. 2013, Arsenault et al. 2014, Poissant et al. 2017, Bajamgnigni Gbambie et al. 2017). The algorithms considered hereafter are generic and are applicable to any objective function. Our choice is made here by convenience and its purpose is only about comparing the relative performance of algorithms. Alternative objective functions, such as the Kling-Gupta Efficiency criterion (KGE; Gupta et al. 2009) or the Root Mean Squared Error (RMSE), could be used if preferred.

4.2.3 Studied watersheds

As part of this study, both calibration problems have been applied to three different watersheds with dissimilar characteristics and located in different regions of Québec (Canada). Figure 4.1 presents the drainage maps of the three river basins with the geolocation of the studied watersheds (grey areas). The meteorological data required as input to the hydrological model includes daily minimum temperature, daily maximum temperature, and

daily precipitation. Daily observed streamflows at the outlets of the three studied watersheds were also required to assess the quality of simulations. The watersheds are:

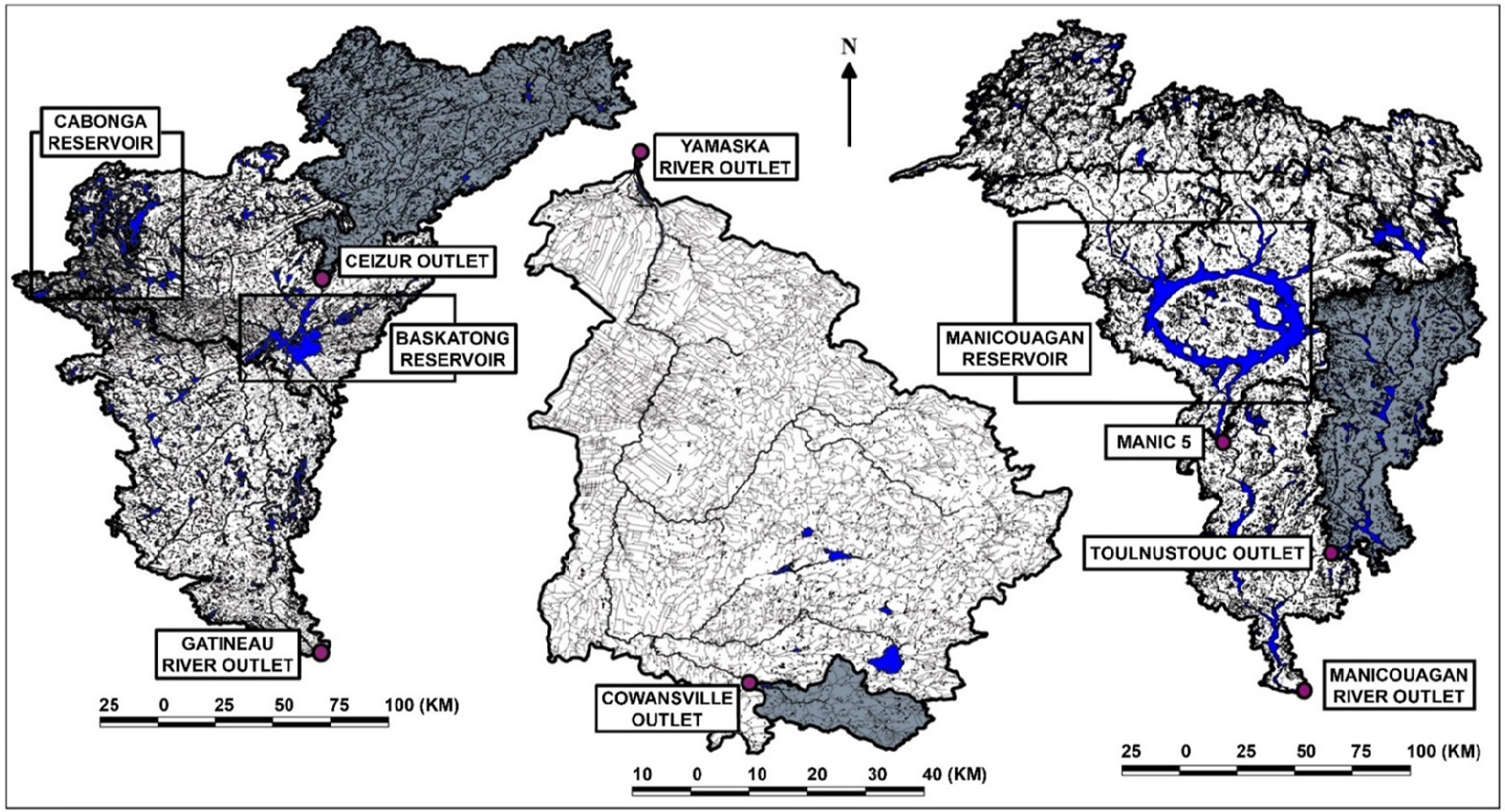
- **Cowansville.** It is an upstream sub-basin of the Yamaska River. Located on the south shore of the St. Lawrence River in the Montérégie region (southern Québec), Cowansville has a small area (215 km²) and is mainly covered by forest and agriculture. Meteorological data were obtained from a gridded dataset at 20-km resolution from DEH (see the acknowledgments section) and the calibration period spans from October 1st, 2000 to September 30, 2005. Daily observed streamflows for the same period were also obtained from DEH. Running a single Cowansville simulation over a 5 years long time-period on a 3.40 GHz Intel Core i7 processor (12 Go of RAM) takes between 55 s to 1 min 20 s.

- **Ceizur.** It is an upstream sub-basin of the Gatineau River basin in the Outaouais region (western Québec), which is mainly forested and has a large area (6,928 km²). Hydro-Québec provided a meteorological gridded dataset at 10-km resolution and daily observed streamflows at the outlet of the sub-basin (see the acknowledgments section). The calibration period is from October 1st, 1988 to September 30, 1992. Running a single Ceizur simulation over a 4 years long time-period on a 3.40 GHz Intel Core i7 processor (12 Go of RAM) takes between 1 min 30 s to 1 min 55 s.

- **Toulnustouc.** This sub-basin is part of the Manicouagan River basin located in the Côte-Nord region (north-eastern Québec), one of the most important rivers for hydroelectric production in Québec (power station Manic 5 is positioned in Figure 4.1). It is an upstream forested sub-basin and has the largest area (8,109 km²) of the studied watersheds. The calibration period spans from October 1st, 1984 to September 30, 1988, and is denoted by a meteorological gridded dataset at 20-km resolution derived from a NLWIS (National Land and Water Information Service – www.agr.qc.ca) dataset and daily observed streamflows are from Hydro-Québec (see the acknowledgement section). Running a single Toulnustouc simulation over a 4 years

long time-period on a 3.40 GHz Intel Core i7 processor (12 Go of RAM) takes between 2 min 5 s to 2 min 45 s.

Figure 4.1. Drainage maps of the three River basins with the studied watersheds (gray area): Ceizur, Cowansville and Toulnustouc



4.3 Preliminary results

Preliminary studies were done to identify the effective optimization strategies of each benchmark optimization algorithm (Huot et al. 2014). Each benchmark method is presented below, followed by a discussion of the results on computational tests.

4.3.1 Benchmark optimization algorithms

Three benchmark algorithms are considered in this study:

- **SCE-UA**. This algorithm is an evolutionary algorithm that generates a first population of model parameter sets (solutions) and makes them evolve in the solution space through "Competitive Complex Evolution" (CCE). This is a heuristic method that performs some reflections of the lower-quality solutions according to a reflection axis drawn through a centroid formed by higher-quality solutions, hence leading the optimization procedure to good-quality zones. SCE-UA works by grouping solutions into complexes that are mixed on the subsequent algorithm iteration to explore different zones of the solution space. The extensive use of SCE-UA for hydrological model calibration in the last three decades makes it a "state-of-the-art" algorithm for comparison with other optimization algorithms (Yapo et al. 1996, Vrugt et al. 2003, Mugunthan et al. 2005, Arsenault et al. 2014, Huang et al. 2014). Provided by the authors, the release that was used in this study includes updates as of September 2004.

- **DDS**. The second algorithm is the "Dynamically Dimensioned Search" (DDS) approach introduced by Tolson and Shoemaker (2007). The main feature of DDS is its ability to automatically scale the search on a user-specified simulation budget. To achieve this, the value of the user-specified budget is integrated into the calculation of the probability of perturbing each search dimension. Therefore, at the beginning of the optimization, DDS perturbs a large number of model parameters (search dimensions), and as the optimization progresses it gradually reduces the number of dimensions to be searched, to finally perturb only one dimension at a time. Thus, the search in two

calibration procedures with near but different simulation budgets will proceed differently because of this scaling. This single-solution based algorithm always searches from the current best solution; all of the worse solutions are rejected and never influence the optimization procedure. Previous studies (Razavi et al. 2010, Arsenault et al. 2014) have shown that the global search strategies used by DDS perform efficiently for the calibration of hydrological models, quickly targeting good-quality solutions. Provided by the authors, release DDS 1.2, which includes updates as of February 2015 was used here.

- **MADS**. The last algorithm is the "Mesh Adaptive Direct Search" (MADS; Audet and Dennis 2006, Abramson et al. 2009). MADS is a direct search method that uses a discretization of the solution space, called the mesh. It performs an adaptive search on the mesh, including controlling refinement and expansion. Each iteration of the algorithm generates $2N$ trial points on the mesh according to orthogonal directions from the current best solution (N being the number of calibration parameters). When iterations fail to improve the solution (after testing the $2N$ trial points), the next iteration is initiated on a finer mesh. On successful iterations, the mesh is coarsened, which may occur as soon as a trial point appears to be better than the current best solution. Unlike the previous two methods, MADS is able to handle general constraints (by the so-called progressive barrier approach (Audet and Dennis 2009)) and is supported by a rigorous convergence analysis, guaranteeing that the final solution will satisfy some necessary optimality conditions (Audet and Dennis 2006). These nonsmooth optimality conditions guarantee the existence of an accumulation point x of the sequence of trials points produced by the MADS algorithm, at which the Clarke

generalized directional derivative of the Lipschitz continuous objective function f is nonnegative in all hypertangent directions d to the domain Ω:

$$f^{\circ}(x;d) := \limsup_{t\searrow 0, y\to x} \frac{f(y+td)-f(y)}{t} \geq 0 \forall d \in T_{\Omega}^{H}(x). \tag{4.2}$$

This is a generalization of the standard smooth optimality condition ensuring that the directional derivatives are nonnegative in the tangent directions, i.e.:

$$f'(x;d) := \lim_{t\searrow 0} \frac{f(x+td)-f(x)}{t} \geq 0 \forall d \in T_{\Omega}(x). \tag{4.3}$$

The interested reader may consult the recent book of Audet and Hare (2017) for the detailed proof, and for contextualization in derivative-free optimization. The NOMAD 3.5 implementation of MADS is used for the present numerical experiments (Le Digabel 2011).

All algorithmic settings are set to the default values as recommended by the authors of each method (Duan et al. 1994, Tolson and Shoemaker 2007, Le Digabel 2011) in the interest of fair comparison.

4.3.2 Computational time

When the spatial discretization of HYDROTEL is combined with a long simulation time period and a high number of meteorological stations, the computational time for a single simulation inevitably becomes intensive. In daily operational conditions, calibrating such models is simply unpractical. Table 4.2 presents the computational times required for calibrating each combination of the “Watershed-Optimization Problem” with the algorithms detailed in Section 3.1. The experiments were carried out on four uniform computers having each 3.40 GHz Intel Core i7 processor with 12 Go of RAM. More than 16,900 hours of computational time have been dedicated to perform all the calibrations required to complete

research investigations and to generate the results presented in this study. Detailed information on these numerical experiments (including figures that show the best objective function value in terms of calls to the simulation) can be found in Huot (2014).

Based on Poulin et al. (2011) and previous works from the authors (Huot et al. 2014), in which the Ceizur and the Cowansville watersheds were studied, the calibration procedure is limited to 2,000 model simulations. This limit is reached here by all three optimization algorithms, whatever the problem. Huot et al. (2014) also provide the standard $1 - NSE$ values (best known local optima) and computational times for both watersheds, here reported in Table 4.2. The three algorithms provide equally accurate calibrations after 2,000 model simulations. Although they work differently, the computational times are identical for all since it is the number of simulations that is decisive (no simulation is done in parallel even if the algorithms may afford it). Since the Toulnustouc watershed is also studied in this paper, the same calibration experiments were performed on it and the results in terms of computational times and standard $1 - NSE$ values are also indicated in Table 4.2 (shaded line).

Table 4.2. Summary of computational times and $1 - NSE$ results from automatic optimization for each "Watershed-Optimization

				HYDROTEL - 10 Parameters		HYDROTEL - 19 Parameters	
Use in this study	**Watershed**	**Area (km²)**	**Simulated Period**	**Computational Times [a]**	**Standard 1-NSE [b]**	**Computational Time [a]**	**Standard 1-NSE [b]**
Implementation	Cowansville	215	5 years	37 hours	0.38	44 hours	0.30
	Ceizur	6,928	4 years	54 hours	0.15	62 hours	0.13
Validation	Toulnustouc	8,109	4 years	63 hours	0.44	77 hours	0.27

[a] Computational time based on a 2,000 evaluations calibration budget

[b] Standard 1 minus Nash-Sutcliffe Efficiency (1-NSE) result from an automatic calibration

4.3.3 Performance comparison

Based on the works from Huot et al. (2014), a series of computational experiments were conducted on the Cowansville and Ceizur watersheds to compare the performance of DDS, MADS and SCE-UA. Each method is used to perform 32 calibration trials on each watershed

for a better statistical representativeness of the results and to better study the behaviors of the optimization algorithms. Each calibration trial is performed with a budget of 2,000 model simulations. The trials are initialized from 32 random model parameter sets, one for each trial, always the same for a trial whatever the algorithm and the watershed used.

Figure 4.2 shows boxplots of the $1 - NSE$ values that summarize the results obtained for the Ceizur watershed for both optimization problems (HYDROTEL 10 and HYDROTEL 19). The results are shown at three checkpoints during the optimization procedure, i.e., after 250, 500 and 1,000 model evaluations, and at the end of the optimization (i.e. when the number of model evaluations reaches 2,000). Similar results were obtained for the Cowansville watershed (not shown).

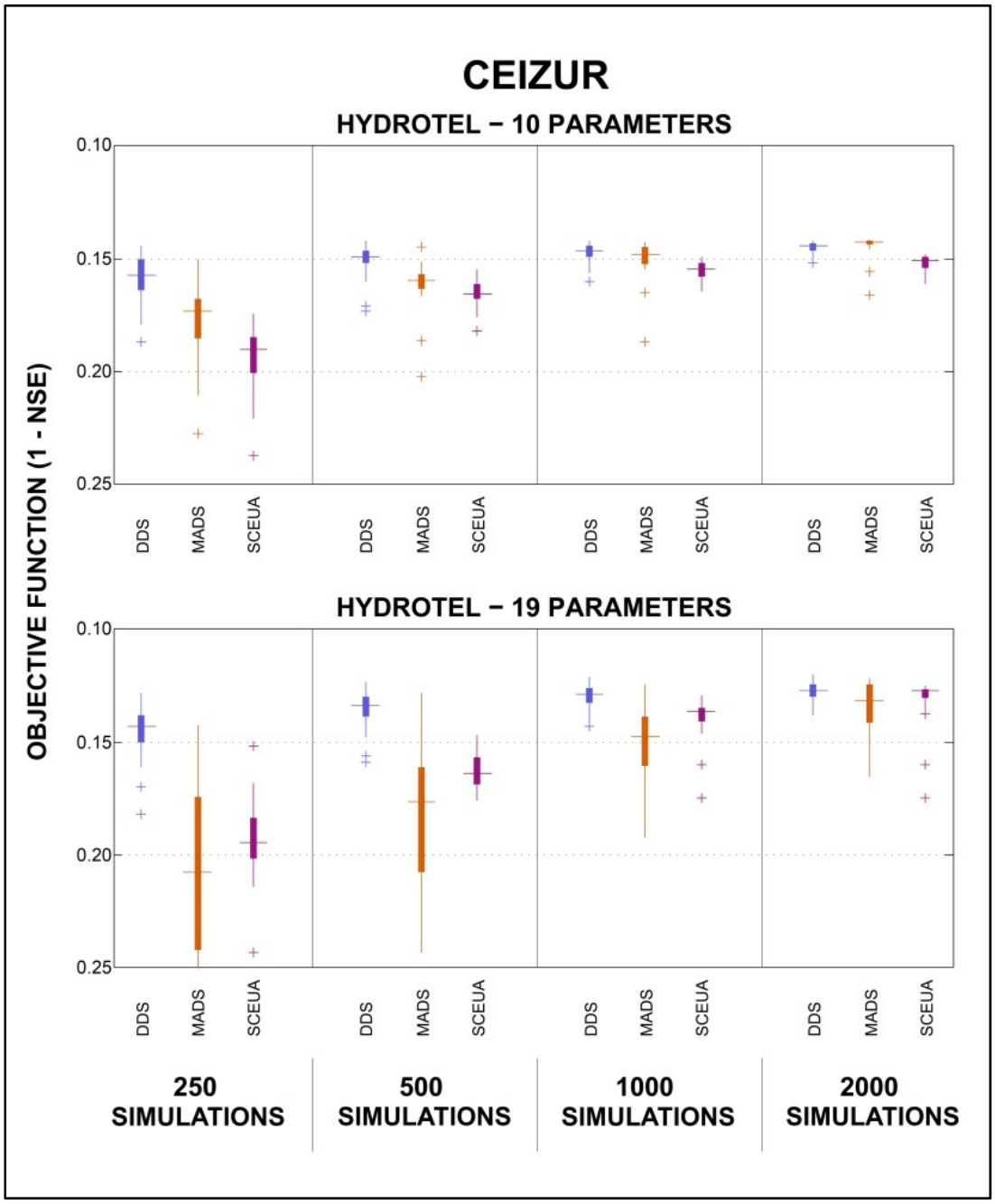


Figure 4.2. Final $1 - NSE$ results from optimizations by DDS, MADS and SCE-UA for the calibrations of HYDROTEL 10 and HYDROTEL 19 on the Ceizur watershed at 250, 500 and 1,000 simulations, and at the end of the optimization

Firstly, Figure 4.2 shows that the DDS algorithm stands out over the course of the calibration, and demonstrates a high potential for reducing the computational time since it finds good-quality model parameter sets very quickly, in only a low number of simulations.

After only 250 simulations, the 32 DDS optimization trials are already in a good-quality zone of the solution space for both HYDROTEL versions. For the remaining 1,750 simulations, the DDS method reduces the distance between the upper and lower quartiles and it slightly improves the median value. The MADS and SCE-UA algorithms are much slower. They need 500 to 1,000 simulations to get solutions that are equivalent to those generated by the DDS in some 250 simulations. However, from a hydrological perspective, all three methods converge on solutions of similar quality at the end. The medians are within a 0.02 interval and the interquartile ranges are of the same order of magnitude for both HYDROTEL problems. Although the DDS definitely shows a better efficiency during the optimization sequence, MADS and SCE-UA still provide high-quality final results after 2,000 simulations.

The results presented in Figure 4.2 highlight some important facts regarding the optimization strategies used by each method. The DDS algorithm offers significant potential for reducing the number of model simulations, as it identifies good-quality solutions very quickly in comparison with MADS and SCE-UA. Its ability to properly explore the solution space reveals interesting optimization strategies that differentiate it from the others. SCE-UA works with a population of model parameter sets that evolves in the solution space and consequently, it wastes a significant number of model simulations because it needs to transport the whole population to good-quality zones, not just a single point as with DDS. In some ways, MADS also wastes model simulations when converging, since it must proceed through $2N$ simulations before decreasing the mesh size of the grid. However, this inefficiency remains limited since MADS terminates rapidly when it reaches a locally optimal solution. The final solutions shown in Figure 4.2 appear to be of good-quality. For the HYDROTEL 10 problem, 65% of the 32 MADS optimization trials were completed before reaching the 2,000-simulations budget, and 15% in the case of HYDROTEL 19.

Results suggest that the DDS algorithm could provide better or equivalent final values of the objective function with lower simulation budgets. Tolson and Shoemaker (2008) mention that a budget of $3N$ to $7N$, where N is the number of calibration parameters, is enough to yield satisfactory results. Nonetheless, the DDS algorithm appears to be inefficient on the

HYDROTEL calibration with such model simulation budgets (corresponding to a maximum of 70 and 133 simulations for HYDROTEL 10 and 19, respectively). Experiments with these recommendations yielded wide interquartile ranges and some outliers for some of the calibration trials (algorithm is trapped in poor-quality zones). According to the authors' experience, a budget of 1,000 to 2,000 simulations is more appropriate for the calibration of HYDROTEL 10 and HYDROTEL 19 problems, respectively, to avoid the DDS algorithm getting trapped in poor-quality zones and produce good-quality solutions.

4.4 Proposed hybrid calibration approach

Two important factors need to be considered when choosing or developing optimization methods. First, the method must be able to perform a global exploration of the space of variables to identify a promising region containing good-quality solutions (exploration). Second, the method needs to perform a local descent within the identified region to approach the local optimum (exploitation) (Audet and Hare 2017). Among the three optimization methods tested (Figure 4.2), DDS outperformed the other two in the global exploration phase, while MADS constantly decreased the objective function value even for the local descent These two complementary features are then merged, and the following hybrid approach is proposed.

4.4.1 A hybrid DDS-MADS calibration approach

The hybrid DDS-MADS calibration approach aims to exploit the most efficient optimization strategies of both methods for reducing the overall computational time. A two-step approach is considered (a DDS step followed by a MADS step), with the effective strategies consisting of: (1) the global exploration ability of the DDS algorithm; (2) the local refinement procedure of the MADS algorithm; and (3) the automatic stopping criterion of MADS, which is relative to the convergence procedure.

In order to adequately merge DDS with MADS, five simple transition features are attached to the source codes of both optimization algorithms. Despite the broad availability of source

codes, no changes were made in an effort to maintain a high level of simplicity and to respect the original optimization strategies of both algorithms. For more details on the optimization strategies used, see Tolson and Shoemaker (2007) for the DDS algorithm and Le Digabel (2011) for the MADS algorithm. These transition features are described below. Figure 4.3 presents the pseudocode of the hybrid DDS-MADS calibration, highlighting in blue the five added transition features.

(1) Cache

The first feature is the addition of a cache file $\{C\}$ over the link between the blackbox simulation and the feedback sent to the optimization algorithm. After each call to the hydrological model, the objective function value $F(X_{new})$ and its associated model parameter sets X_{new} are registered in a cache file. This cache contains the historical record of the optimization procedure. Its implementation serves two purposes: (1) to avoid launching twice any simulation with the same model parameter sets; and (2) to provide information on the sensitivity of the objective function to each calibration parameter. This cache file feature is identified with the (a) symbol in Figure 4.3.

(2) Fine-tuning of the initial mesh size parameter

The MADS algorithm requires an initial value of the mesh size parameter Δ_{k0}. Instead of using the default value, the hybrid method analyses the results in $\{C\}$ from the DDS step to propose a more appropriate value of Δ_{k0}. This algorithm parameter dictates the initial step lengths performed by MADS. If a good-quality solution $X^{DDS}{}_{best}$ is provided to the MADS step, the initial mesh size Δ_{k0} must be tuned on a local domain. Alternatively, if a poor-quality solution is provided to the MADS step, the initial mesh size Δ_{k0} must be tuned on a larger domain to allow the method to escape from it. In both situations, the Δ_{k0} fine-tuning is achieved using a spatial variability analysis from the model parameter sets found in the cache file $\{C\}$. The initial mesh size Δ_{k0} is determined according to the two following steps, which are identified with the (b) symbol in Figure 4.3.

The first step deals with selecting the model parameter sets from the cache file $\{C\}$. All model parameter sets X having an objective function value $F(X)$ within a prescribed tolerance gap from $F({X^{DDS}}_{best})$ are selected to be part of the reduced cache $\{C^*\}$ serving the fine-tuning procedure. This gap is set to 10% of the difference between the worst X in $\{C\}$ and the value of ${X^{DDS}}_{best}$. This gap leads to a local domain selection when ${X^{DDS}}_{best}$ is already in a good-quality zone, and to a larger domain selection when the quality of ${X^{DDS}}_{best}$ is poorer. This strategy allows to include more than one good-quality zone if $\{C^*\}$ is composed of lower-level good-quality solutions positioned inside other good-quality zones far from the value of ${X^{DDS}}_{best}$.

The second step concerns the spatial variability analysis of X in $\{C^*\}$. Since the range of values may vary greatly from one calibration parameter to another, the initial mesh size Δ_{k0} must be fine-tuned anisotropically, i.e., differently for each model parameter. The maximum and minimum values of each calibration parameter j are used for setting the initial parameter mesh size ${\Delta^j}_{k0}$. Three different situations are possible. (1) The value of j is the same for all X in $\{C^*\}$. The mesh size is then set to its default value, i.e., 10% of the interval defined by the upper and lower bounds on the allowed values of j. (2) The values of j mostly cover the interval defined by the bounds. The mesh size is then set to 50% of the bound interval to maintain an efficient size of the mesh. (3) In all the other cases, the initial mesh size is set to the ratio between the range of the values taken by j and the interval length defined by the bounds on j.

(3) Stagnation parameter

With a two-step calibration approach, identifying the appropriate timing for interrupting the DDS step and initiating the MADS step is a major issue to ensure efficient calibration. When DDS reaches a zone of good-quality solutions, its optimization strategies become less efficient for local improvement. Stagnation occurs and often results in several hundreds of simulations without any improvement of the quality of the objective function value. Such stagnation suggests the exhaustion of the DDS step optimization strategies and indicates the starting point for the MADS step. To ensure the transition between the two calibration steps,

a stopping criterion represented by the stagnation parameter $stag$ is inserted to limit the number of consecutive unsuccessful simulations. The symbol (c) in Figure 4.3 refers to the stagnation parameter feature.

The stagnation parameter $stag$ must be set to a threshold value that allows a smooth transition to the MADS step at the right timing. The proposed value is $2N$, where N is the number of calibration parameters. There are two justifications for this choice: (1) the local search strategy of the MADS algorithm proceeds to the refinement of the mesh size after $2N$ unsuccessful model simulations; and (2) a sensitivity analysis was performed on HYDROTEL calibration problems, and showed that $2N$ is an appropriate threshold for the stagnation parameter $stag$. As a result, the DDS step benefits from the same opportunity to improve the quality of the objective function value as the MADS step does for the refinement of the mesh size.

After each completed simulation, the value $F(X_{new})$ is compared to the best current value F_{best} to update the number of consecutive unsuccessful simulations s. Three scenarios may occur: (1) if $F(X_{new})$ improves the objective function value, the stagnation counter is reset to $s = 1$ and $F_{best} = F(X_{new})$; (2) if $F(X_{new})$ equals F_{best}, the stagnation counter remains unchanged; and (3) if F_{best} remains superior to $F(X_{new})$, then the simulation is declared unsuccessful and the stagnation counter increases to $s = s + 1$. When the stagnation counter s reaches the threshold of the stagnation parameter $stag$, the DDS step is stopped and the MADS one takes over.

(4) Total simulation budget

The user must provide the total simulation budget b_t, which determines the overall allowed computational time for the optimization procedure. The two-step approach is designed to allow the optimization strategies of both methods using the necessary simulation budget in terms of their own capacity to decrease the objective function value. The DDS step is initially set with the total simulation budget b_t to globally explore the solution space until the stagnation threshold $stag$ is reached. The MADS step then continues the optimization

procedure with the remaining budget $b2 = b_t - b1$, where $b1$ represents the number of simulations completed in the DDS step (see the symbol (d) in Figure 4.3). It is important to keep in mind that the remaining budget $b2$ is not necessarily used completely if the MADS step terminates the optimization according to its own stopping criterion.

(DDS step)
DDS algorithm

STEP 1 - DEFINITION OF DDS INPUTS

- Total simulations budget: b_t
- Neighborhood perturbation size parameter: $r = 0{,}2$
- Vector of lower bounds: X_{min}
- Vector of upper bounds: X_{max}
- Initial solution: X_0
- (c) **Stagnation parameter:** $stag = 2 \cdot length(X_0)$

STEP 2 - INITIAL SIMULATION

- Set simulation counter to: $b1 = 1$
- (c) **Set stagnation counter to:** $s = 1$
- Evaluate objective function at initial solution: $F_{best} = F(X_0)$ and $X_{best} = X_0$
- (a) **Store $F(X_0)$ and X_0 in cache $\{C\}$**

STEP 3 - SELECTION

- Calculate probability of the current iteration:
  $P(b1) = 1 - (Ln(b1) / Ln(b_t))$
- Selection of disturbed decision variables:
  $FOR\ j = 1, \ldots, length(X_{best})$, add X^j_{best} to $\{Y\}$ with probability $P(b1)$
  IF $\{Y\}$ is empty, select one random X^j_{best} and add to $\{Y\}$

STEP 4 - PERTURBATION

- Perturbation of all decision variables:
  FOR each decision variable of $\{Y\}$
  $X^j_{new} = X^j_{best} + \sigma^j \cdot N(0,1)$, where $\sigma^j = r \cdot (X^j_{max} - X^j_{min})$
- Check for exceeding lower or upper bound:

$$X^j_{new} = \begin{cases} X^j_{min} + (X^j_{min} - X^j_{new}) & X^j_{new} < X^j_{min} \\ X^j_{max} - (X^j_{new} - X^j_{max}) & X^j_{new} > X^j_{max} \end{cases}$$

STEP 5 - NEW SOLUTION SIMULATION

- Evaluate the new solution: $F(X_{new})$
- (a) **Store $F(X_{new})$ and X_{new} in cache $\{C\}$**
- Update the current best solution if necessary:
  (c) $F_{best} = F(X_{new})$ AND $X_{best} = X_{new}$ $|$ $IF\ F(X_{new}) \le F_{best}$

$$s = \begin{cases} 1 & IF\ F(X_{new}) < F_{best}, \\ s & ELSE\ IF\ F(X_{new}) = F_{best}, \\ s+1 & otherwise; \end{cases}$$

STEP 6 - STOPPING CRITERION

- Update the simulation counter: $b1 = b1 + 1$
- Check stopping criteria:
  (c) $IF\ b1 = b_t$ or $s = stag$, $STOP$ the optimization process
  $ELSE$ go to $Step\ 3$

$START$ optimization with $MADS$ from X^{DDS}_{best}

(MADS step)
MADS algorithm

▶ **STEP 0 - DEFINITION OF MADS INPUTS**

- (d) **Number of function simulations (remaining budget):**
  $b2 = b_t - b1$
- Vector of lower bounds: X_{min}
- Vector of upper bounds: X_{max}
- Initial solution: $X_0 = X^{DDS}_{best}$
- Initial mesh size parameters:
  (a) **All lower-quality solutions in $\{C\}$ are raised $\{C^*\}$**
  (b) $FOR\ j = 1, \ldots, length(X_0)$, $max^j = max(\{C^{j*}\})$ and $min^j = min(\{C^{j*}\})$

$$\Delta^j_{k0} = \begin{cases} 0.1 \cdot (X^j_{max} - X^j_{min}) & IF\ max^j = min^j, \\ 0.5 \cdot (X^j_{max} - X^j_{min}) & ELSE\ IF\ \dfrac{(max^j - min^j)}{(X^j_{max} - X^j_{min})} \ge 0{,}5, \\ (max^j - min^j) / (X^j_{max} - X^j_{min}) & otherwise; \end{cases}$$

- Mesh update basis parameter: $\tau = 4$
- Evaluate objective function at initial solution:
  $F_{best} = F(X_0)$ and $p_k = X_0$

STEP 1 - SEARCH (skip this step on first iteration)

- Random generation of new solutions around the best solution found:
  $\{R\} = p_k + \Delta_k \cdot Z'$
  where Z' is random orthogonal directions unvisited during the poll step
- Evaluate objective function at all solutions:
  $FOR\ EACH\ R^t \in \{R\}$, $F(R^t)$
- Update the current best solution if necessary:
  $IF\ F(R^t) \le F(p_k)$, $F(p_k) = F(R^t)$ and $p_k = R^t$, go to $Step\ 3$
  $Step\ 1$ is declared successful

STEP 2 - POLL

- Generation of new solutions on the mesh around the best solution found:
  $\{L\} = p_k + \Delta_k \cdot Z''$
  where Z'' is $2 \cdot N$ orthogonal directions and $N = length(X_0)$
- Evaluate objective function at all solutions:
  $FOR\ f = 1, \ldots, 2 \cdot N$, $F(L^f)$
- Update the current best solution if necessary:
  $IF\ F(L^f) \le F(p_k)$, $F(p_k) = F(L^f)$ and $p_k = L^f$, go to $Step\ 3$
  $Step\ 2$ is declared successful

STEP 3 - UPDATING PARAMETERS

- Update the mesh size:
  $IF\ Step\ 1$ or $Step\ 2$ is declared successful, $\Delta_k + 1 = \Delta_k \cdot \tau$
  $ELSE\ \Delta_{k+1} = \Delta_k / \tau$
- Check stopping criteria:
  IF simulation counter $= b2$ or $\Delta_k \le \Delta_k^{min}$, $STOP$ the optimization process
  $ELSE$ go to $Step\ 1$

Figure 4.3. Pseudocode of the hybrid DDS-MADS calibration approach

(5) Rounding the objective function value

The last feature addresses the attenuation of the numerical noise on the objective function surface by rounding the values. A very desirable property of the MADS algorithm is its capacity to adapt the mesh size to nonsmooth objective function surfaces, regardless of the noise level. While the objective function surfaces of the hydrological models are likely characterized by noise, from a hydrological perspective on a long simulation period, a difference of 0.005 between two values of $1 - NSE$ does not allow an absolute differentiation of the quality of the model parameter sets in a validation or operational simulation. The DDS and MADS algorithms both deal with objective function values having a precision level of up to 8 digits. However, a continuous adaptation from refinement to expansion of the mesh size for slight gains less than a 0.005 interval of $1 - NSE$ would lead to unnecessary computational efforts. To avoid this situation, the noise level is attenuated with appropriate rounding. Figure 4.4 presents a sensitivity analysis on the number of digits (2, 3, 4 and 5 digits) that was performed on the Ceizur watershed for the HYDROTEL 10 problem with four different trials for each number of digits. This figure shows that rounding the objective function values conservatively to 3 digits (± 0.001) ensures a faster convergence to a satisfactory solution without stretching the continuous adaptive size of the mesh for useless objective function gain (referred in Figure 4.4 to trials reaching 1,300 simulations). This rounding feature impacts the number of simulations completed by the hybrid approach (an economy of approximately 300 simulations in comparison with 4 or 5-digit results). The same conservative logic may be applied when other objective functions are used, but depends on the desired precision level, and on the characteristics of the objective function surface. By handling this directly through the $F(X)$ values returned by the hydrological model simulation, both the DDS and MADS steps can take advantage of this rounding.

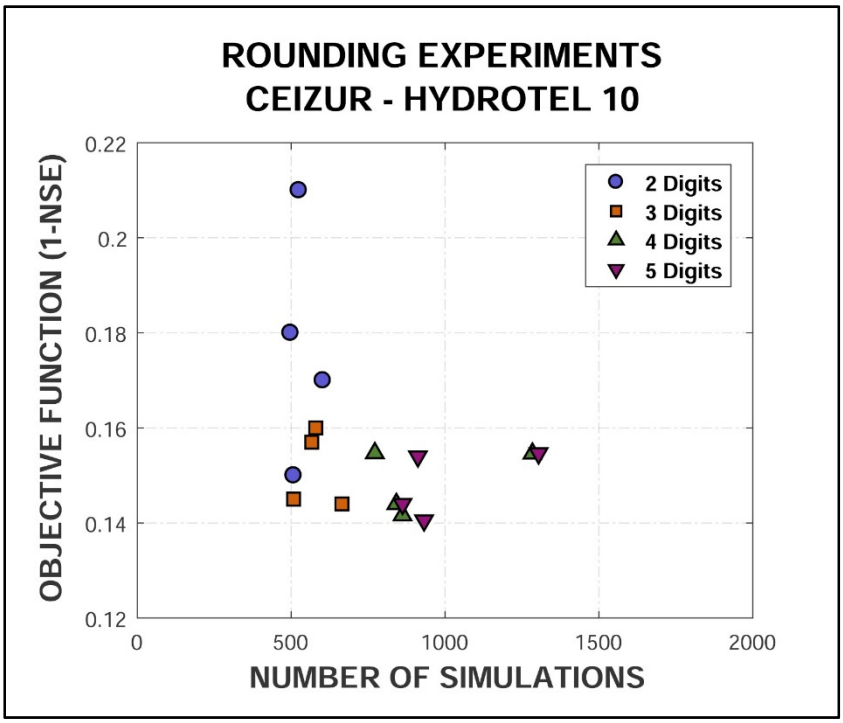


Figure 4.4. Experiments on rounding the objective function value (number of digits) for the Ceizur watershed on HYDROTEL 10 problem

4.4.2 Numerical results

Figures 4.5 and 4.6 present the final results from 32 calibration trials using the hybrid DDS-MADS approach for the Cowansville and Ceizur watersheds modelled by HYDROTEL 10 and 19, in comparison with the benchmark DDS algorithm set to a total budget of 2,000 simulations (hereafter called DDS-Bench). DDS-Bench is used as a basis for comparison since it provided the best performance in preliminary studies (Section 3). The hybrid method and DDS-Bench both start with the same 32 random model parameter sets. In Figures 4.5 and 4.6, the grey curves in the background represent results obtained with DDS-Bench. The blue curves represent the DDS optimization step within the hybrid DDS-MADS approach, while orange curves represent the MADS optimization step. The blue dots are the endpoints of the DDS step and the starting points of the MADS step. The orange points correspond to the final objective function values obtained with the hybrid approach at the end of the optimization.

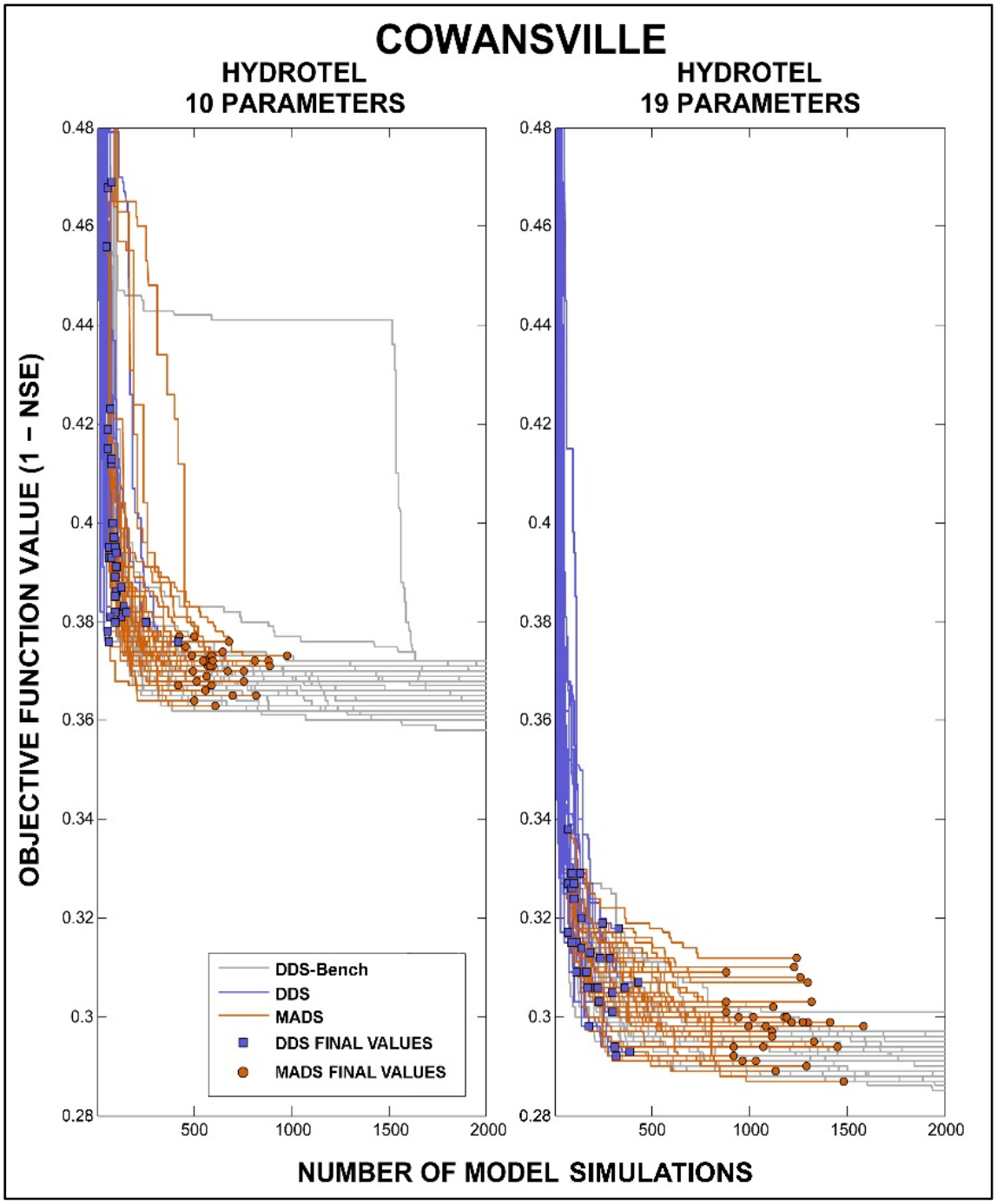


Figure 4.5. Evolutions of the best objective function values obtained ($1 - NSE$) with the hybrid DDS-MADS approach compared to DDS-Bench values. Optimization was performed for the HYDROTEL 10 and 19 problems on the Cowansville watershed

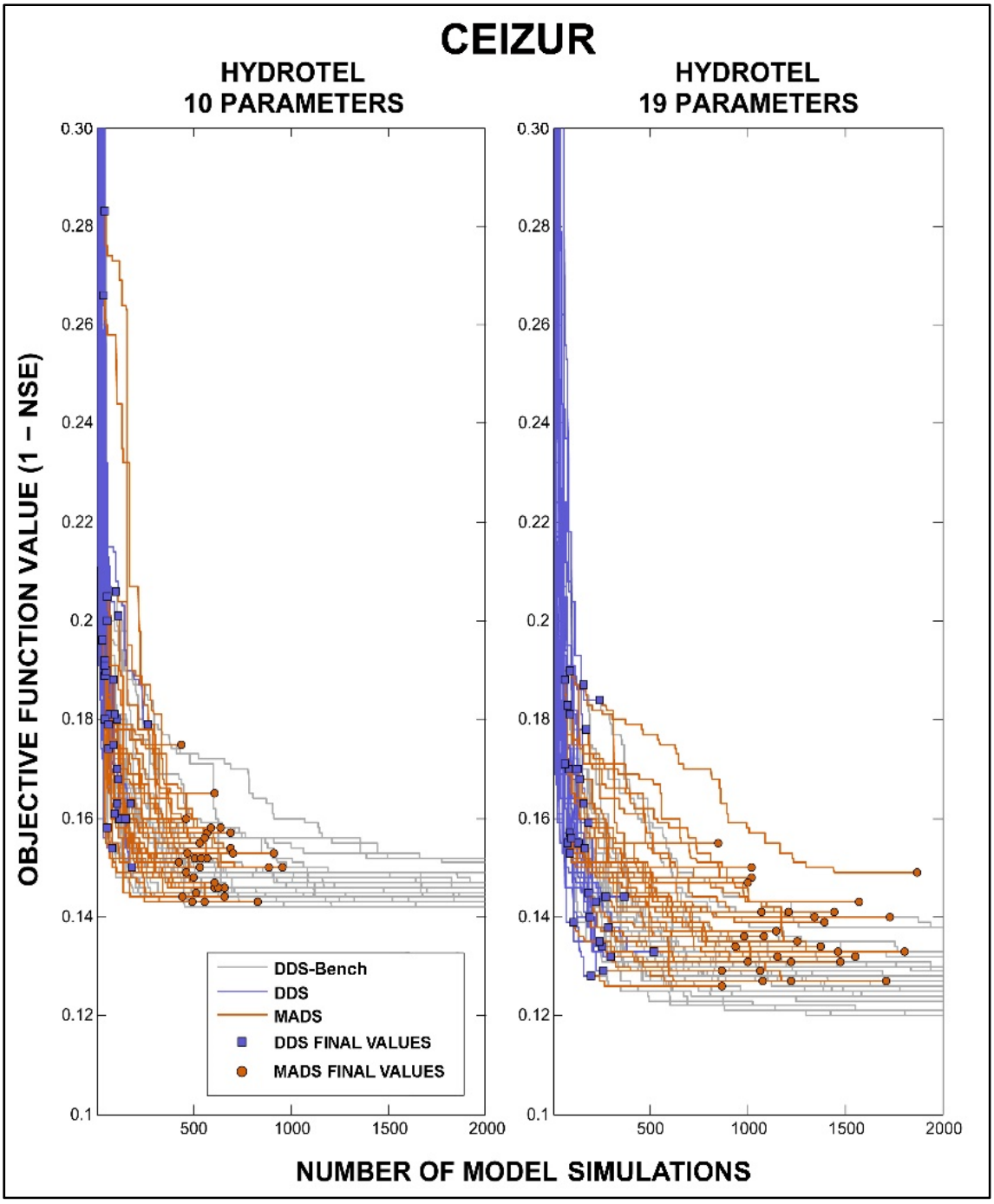


Figure 4.6. Evolutions of the best objective function values obtained ($1 - NSE$) with the hybrid DDS-MADS approach compared to DDS-Bench values. Optimization was performed for the HYDROTEL 10 and 19 problems on the Ceizur watershed

Figures 4.5 and 4.6 show that the evolutions of the best objective function values of the hybrid approach are comparable to those obtained with DDS-Bench. However, each calibration trial of DDS-MADS stopped before reaching the 2,000-simulation budget, providing a reduction in computational time. One can observe that the DDS step tends to greatly reduce the value of the objective function within a limited number of simulations. As discussed in Section 3.3, the DDS step demonstrates remarkable efficiency in global exploration of the solution space, quickly locating good-quality zones. Then, when the DDS optimization strategies are dwindling, the MADS step takes over and continues to decrease the objective function value. The MADS step allows the calibration procedure to finish quickly, while ensuring that a local optimum can be reached on a limited area of the solution space.

To ensure that no model simulation is wasted in the transition from DDS to MADS, the initialization of the MADS step must adequately control the adaptive search on the mesh (expansion and refinement procedures). Figure 4.7 presents the number of model simulations required by the MADS step to make a first improvement on the objective function value. The horizontal axis shows the number of model simulations of the MADS algorithm corresponding to $2N$ trial points generated on the mesh (N being the number of calibration parameters). Figure 4.7 shows that MADS immediately improve the solution for both HYDROTEL problems. Remind that DDS was interrupted after $2N$ unsuccessful model simulations to start the MADS step. This shows that there is no slow-down in the decreasing curve when the MADS step takes over, an indication that the initial mesh size is a good fit based on the historical results of the previous DDS step. There are some calibrations that obtain no improvement with the MADS step (see right side of Figure 4.7). This shows that the DDS step has previously completed a high-quality global search in the calibration procedure, while the MADS step confirms the satisfaction of optimality conditions with the objective function value rounded to 3 digits. This contributes to decreasing the time still providing final solutions that satisfy necessary optimality conditions.

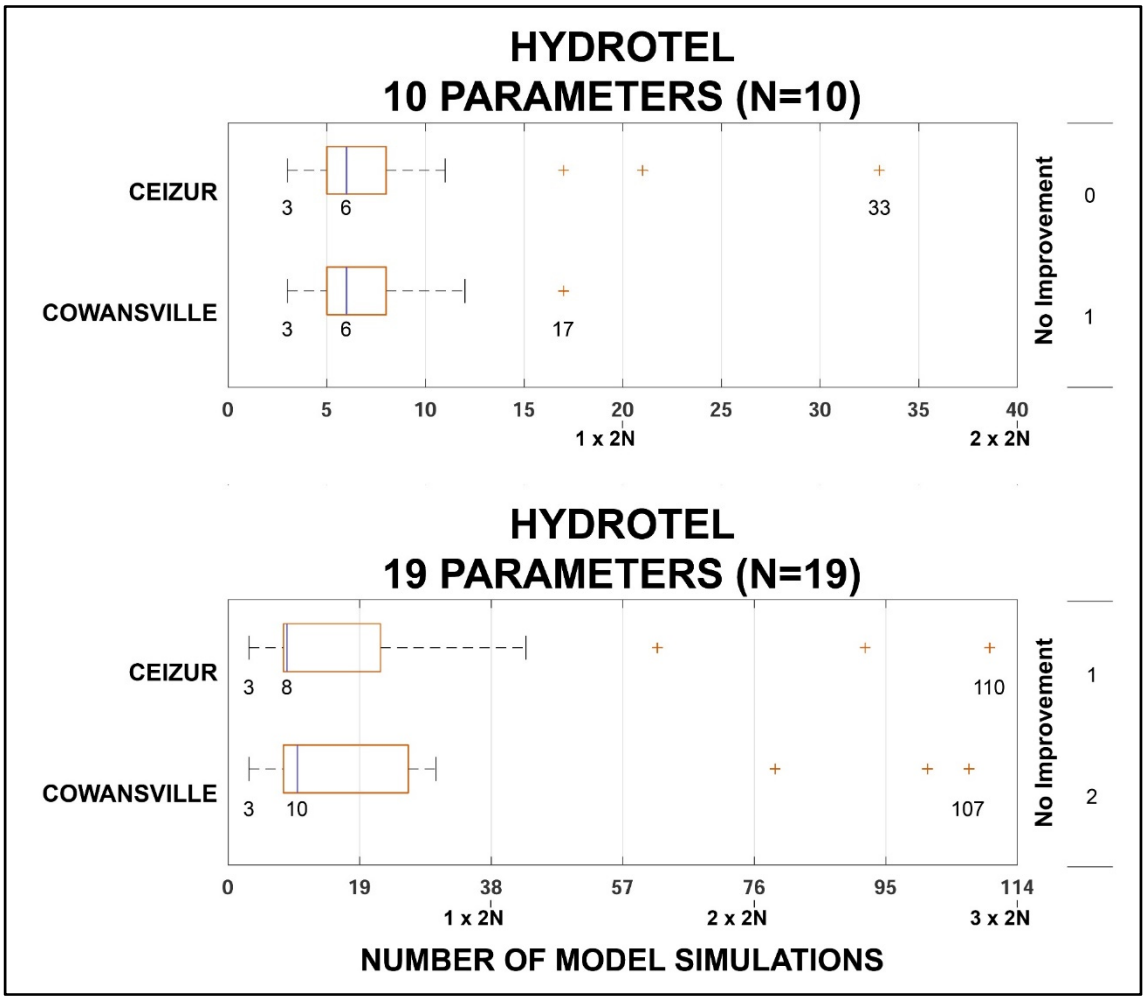


Figure 4.7. Number of model simulations required by the MADS step to improve the endpoints of the DDS step on HYDROTEL 10 and HYDROTEL 19 for the Ceizur and Cowansville watersheds

Figures 4.8 and 4.9 are about the distributions of the final objective function values obtained for the Cowansville and Ceizur watersheds, respectively, and for both HYDROTEL problems. These figures show that the range between the final values of DDS-Bench (grey box plots) and those of the hybrid DDS-MADS approach (orange box plots) is less than 0.02 of the $1 - NSE$. In a hydrological modelling context running on long simulation periods, this range is too small to favour one optimization method over the others, especially when running the model in validation or operational simulation modes (Arsenault et al 2018). The low variability of the 32 final solutions obtained with DDS-MADS provides an indication of the robustness of this calibration approach.

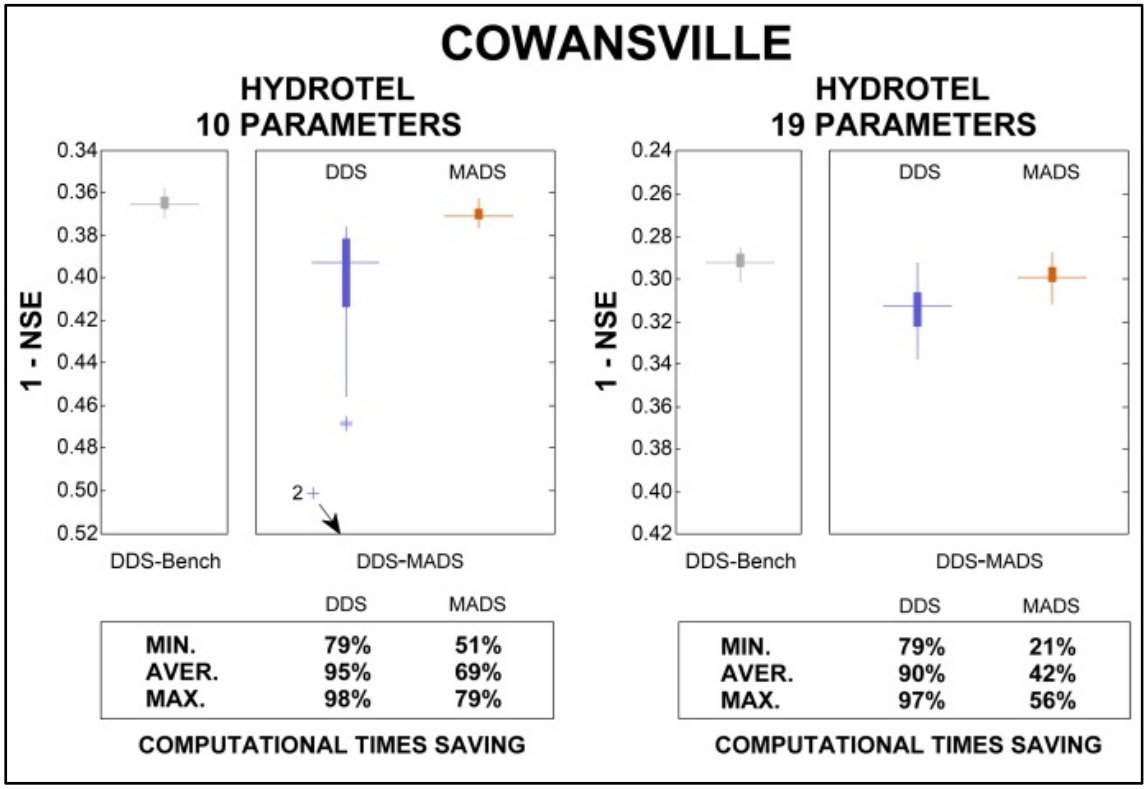


Figure 4.8. Final $1 - NSE$ optimization results and computational time savings of DDS-MADS (separated into DDS and MADS steps) compared to DDS-Bench results. Optimization was performed for the HYDROTEL 10 and 19 problems on the Cowansville watershed

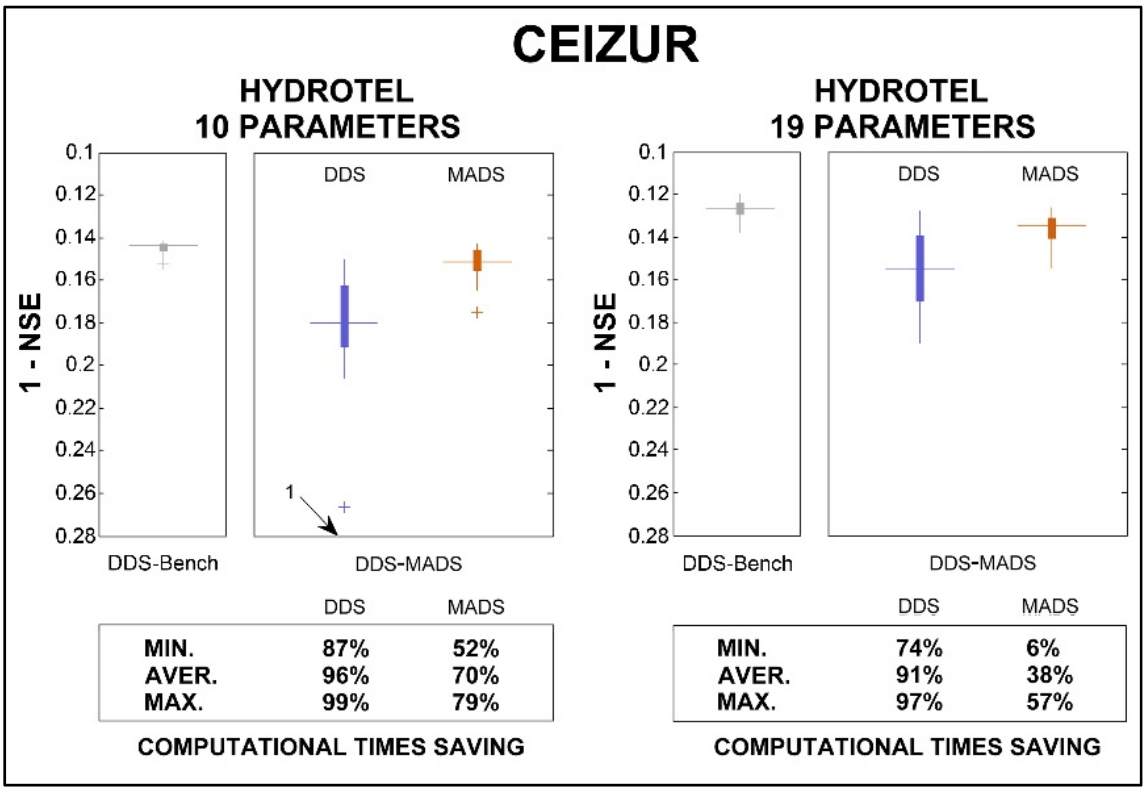


Figure 4.9. Final $1 - NSE$ optimization results and computational time savings of DDS-MADS (separated into DDS and MADS steps) compared to DDS-Bench results. Optimization was performed for the HYDROTEL 10 and 19 problems on the Ceizur watershed

The wider variability obtained by the DDS step (blue box plots) in comparison with that of the MADS step (orange box plots) is also noteworthy. This wide variability suggests that it is not a good strategy to terminate the calibration immediately after the DDS step when the stagnation parameter threshold has been reached. The MADS step has the capacity to reduce the variability of the objective function values, which is a valuable complement to the efficiency of the global exploration capacity of the DDS step.

The computational time savings indicated in Figures 4.8 and 4.9 are expressed in terms of percentage based on one minus the ratio of the number of simulations executed by the algorithm to the total budget of 2,000 simulations. Using the hybrid DDS-MADS approach offers a computational time reduction of 69% on average compared to the DDS-Bench results for the calibration of HYDROTEL 10, and of 42% on average for HYDROTEL 19.

This results in a reduction of 25 and 18 hours, respectively, on the computational times indicated in Table 4.2 for the two Cowansville HYDROTEL problems. Increasing the number of model parameters from $N = 10$ to $N = 19$ almost doubles the computational time. This is mostly due to the fact that each iteration of MADS launches $2N$ simulations at each iteration to eventually refine the mesh size and converge to a nearby local optimum. Note that the MADS algorithm allows changing the $2N$ number of orthogonal directions visited on the mesh to $N + 1$ directions (for example). This can reduce the number of necessary simulations to converge to a nearby local optimum, but slightly lower-quality final solutions may be obtained. Similar interesting results were obtained on the Ceizur watershed, i.e., 70% computational time reduction on average for the calibration of HYDROTEL 10 and 38% on average for HYDROTEL 19, representing a decrease of 37 and 23 hours respectively on the computational times in Table 4.2.

4.5 Validation

Calibrations were performed in Section 4 on the Cowansville and Ceizur watersheds. The Toulnustouc watershed is now used as validation for the hybrid DDS-MADS approach. This third watershed is completely independent from the development of the DDS-MADS approach. It also has physiographic characteristics that differ from the two others and is located in a different region of the province. This was motivated by the need of validating the versatility of the hybrid approach while also proving a non-overfit to some specific optimization problems. Both problems HYDROTEL 10 and 19 are solved for Toulnustouc. Figure 4.10 shows the results on the watershed as in Figures 4.5 and 4.6.

The performance of the hybrid approach is equally effective as it was on the other two watersheds. The quality of the final objective function values is comparable to those obtained with DDS-Bench, and the variability of final solutions is within the 0.01 and 0.02 ranges of the objective function for HYDROTEL 10 and 19 respectively. All calibration trials are stopped before achieving the 2,000-simulation budget. Based on Table 4.2, the computational time savings is slightly higher on these validation problems; 75%, corresponding to 45 hours

on average on HYDROTEL 10, and 45%, corresponding to 35 hours on average on HYDROTEL 19. The results on the Toulnustouc watershed show robustness in terms of the quality of the final solutions, and the computational time efficiency is preserved.

The HYDROTEL 19 results in Figure 4.10 show that the DDS step was trapped in poor-quality zones for two calibration trials with a value of $1 - NSE$ greater than 0.39. A similar behaviour is observed for three calibration trials in Figure 4.5 on HYDROTEL 10. Results show that DDS-MADS is able on these five trials to exit these zones of poor-quality due to its refinement and expansion procedures. Although this does not occur here with DDS-Bench, it is likely to happen. Starting multiple simultaneous calibrations reduces the probability of DDS being trapped in poor-quality zones. This can be an issue when only single calibrations are possible or allowed. The MADS step prevents this and maintains the optimization procedure even on single calibration.

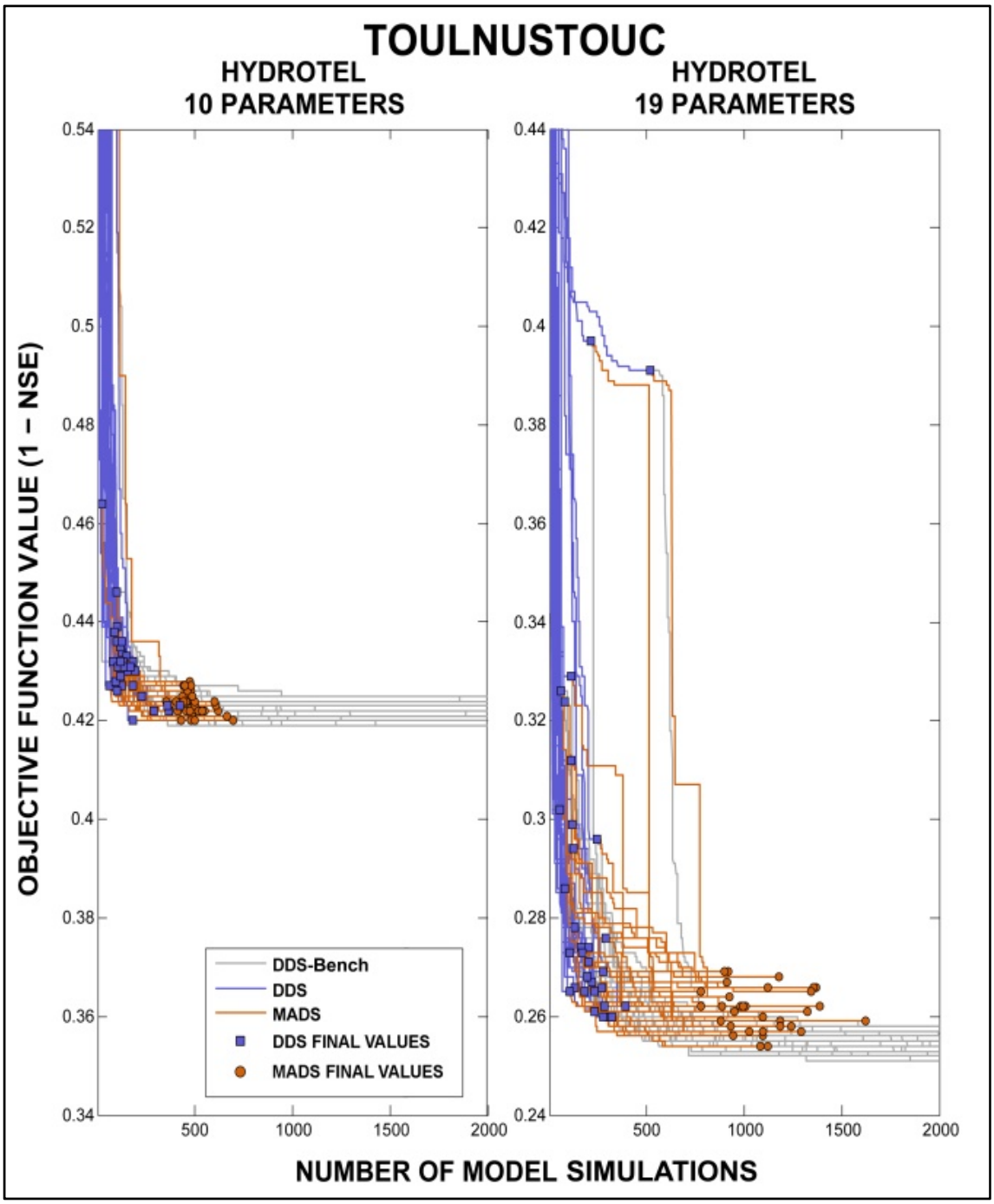


Figure 4.9. Evolutions of the best objective function values obtained ($1 - NSE$) with the hybrid DDS-MADS approach compared to DDS-Bench values. Optimization was performed for the HYDROTEL 10 and 19 problems on the Toulnustouc watershed

On all three watersheds, average computational time savings of 70% on HYDROTEL10 and 40% on HYDROTEL 19 were achieved in comparison with the DDS-Bench reaching 2,000 model simulations. This means a saving of 1,400 and 800 model simulations for HYDROTEL 10 and 19, respectively. In the interest of fair comparison, the next validation experiments compare the DDS-MADS approach with DDS-Bench and the DDS algorithm with simulation budgets equivalent to the total number of simulations executed by DDS-MADS. The DDS algorithm is then run with 600 and 1,200 total simulation budgets for HYDROTEL 10 and 19, respectively (named hereafter DDS-600 and DDS-1200). A "classical" hydrological model validation is then performed by applying the above model parameter sets on a different time period. Figure 4.11 presents a comparison of the final $1 - NSE$ values obtained in calibration versus validation with DDS-Bench, DDS-600 (only with HYDROTEL 10), DDS-1200 (only with HYDROTEL 19) and DDS-MADS for the Toulnustouc watershed applied to both HYDROTEL problems. The validation period extends from October 1st, 1988 to September 30, 1998 (10 years), and all 32 final model parameter sets obtained in calibration trials (4 years) are validated on this new time-period. Similarly to Figures 4.8 and 4.9, Figure 4.11 shows the short interquartile ranges of the final values of DDS-Bench, DDS-600 and those of the hybrid DDS-MADS approach obtained in calibration for HYDROTEL 10. Medians of the objective function values from all three approaches are also within a 0.01 range. Only DDS-600 presents four final $1 - NSE$ values that are considered as outliers and three of these are within 0.02 of the $1 - NSE$ median value. Similar observations can be made in the validation period for HYDROTEL 10. In the case of HYDROTEL 19, a slightly wider difference between the three medians is obtained in calibration and validation, but the difference remains smaller than 0.02 of $1 - NSE$. Despite similar results to those from DDS-MADS in calibration and validation, DDS-1200 provides as for DDS-600 some calibration trials with lowest final $1 - NSE$ values.

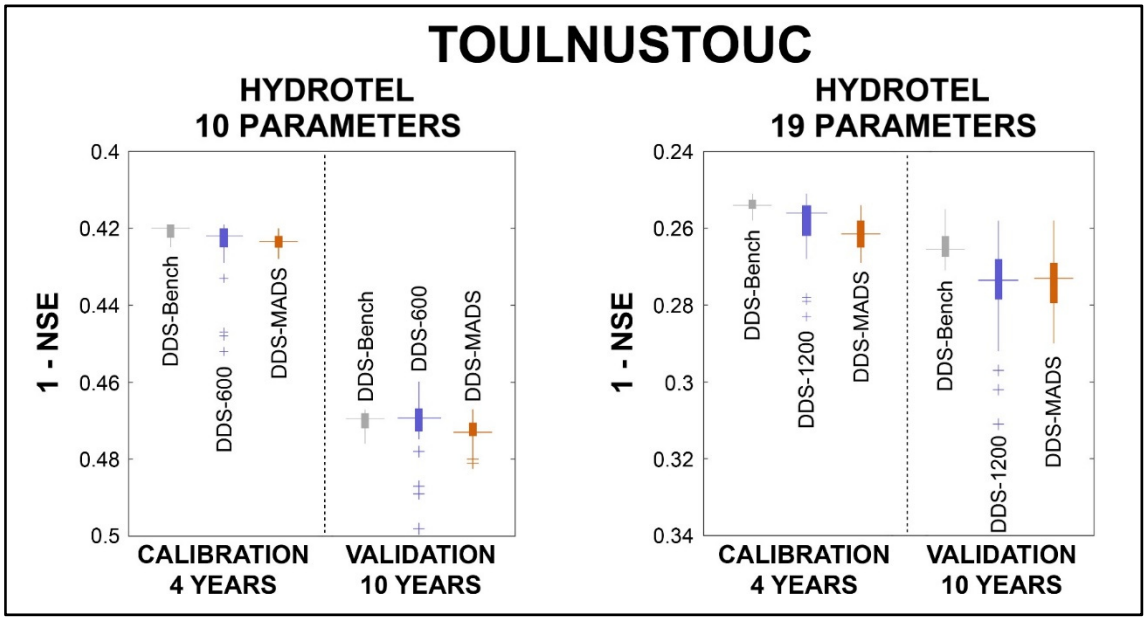


Figure 4.11. Final $1 - NSE$ optimization of DDS-MADS compared to DDS-Bench and DDS-600 results in the calibration and validation periods for the HYDROTEL 10 on the Toulnustouc watershed. Final $1 - NSE$ optimization of DDS-MADS compared to DDS-Bench and DDS-1200 results in calibration and validation periods for the HYDROTEL 19 on the Toulnustouc watershed

The overall results show that the hybrid DDS-MADS approach is comparable, in terms of final solutions obtained, to the DDS-Bench in calibration and validation. The slightly lower quality of the final solutions is compensated by a significant reduction in the number of necessary simulations; i.e. a significant reduction in computational time. Some calibration trials executed with DDS-600 and DDS-1200 are trapped in poor-quality zones. The DDS-MADS approach provides more robust and stable final $1 - NSE$ values with same overall simulation budget.

4.6 Discussion and Future Work

Figures 4.5 to 4.11 and preliminary results show that the DDS-Bench algorithm performs effectively at the beginning of the calibration, but consumes many unnecessary simulations after having found a good-quality solution. In several of these runs, more than 1,000 (unnecessary) simulations improve the quality of the final solution by only a marginal

amount. This behavior suggests beginning the calibration procedure using the DDS algorithm and then terminating it as soon as a good-quality solution is found. However, this approach requires constant monitoring or the implementation of a stopping criteria specified by the user to cut off the optimization procedure with good timing. In both cases, the final solutions are subject to the knowledge of the modelled watershed by the user.

Another possibility is to use the DDS algorithm with a lower simulation budget. The capacity to adapt the optimization to the user-specified budget is an interesting feature of the DDS algorithm. The developers of this algorithm (Tolson and Shoemaker 2007) prove that a lower total simulation budget normally leads to a faster decrease of the objective function value. However, Figure 4.11 shows that the probability of being trapped in a poor-quality zone increases when the number of simulations dedicated to the global exploration is reduced due to the adaptive feature. As mentioned in Section 5, multi-start strategies and parallelism may help to prevent such trapping issues, but that would involve the use of a multi-core processor. Another strategy is to make a specific sensitivity analysis on the simulation budget assigned to the DDS-Bench algorithm. A specific optimal simulation budget may be found for each optimization problem, but high computational times would then be involved.

The hybrid DDS-MADS calibration approach has four important advantages. First of all, it does not require any tuning of optimization algorithm parameters, or sensitivity analysis of those parameters. This means that a low level of a priori knowledge of the optimization problem is required, and that no computational time needs to be consumed before running the calibration procedure. The MADS step within the proposed hybrid approach is directly involved in the strategy to stop the calibration automatically depending on the quality of the objective function value. No constant monitoring and no manual cut-off are needed. Secondly, the hybrid approach is designed to run on a single-core processor without monopolizing all the computational power of a standard laptop. These functionalities contribute to a simplified implementation for operational usage and for research purposes. Thirdly, it is observed that the DDS and MADS algorithms are complementary in this hybrid approach. The capacity of DDS to globally explore the solution space quickly leads to a

good-quality zone of the parameter space. Then, when stagnation occurs in the global exploration step, MADS ensures local tuning of the calibration parameter set and converges to a nearby local optimum. The five transitional features are therefore proposed to smooth out the calibration procedure from global to local search. Moreover, the MADS step may recover from global exploration failures from the DDS step when it gets trapped in poor-quality zones due to its initial mesh size adjustment feature (as illustrated in Figures 4.5 and 4.9). Fourthly, the main feature of this hybrid calibration approach is its significant reduction in computational time without sacrificing robustness and stability of the final objective function values. For the three studied watersheds combined, average time savings of 70% on HYDROTEL 10 and 40% on HYDROTEL 19 were achieved. This should facilitate the use of computationally intensive hydrological models in operational or research contexts.

Another important point about the DDS and MADS algorithms is the accessibility of their source codes and the technical support offered. In both cases, these algorithms are freely and easily available on the authors' websites. The Matlab source codes for the DDS algorithm are available in open source by emailing the first author of Tolson and Shoemaker (2007). Releases are updated and technical support is always available. The latest release allows solving box-constrained optimization problems with continuous, integer or mixed decision parameters. With the large community around the MADS algorithm, the authors service a website dedicated to this optimization tool. The MADS algorithm is downloadable from the Matlab OPTI Toolbox (open source code), as well as in C++ for multi-platform applications. Multiple functionalities are continuously added to MADS and it is accompanied by a clear user guide. Different types of decision parameters including categorical variables, Variable Neighborhood Search (VNS) and quadratic or surrogate models are just a short list of the additional functionalities.

The hybrid DDS-MADS approach makes possible the transposition of the hybrid approach on biobjective and/or parallel computing networks optimization. The authors of both methods propose different versions or modalities offering these possibilities. The DDS authors propose the multi-objective "Pareto Archived Dynamically Dimensioned Search" (PADDS;

Asadzadeh and Tolson 2013) version, the "Asynchronous Parallel Dynamically Dimensioned Search" (PDDS; Tolson et al. 2007), a parallel computing networks method, and the "Parallel Pareto Archived Dynamically Dimensioned Search" (ParaPADDS; Tolson et al. 2015), combining multi-objective optimization and parallel computing. The NOMAD software package also provides tools for biobjective optimization (Audet et al. 2010) as well as for its utilization in a parallel environment (Audet et al. 2008).

Due to intensive computational cost of all experiments conducted in this study, the development of the hybrid DDS-MADS calibration approach is based on some methodological choices which impact the results presented in this study. Additional computationally intensive hydrological models and other optimization problems will be tested in future works to extend the conclusions and the potential use of DDS-MADS to wider domains. The Nash-Sutcliffe criteria was chosen since it is a "state-of-the-art" objective function. Many other objective functions are used in hydrology, and have different properties. The impact of the objective function criterion on calibration results would also be interesting to study to better investigate the dependence between the objective function and the optimization method used.

As mentioned in the literature review, Razavi et al. (2010) introduced four different approaches for dealing with the issue of computationally intensive problems. Future research efforts will focus on the last one: the implementation of surrogate models within the optimization procedure. Still working with computationally-intensive hydrological models, two broad families of surrogate models will be analyzed: (1) the lower-fidelity physically based surrogates, which are a simplification of the original model; and (2) the response surface surrogates, which are statistical or empirical data-driven models reproducing the response of the original model (Razavi et al. 2012). Retaining the goal of reducing the computational time, future research will evaluate the representativeness of a wide range of surrogate models built from the original HYDROTEL model. The representative surrogate models will then be implemented inside different optimization algorithms or calibration

approaches. Mixing surrogate models and more computationally-efficient optimization approaches is not ruled out.

4.7 Acknowledgements

The authors would like to sincerely thank Mr. Bryan A. Tolson for providing the source codes of the DDS algorithm and for multiple wonderful discussions about hydrological models calibration. We would also like to thank Mr. Christophe Tribes and Mr. Sébastien Le Digabel for their precious help with the NOMAD software. Final thanks to the reviewers and to the associate editor for constructive comments which helped to substantially improve the quality of this paper.

The meteorological gridded datasets and daily observed streamflows for the studied watersheds were provided by the Direction de l'Expertise Hydrique (DEH) and Hydro-Québec. Funding for this study was provided by a PhD scholarship from the National Sciences and Engineering Research Council (NSERC) of Canada.

CHAPITRE 5

LOW-COST AND REPRESENTATIVE SURROGATE HYDROLOGICAL MODELS, PART I: CONSTRUCTION OF SURROGATES

Pierre-Luc Huot[1], Annie Poulin[1], Charles Audet[2] et Stéphane Alarie[3]

[1] Département de la Génie de la Construction, École de technologie supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3.
[2] Département de Mathématiques et Génie Industriel, Polytechnique de Montréal et GERAD,
2900 boul. Édouard-Montpetit, Montréal, Québec, Canada H3T 1J4.
[3] Institut de Recherche d'Hydro-Québec,
1800 boul. Lionel Boulet, Varennes, Québec, Canada J3X 1S1.

Article soumis pour publication dans la revue « Environmental Modelling & Software »,
décembre 2018

Abstract

Dealing with computationally-intensive calibration processes is still common in distributed hydrological modelling despite the computing power growth. Computational time for one single distributed hydrological simulation may easily consume more than 1 minute, and the calibration process can require thousands of simulations. The use of surrogate models that are low-cost and representative of the calibration process is an interesting avenue to reduce the computational time. The first part of this study explores three possibilities to construct reduced-fidelity surrogate models from the HYDROTEL model, a computationally-intensive hydrological model. The relevance of these three types of surrogates and their combination within a calibration process is evaluated according to the best compromise between representativeness and a decrease of CPU time. In a second paper, surrogate models are implemented within an existing efficient calibration process to significantly reduce the computational time.

Keywords: Distributed hydrological model, Computationally-intensive simulation model, Efficient Calibration, Surrogate model, Reduced-fidelity model

5.1 Introduction

Increasing the efficiency of hydroelectricity production, delimiting flood zones territories near water systems, managing reservoir water levels for water accessibility and environmental constraints, and studying impacts of climate change on water resources are expert domains that depend on the comprehension of the water cycle processes (Singh & Woolhiser 2002, Moradkhani & Sorooshian 2008). Nowadays, the acquisition of water resources knowledge involves the use of simulation tools, such as hydrological models that simulate various hydrological processes at the catchment area or finer scale (Singh & Woolhiser 2002). These models are essential because they allow to better understand the spatialization and temporal distribution of key resources on a specific territory. In addition, these tools provide suitable information to guide decision-makers for current and future water resource issues. The work proposed in this study aims at improving and developing hydrological modelling tools that affect fundamental issues for an efficient development of modern society.

Diversification in hydrological model applications has led to the design of several types of models according to their level of spatial discretization (Moradkhani & Sorooshian 2008). Ranging from a single and global territory entity (global models) to a fully-distributed area into fine mesh elements (distributed models), each spatial discretization levels has its pros and cons depending of its use. Global models require a small amount of input data for modeling and are extremely fast to compute, but the spatial variability in the catchment is ignored. Distributed models often rely on the use of fine-resolution databases from remote sensing or geographic information systems (GIS) to spatialize catchment scale modelling and offer a good representation of the spatial variability of hydrological processes (Singh & Whoolhiser 2002, Moradkhani & Sorooshian 2008, Das et al. 2008, Pechlivanidis et al. 2011). Nonetheless, this advantage can lead to an important cost in computational time depending on the chosen spatial discretization level. Knowing that a single simulation may require several minutes of computing time, a full calibration process can become computationally intensive (Mugunthan et al. 2005, Zhang et al. 2009, Razavi et al. 2010,

Huot et al. 2017). This can provide a major impact in operational or research contexts, which forces users to choose computationally efficient calibration methods.

5.1.1 Literature review

Razavi et al. (2010) described four avenues to deal with computationally-intensive calibration processes: (1) the development of efficient optimization strategies to reduce the number of simulations executed, (2) the use of model preemption, which interrupts model simulations identified as being of low-quality, (3) the use of low-cost surrogate models to evaluate the potential quality of solution points, (4) the use of parallel computing networks. A fifth avenue to be considered consists in the technology progress in computing power. A simulation model running today is clearly less computationally-intensive than it was 10 years ago. For instance, the CPU time that was 3 hours and 45 minutes (Sun Blade 100, 500 MHz) for solving a collection of 55 optimization problems 10 years ago is now 5 minutes and 40 seconds on a recent personal computer (Intel Core i7, 3.40 GHz) with the same optimization algorithm (Audet & Orban 2006, Audet et al. 2018).

The literature exposes several uses for each of the first four avenues to deal with computationally intensive problems (Booker et al. 1999, Shoemaker et al. 2007, Razavi et al. 2010, Razavi et al. 2012, Regis & Shoemaker 2013, Razavi & Tolson 2013, Huang et al. 2014) and other studies combine some of these approaches (Regis & Shoemaker 2009, Razavi et al. 2010, Regis & Shoemaker 2013, Audet 2014). The present study focuses on mixing the use of the low-cost surrogate models and the development of more efficient optimization strategies.

Razavi et al. (2012) define surrogate models as a simplification or approximation of a simulation model and having the characteristic of being less intensive in terms of computational time than the original simulation model. Audet and Hare (2017) define a surrogate as a function that shares similarities with the objective function of a simulation

model, which is less intensive in computational time and which is useful for reducing the number of simulations within an optimization process.

There are two broad families for designing surrogate models: the use of response surface surrogates emulating the objective function surface from a history of solution points (sets of parameters), or the use of reduced-fidelity physically-based models a priori designed, also called low-fidelity models or emulated models (Booker et al. 1999, Regis & Shoemaker 2007, Razavi et al. 2012, Wang et al. 2014). Le Digabel (2011) also distinguishes these two types of surrogate models: adaptive functions, corresponding to response surface surrogates, and non-adaptive functions referring to the reduced-fidelity models.

The conclusion of the survey paper of Razavi et al. (2012) exposes the underutilization of reduced-fidelity models within the water resource community in spite of its high promising potential to help users of time-consuming models, and there are multiple advanced strategies in broader research community. The reason behind this may be related to a lack of versatility by the reduced-fidelity surrogates. These non-adaptive surrogates are clearly not polyvalent because each surrogate is a particular simplification of a specific optimization problem and can not be easily transposed to any other optimization problems, even if many similarities are observed (Razavi et al. 2012, Wang et al. 2014, Leifsson & Koziel 2015). Another reason may be the necessity of a high level of computer programming knowledge, computer simulation sequences and domain-specific principles to provide the best compromise between representativeness and a decrease of CPU time.

5.1.2 Research objectives and contributions

The present work is part of a larger research project composed of two papers aiming to combine the use of surrogate models inside more efficient optimization strategies to deal with the computationally-intensive calibration process of a hydrological model. The first part studies the development of reduced-fidelity models from the hydrological model HYDROTEL (Fortin et al. 2001a). This paper explores, evaluates and combines three

different methods to construct surrogate models from the original hydrological model. The first type of surrogate model reduces the number of meteorological gridpoints from a gridded dataset on the territory of the modeled watershed, the second type reduces the calibration time-period and the third one reduces the watersheds spatial discretization by decreasing the number of simulation units within the model application. Representativeness and computational time for each type of surrogate model and for the combination of all of them are evaluated and analysed. A second paper presents the implementation of such reduced-fidelity surrogate models within the hybrid optimization approach DDS-MADS (Huot et al. 2017), which has proven to be efficient for the calibration of HYDROTEL.

5.1.3 Paper organization

This paper is organized as follows. Section 5.2 presents the hydrological model HYDROTEL and its particular model structure, followed by the method used to evaluate the representativeness and the computational time consumed by the surrogate models. Section 5.3 exposes the hydrological modelling methodology: modeled watersheds, computational time benchmark, necessary input data and objective function. Section 5.4 details the computational experiments evaluating the representativeness and computational time for each type of reduced-fidelity surrogate models and all combined. Connections with the follow-up paper are discussed in the closing section.

5.2 The HYDROTEL problems

This project considers the distributed, physically-based and computationally intensive hydrological model, named HYDROTEL, as the main model to be calibrated (Fortin et al. 2001a, b). This model is based on a modular approach with various submodel options, and is compatible with remote sensing and/or GIS data. The spatial variability of land cover, land use, natural hydrographic system, soil type and topography is distributed on the territory into a gridded mesh at a low distance resolution. Cells of this gridded mesh are rallied in several simulation units named RHHUs (Relatively Homogenous Hydrological Units), similarly to a division of the territory into many sub-basins. All hydrological processes are simulated

independently on each RHHU except the channel water routing on existing hydrological system. In a first paper, Fortin et al. (2001a) describes mathematical equations governing each hydrological submodels available, as much as possible physically-based, but also conceptual or empirical in nature. A second paper gives a step-by-step watershed modelling example (Fortin et al. 2001b).

Many of the HYDROTEL submodels parameters can be fixed either because they are optional, but in some cases can be useful as corrective factors, or because they have a low-sensitivity impact on the quality of simulations. Ricard et al. (2012) and Poulin et al. (2011) present a 12 calibration parameters version which includes the most sensitive parameters. In this study, two HYDROTEL problems are used in order to evaluate the impact of dimensionality on the representativeness and CPU time of potential surrogate models. The first one includes the 10 most sensitive parameters (HYDROTEL 10) and the second problem has 9 additional parameters (HYDROTEL 19). Appendix I (referred here to Table 4.1) presents the selected parameters and their bounds for each one of these two HYDROTEL calibration problems. More specific information about both HYDROTEL problems can be found in Huot et al. (2017) and Fortin et al. (2001a, b).

5.2.1 Model structure

The HYDROTEL modular structure provides flexibility on simulating various available submodels and allows the addition of new ones by the user. Hydrological submodels used in this study relate to the three computing steps presented in Table 5.1. The HYDROTEL structure sequences six submodels simulating the hydrological processes: (1) interpolation of meteorological data, (2) snow cover estimation, (3) potential evapotranspiration, (4) vertical water balance, (5) overland water routing and (6) channel water routing on the existing hydrographic system. The first four submodels are simulated independently on each RHHU and represent the first computing step of the hydrological model routine. The second step consists of the overland water routing simulated on each RHHUs towards its associated river reach. The third one calculates the channel water routing on the entire existing hydrographic

system. The three sequential computing steps provide the opportunity to re-simulate the following computing step without having to simulate again the previous one when no calibration parameter was modified.

Table 5.1. Submodels used in each computing steps of the HYDROTEL routine and impact on computational time consumed when using the proposed surrogate models

			Increasing or Decreasing the Computational Time by Reducing :		
Computing Steps	**Hydrological Process**	**Submodels Used**	Simulated Period	Number of RHHUs	Number of Meteo Stations
STEP 1	Interpolation of Meteorological Data	Thiessen Polygons	↓	↓	↓
	Snow Cover Estimation	Mixed (degree-day) energy-budget method	↓	↓	
	Potential Evapotranspiration	Hydro-Quebec method	↓	↓	
	Vertical Water Balance	BV3C	↓	↓	
STEP 2	Production of Geomorphological Hydrographs when relative parameters are modified	- N/A -	↓	↑	
	Overland Routing	Kinematic Wave Equation	↓	↓	
STEP 3	Channel Routing	Kinematic Wave Equation	↓		

A particularity in the overland water routing (step 2) requires special attention. First, the production of a specific geomorphological hydrograph based on the routing of a reference water depth over all cells in the flow direction structure of each RHHU is necessary (shaded section in Table 5.1, step 2). Then, the water depths available, at each time step, are distributed according to these hydrograph patterns to provide the lateral inflow to the hydrographic reach to which the RHHU is associated. The production of the geomorphological hydrograph is computationally-intensive especially because it is independently calculated on each RHHU. The creators of HYDROTEL established a file storage procedure to decrease this time-consuming step. On every simulation, the HYDROTEL routine chooses the appropriate RHHU hydrograph file according to the 4 related calibration parameters of the overland water routing (see Table 4.1 for their definition). If a new combination of these related calibration parameters is to be simulated, a new RHHU geomorphological hydrograph is calculated for each RHHU and they are all

stored in a single new file. In the HYDROTEL 10 problem, none of these 4 related calibration parameters are included, but 2 are included in the HYDROTEL 19 problem. So, a geomorphological hydrograph for each RHHU will be produced only on the first simulation of HYDROTEL 10. New hydrographs are produced on each new combination of these 2 related calibration parameters for the HYDROTEL 19 problem. This situation can generate a 5% to 25% CPU time increase for a simulation with HYDROTEL 19 versus HYDROTEL 10.

5.2.2 Representativeness of surrogates versus computational time

It is necessary to clearly define what is a good surrogate model from objective criteria. Going back to the definitions presented above, two conditions are mandatory: (1) a high level of similarity/representativeness between the objective function of the surrogate and the original simulation models and (2) an important decrease in computational time. In other words, a high representativeness and a low time-consuming simulation lead to the ideal surrogate model. Both are discussed below.

Representativeness of surrogates

The representativeness of forthcoming surrogate models will be assessed by simulating 5,000 uniformly distributed parameter sets in the parametric space generated with the Latin Hypercube Sampling (LHS; McKay et al. 1979), comparing both objective function values of the original and surrogate models. As for Razavi and Tolson (2013), the R-Square coefficient (R^2) is commonly used for qualifying the level of representativeness since it provides information on the goodness of fit between the original and the surrogate models. As a guideline, Toal (2015) suggests to use surrogates that have a reasonably high R^2 (greater than 0.9). It is important to note that this is a guideline for response surface function surrogates, thus not directly specified for reduced-fidelity models. However, as can be seen in Figure 5.1, a high R^2 value does not guarantee a high representativeness near the optimal zones. Overconfidence in the R^2 obtained by sampling 5,000 points from the parametric space could lead to a surrogate that quickly becomes useless in a calibration process, particularly if the

optimal zones are not well represented. Figure 5.1 presents two hypothetical cases in which the surrogate function A gives a better R^2 coefficient than the surrogate B. Despite this, important gaps are observed between the different optima on surrogate A. It would be preferable to use surrogate B in a calibration process since the optima of the surrogate and original functions are aligned. The surrogate model must share similarities with the original model meaning that the behavior of the surrogate model must represent as much as possible the behavior of the original model, including the alignment and the ordering of the optimal zones. Hence, sharing similarities does not mean that the objective function values obtained on the surrogate versus the original must be as similar as possible.

To counteract this situation, a second coefficient, named the Spearman's rank correlation coefficient (Rs), assesses how well two variables are monotonically related. In other words, the Spearman coefficient measures a statistical dependence between the rank values (best rank to worse) of two variables (Fieller et al. 1957); in this study, variables are the objective function values of the original and surrogate models. The Spearman coefficient (Rs) ranges from -1 to 1, where $Rs = 1$ represents a perfect ranking fit between original and surrogate model, and $Rs = -1$ is a fully opposed correlation. Figure 5.1 illustrates that situation where surrogate B outperforms surrogate A in terms of the Spearman coefficient while it's the opposite result with the R^2 coefficient. These hypothetical cases illustrate that both correlation coefficients are thus useful and complementary to assess the representativeness of the surrogate models. The book of Audet and Hare (2017) discusses the similarities needed to properly exploit surrogate models within an optimization process.

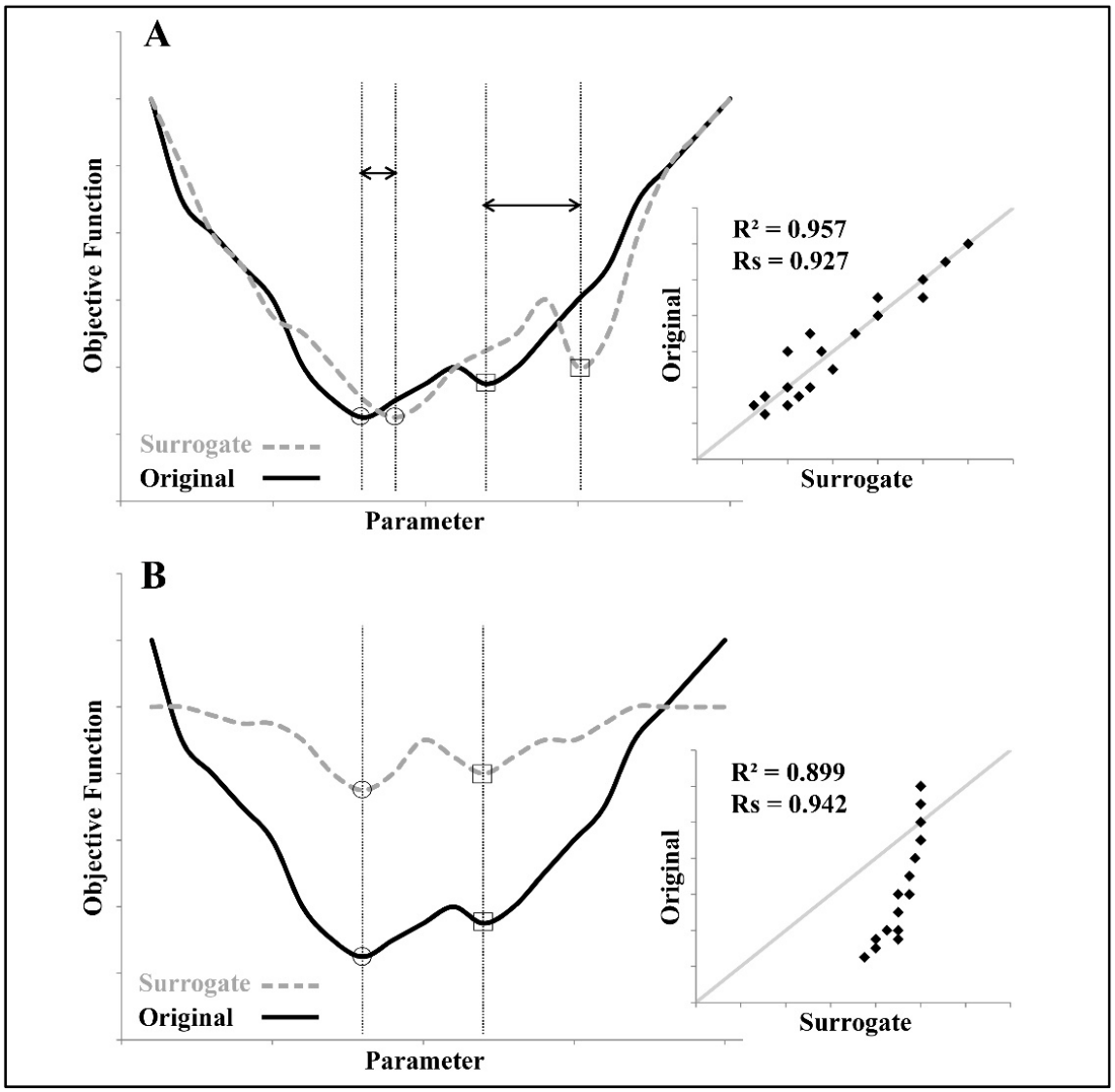


Figure 5.1. Representativeness of two hypothetical cases illustrating original versus surrogate functions (A & B) with R-Square coefficient (R^2) and Spearman's rank correlation coefficient (Rs)

Finally, representativeness can also be graphically evaluated by drawing a scatter plot of the 5,000 objective function values of the original versus the surrogate models. Outlier points and an extent of discrepancy in the scatter patterns can be visually identified but often toned down in correlation coefficient. This observation could be an indicator of an ill-represented optimal zone in the surrogate model or of the addition of a non-existent one. Both cases complicate the use of the surrogate in the calibration process.

Computational time

Each surrogate model designed will be simulated 1,000 times to get an average computational time to eliminate possible variations in computer performance and computational time related to the production of geomorphological hydrographs or not. Obviously, the lower the computational time is, the more interesting the surrogate model is with respect to a full calibration process. The ratio between the CPU time of the original and surrogate models, named hereafter the CPU time ratio, has a major impact on the potential for the use of each surrogate model. The higher the CPU time ratio is, the more surrogate simulations can be rooted without weighing down the calibration process. Razavi and Tolson (2013) proposed a hydrological model calibration framework exploiting short but representative data time series to calibrate a problem for which longer data time series are available. The CPU time ratio between short data time series model (surrogate model) and long data time series model (original model) may be evaluated between 6 to 8 times less intensive in computational time. This level of CPU time ratio conducts to more efficient calibration tools.

In this paper, five ranges of CPU time ratios are considered. Let r be the ratio of the CPU time used by the original model over that of the surrogate model. The following adjectives will be used to qualify this ratio:

$$\begin{aligned} Unusable &\quad if\ r \leq 1.5 \\ Poor &\quad if\ 1.5 < r \leq 4 \\ Medium &\quad if\ 4 < r \leq 8 \\ Superior &\quad if\ 8 < r \leq 15 \\ Ideal &\quad if\ r > 15 \end{aligned}$$

5.2.3 Potential surrogate models

We propose three potential simplifications to design surrogate models: (1) reducing the input meteorological dataset (i.e. reducing the spatial resolution of the gridded dataset) to be interpolated by HYDROTEL on the watersheds, (2) using a shorter calibration time-period,

(3) decreasing the spatial discretization by reducing the number of RHHUs (sub-basins). Individual and combinations of these simplifications are also assessed to obtain the best CPU time ratio available while maintaining a high level of representativeness.

Number of Meteorological Stations

The HYDROTEL problems are designed with a meteorological gridded dataset at a low distance resolution, which results in a large number of "pseudo-meteorological" stations (hereafter simply called stations or meteorological stations) on the territory. In operational or research contexts, this of course can contribute to an adequate simulation of the streamflows, but for calibration purposes, the surrogate model does not need the same accuracy. Table 5.1 shows that reducing the number of meteorological stations on the territory provides a decrease in computational time only in the interpolation of the meteorological data submodel. Therefore, one could expect a poor or medium CPU time ratio. A checkerboard pattern is used to withdraw meteorological stations in order to maintain a good distribution of the dataset on the territory and to simplify the design of this surrogate model type.

Calibration Time-Period

Obtaining gain in computational time with a shorter calibration time-period is not surprising. However, a high level of representativeness may be more difficult to reach considering the interannual variability of the hydrologic cycle. Initial conditions on the watershed of a short calibration time-period have also an important impact on representativeness, so these initial conditions need to be adequately set on every surrogate models. On the HYDROTEL problems, running a short calibration time-period should decrease the computational time for all submodels resulting in a medium or superior CPU time ratio (see Table 5.1). For this type of surrogate model, the representativeness level is more challenging than obtaining interesting CPU time ratios.

Number of RHHUs

In the HYDROTEL problems, the division of the territory into several RHHUs is based on the overland flow directions structure obtained with the modified digital elevation model and

the digitized river network. The RHHUs division is automatically produced in the preparation of a watershed database using a threshold, specified by the user, indicating the maximum number of cells upstreaming each RHHUs. This preparation parameter gives the opportunity to modify the number of RHHUs within the watershed territory by adjusting its value. As the hydrological processes are simulated on each RHHU, a smaller number of RHHUs could speed-up the simulation for all submodels. However, Table 5.1 suggests the opposite for the production of the geomorphological hydrographs. A smaller number of RHHUs results in RHHUs with a larger distance between the upstream and the downstream cells. Routing the reference water depth over all cells on a longer flow distance on each RHHU is more time-consuming. Utilization of this type of surrogates leads to the following question: could the gain in computational time obtained in the routine of all submodels with a smaller number of RHHUs compensate the intensive computational time assigned to the production of longer geomorphological hydrographs? Obviously, this situation is relevant to HYDROTEL 19 since the hydrographs are frequently re-calculated. With a high level of uncertainty in this surrogate type, poor CPU time ratio for HYDROTEL 19 may be anticipated, but a medium or superior ratio for HYDROTEL 10.

5.3 Original calibration problems

This section describes the six original calibration problems studied in this paper. The two formulations of the HYDROTEL problem with 10 and 19 calibration parameters in Section 5.2 are tested on three watersheds with natural streamflows, all located in the province of Québec (Canada): the Ceizur watershed with a 6,928 km² area which is an upstream sub-basin of the Gatineau River (west of the province), the Cowansville watershed with a 215 km² area which is an upstream sub-basin of the Yamaska River (south of the province) and the Toulnustouc watershed with a 8,109 km² area which is an upstream sub-basin of the Manicouagan River (north of the province). All three watersheds differ in their soil types, land cover, land uses and topography. The six combinations of “Watershed-HYDROTEL problem” establish the original calibration problems and are originally developed by Huot et al. (2017).

5.3.1 Computational time

As for the potential surrogate models in section 5.2.3, the computational time for one single simulation of the original hydrological model is obtained by averaging the computational time of 1,000 simulations with random calibration parameters on a 3.40 GHz Intel Core i7 processor with 12 Go of RAM. Table 5.2 summarizes all averaged computational times for all combinations of "Watershed-HYDROTEL problem". Recall that HYDROTEL 19 includes the production of the geomorphological hydrographs, which is intensive in computational time when related calibration parameters are modified.

5.3.2 Meteorological data and calibration time-periods

Meteorological data, including daily minimum and maximum temperatures and daily precipitations, were obtained from gridded datasets (see the acknowledgements section for data sources) and Table 5.2 shows the number of meteorological stations in each watershed. Calibration time-period spans from October 1, 1988 to September 30, 1992 for the Ceizur watershed (4 years), from October 1, 2000 to September 30, 2005 for the Cowansville watershed (5 years) and from October 1, 1984 to September 30, 1988 for the Toulnustouc watershed (4 years). Daily streamflow at the outlets are observed on the same time-periods. The initial number of RHHUs for each watershed was previously designed by industrial partners (see acknowledgments section for data source) in order to optimize the quality of modelling and to maintain moderate computational time. The obtained numbers of RHHUs are also shown in Table 5.2.

Table 5.2. Summary features of the six original simulation model problems relative to three studied watersheds modeled on HYDROTEL 10 and HYDROTEL 19

Watersheds	HYDROTEL 10 Computational Times [a b]	HYDROTEL 19 Computational Times [a c]	Number of Meteo Stations	Simulated Period	Number of RHHUs
Ceizur	120.84 sec	116.18 sec	210	4 years	230
Cowansville	26.81 sec	40.96 sec	8	5 years	89
Toulnustouc	49.43 sec	89.05 sec	103	4 years	197

[a] Computational times for one single simulation on a 3.40 GHz Intel Core i7 processor with 12 Go of RAM.

[b] Computational times do not include the production of the geomorphological hydrographs.

[c] Computational times include the production of the geomorphological hydrographs.

5.3.3 Objective function

The research context of this paper uses of hydrological models for streamflow predictions at the discharge of the watershed outlet. The quality of each simulation is thus assessed by comparing the streamflow predictions from hydrological models with the observed streamflow data on the same time-period at the watershed outlet. Moriasi et al. (2007) and Servat and Dezetter (1991) recommend the Nash Sutcliffe Efficiency (NSE; Nash & Sutcliffe 1970) as quantitative statistical criterion (objective function) because it reflects the overall fit between predicted and observed streamflows and it is very commonly used in hydrological modelling (Mugunthan et al. 2005, Razavi et al. 2010, Pechlivanidis et al. 2011, Pushpalatha et al. 2012, Ricard et al. 2012, Arsenault et al. 2014, Huot et al. 2017, Poissant et al. 2017, Bajamgnigni Gbambie et al. 2017). The objective function used in this paper consists of minimizing 1 minus the Nash-Sutcliffe Efficiency ($1 - NSE$) and is described as follows:

$$1 - NSE = \frac{\sum_{i=1}^{N}(O_i - S_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O})^2} \tag{5.1}$$

on daily time-steps $i = 1,2,3,\dots,N$ of the calibration time-period, and where the residual variance is calculated between the observed streamflow value O_i and the simulated (predicted) streamflow value S_i and the observed data variance is calculated between the observed streamflow value O_i and the mean of observed streamflows $\bar{O}$. The value $1 - NSE$

determines the ratio of the residual variance (noise) to the observed data variance (information). The value $1 - NSE$ ranges from 0 to ∞, with $1 - NSE = 0$ indicating a perfect fit between observed and simulated values. When $1 - NSE > 1$, the mean of observed streamflows displayed on all the time-period is considered as a better quality hydrograph than the simulated streamflows.

5.4 Computational experiments

Each computational experiment presented below is based on simulation results of 5,000 parameter sets uniformly distributed in the parametric space generating with the LHS. The objective function values $(1 - NSE)$ obtained with the original (horizontal axis) and surrogate (vertical axis) models are graphically compared in scatter plots supporting by the Spearman coefficient (Rs) and the R-Square coefficient (R^2) to evaluate the representativeness. As mentioned previously, it is more important to focus on the representativeness of the optima zones in a perspective of using surrogate models in calibration processes. Therefore, the scatter plots inside figures are presented on adjusted axis zooming on optima zones. A good representativeness is considered when the 0.9 value is reached by Rs and R^2 as recommended by Toal (2015). The slanted line represents a perfect fit between original and surrogate models. The gray area is delimited by the best known value of the objective function for each original "Watershed-HYDROTEL problem". The average computational time by one single surrogate simulation and the CPU time ratios are presented in following tables. Results are presented on the Toulnustouc watershed because it represents the worst watershed candidate in terms of the representativeness for all types of surrogates.

5.4.1 Reducing the number of meteorological stations

Figure 5.2 illustrates the representativeness between surrogates having different levels of meteorological station removal and the original models for the Toulnustouc watershed on HYDROTEL 10 and 19. The withdrawing process of meteorological stations generates many possibilities of dataset distribution on the territory and variability into the number of

meteorological stations selected. Four different instances have been made on each level of meteorological station removal. Consequently, this figure presents the worst and the best instances for each level of meteorological station removal in terms of the values of the Rs and R^2 coefficients.

First, Figure 5.2 shows both Rs and R^2 coefficient tend to decrease as the number of meteorological station decreases. This deterioration is important when meteorological data is highly heterogeneous, especially on large territories. The two worst instances with 87,5% removed stations of the Figure 5.2 present this case. It should be noted that these two worst instances, which are below the 0.9 coefficient limit value, share exactly the same selected meteorological stations. The worst instance for HYDROTEL 10 at 87.5% of meteorological stations removed provides Rs and R^2 reaching their lowest values: 0.638 and 0.411, respectively. After investigation on these two worst instances, all meteorological stations have been removed in the West side of the territory due to the checkerboard pattern withdrawing process. Input data in this area are more heterogeneous then in the other parts of the territory. This could explain why the representativeness coefficients are lower for these two instances.

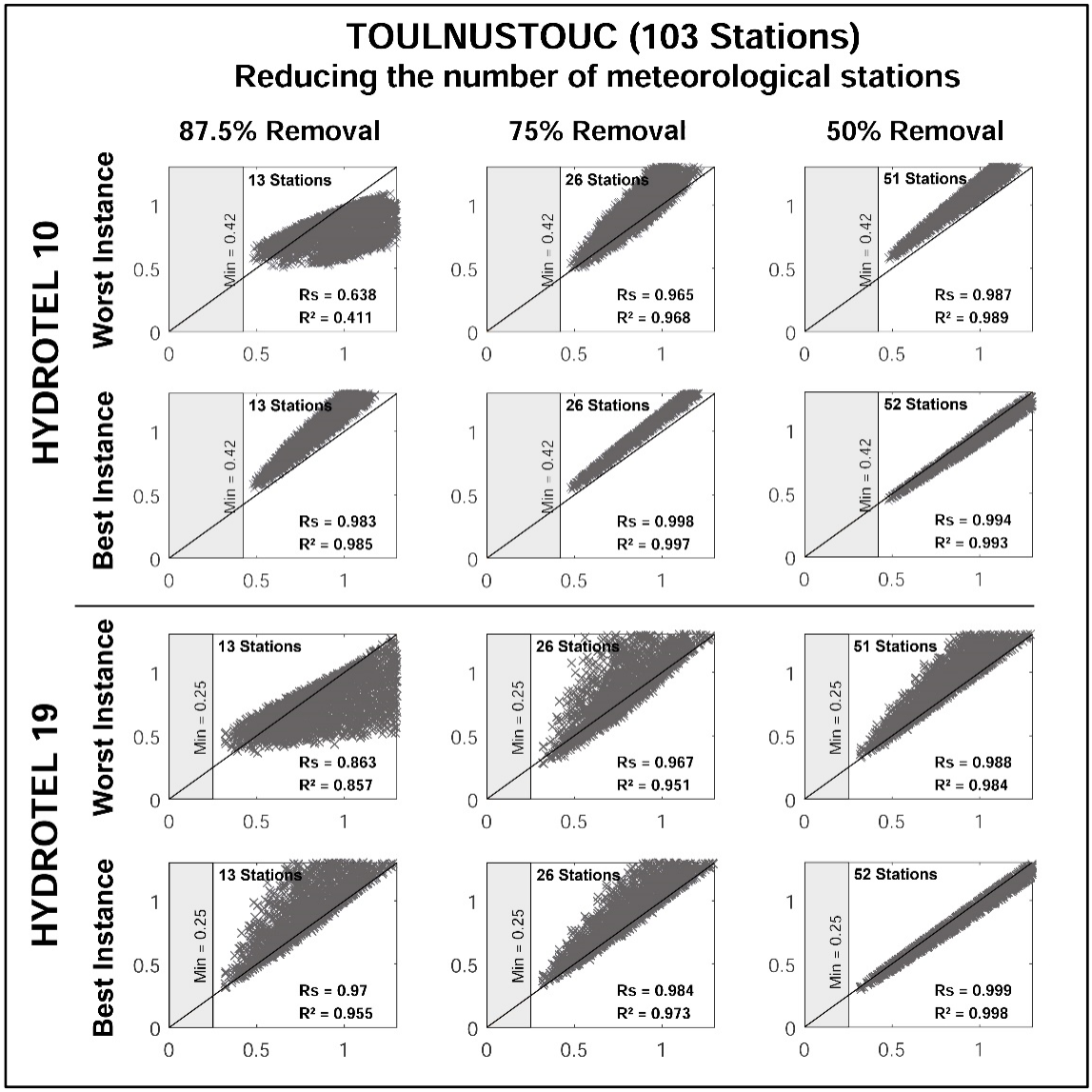


Figure 5.2. $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Toulnustouc watershed comparing the surrogates having a reduced number of meteorological stations with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation

Second, Figure 5.2 shows that when more meteorological stations are removed, the representativeness may decrease significantly (worst instances) or slightly (best instances), a sign of some variability in the results. For example, a wide difference in representativeness is obtained between worst and best instances at 87.5% level of stations removal for HYDROTEL 10. Indeed, as more stations are removed, the selection of meteorological

stations must be made with vigilance. Importance is granted to a systematic station removing process to maintain a good distribution of the data on the territory and to integrate the geographic and meteorological knowledge of the user on the territory.

Finally, except for the two cases mentioned above, interesting representativeness at 50%, 75% and 87,5% of removed meteorological stations are obtained with Rs and R^2 over the 0.9 limit value on every instance. The best instance at 50% stations removal on HYDROTEL 19 yields a representativeness close to perfect coefficients. Better results are obtained on the Ceizur and Cowansville watersheds where Rs and R^2 are higher than 0.9 for any instance at any level of stations removal (Annexe II).

Tables 5.3 and 5.4 suggest that this type of surrogate decreases the computational time but provides only poor CPU time ratios for any instance. Obviously, better CPU time ratios are associated with the surrogates having 87.5% of meteorological stations removed. Surrogates on the Ceizur watershed offer the best CPU time ratios, then the Toulnustouc and finally, Cowansville. This ranking is related to the initial number of meteorological stations on the original models; i.e. the interpolation of meteorological data from a huge number of stations consumes a big portion of the total computational time of the HYDROTEL simulating routine, as for the Ceizur watershed. Even if the percentage of meteorological stations removed is the same, the CPU time consumed decrease further on a huge number of stations. Conversely, the interpolation of meteorological data from a small number of stations as on the Cowansville watershed represents a very short computational time, about a few seconds, in the HYDROTEL simulating routine.

Table 5.3. Average CPU time for one single simulation and CPU time ratio of surrogates having a reduced number of meteorological stations for all "Watershed-HYDROTEL 10" combinations

		HYDROTEL - 10 Parameters		
		CEIZUR	**COWANSVILLE**	**TOULNUSTOUC**
Benchmark Models		120.84 sec	26.81 sec	49.43 sec
Surrogates Number of Meteo Stations	50% removed	79.75 sec poor	26.73 sec unusable	38.27 sec unusable
	75% removed	69.90 sec poor	26.20 sec unusable	36.15 sec unusable
	87.5% removed	66.13 sec poor	26.44 sec unusable	25.34 sec poor
CPU Time Ratio:		-∞, 1.5] = unusable,]8, 15] = superior,	]1.5, 4] = poor,]15, +∞ = ideal	]4, 8] = medium,

Table 5.4. Average CPU time for one single simulation and CPU time ratio of surrogates having a reduced number of meteorological stations for all "Watershed-HYDROTEL 19" combinations

		HYDROTEL - 19 Parameters		
		CEIZUR	**COWANSVILLE**	**TOULNUSTOUC**
Benchmark Models		116.18 sec	40.96 sec	89.05 sec
Surrogates Number of Meteo Stations	50% removed	92.35 sec unusable	40.45 sec unusable	77.22 sec unusable
	75% removed	85.51 sec unusable	39.53 sec unusable	74.78 sec unusable
	87.5% removed	76.13 sec poor	40.66 sec unusable	64.28 sec unusable
CPU Time Ratio:		-∞, 1.5] = unusable,]8, 15] = superior,	]1.5, 4] = poor,]15, +∞ = ideal	]4, 8] = medium,

With the results presented above, the best compromise for all "Watershed-HYDROTEL problem" combinations between representativeness and a decrease of CPU time is the surrogates having 87.5% of meteorological stations removed. High representativeness may be obtained if station removing process is completed thoughtfully, but CPU time ratios are

clearly not as good as we would like to use this type of surrogates inside efficient calibration processes.

5.4.2 Reducing the calibration time-period

Figure 5.3 illustrates the representativeness between surrogates with four different calibration time-periods: 6, 9, 12 and 18 months, and the original models for the Toulnustouc watershed on HYDROTEL 10 and 19. The original models for Toulnustouc include a calibration time-period of four years from 1 October 1984 to 30 September 1988. This same period has been divided into many calibration time-periods to perform the surrogate models. Start and end dates are carefully chosen to avoid starting or interrupting a hydrological event, especially on calibration time-period under 1 year. All calibration time-periods of surrogate models begin on a dry period and finish after a flow recession period. No initialization period has been rooted first before the calibration time-period considered because the initial known conditions are carefully set inside the HYDROTEL routine. Two instances with 18 months are selected: October 1, 1984 to March 31, 1986 and October 1, 1986 to March 31, 1988. Four instances with 12 months are considered beginning on October 1 of each year and ending on September 30 of the following year. Four instances with 9 months beginning on October 1 of each year and ending on June 30 of the following year and four instances with 6 months beginning on January 1 of each year and ending on June 30 of the same year are also considered. As the three watersheds chosen are highly influenced by the accumulation and melting of snow, calibration time-periods of 6 and 9 months focus on the period where spring flood occurs of peak flows. Other instances, not included in this paper, have been tested on short time-periods of 6 or 9 months focusing on other hydrological events (dry period) and worst representativeness have been obtained on every trial. Considering that several sub-time-periods can be partitioned inside the original 4-year time-period, Figure 5.3 presents only two different instances (the worst and the best values of Rs and R^2) on each calibration time-period previously described.

Figure 5.3 shows that the representativeness of the surrogate models may be importantly impacted by a calibration time-period under 1 year. Surrogate models with a calibration time-period of 6 months never have a Rs value exceeding 0.512, nor a R^2 value exceeding 0.392. The scatter plots also present a wide extend of discrepancy. With such results, it is clear that a short time-period of 6 months covering only part of the snow accumulation period (November and December snow is missing) and the melting period does not adequately represent 4 completed annual hydrological cycles of the original calibration time-period.

For the calibration time-period of 9 months, Figure 5.3 shows that acceptable representativeness may be obtained on both HYDROTEL problems achieving Rs and R^2 coefficients both close to the 0.9 limit value. However, much lower coefficients are also obtained. This lack of consistency makes difficult it to consider this calibration time-period as a potential useful surrogate model inside an efficient calibration process. Conclusions obtained on the Ceizur and the Cowansville watersheds for this two time-periods are similar.

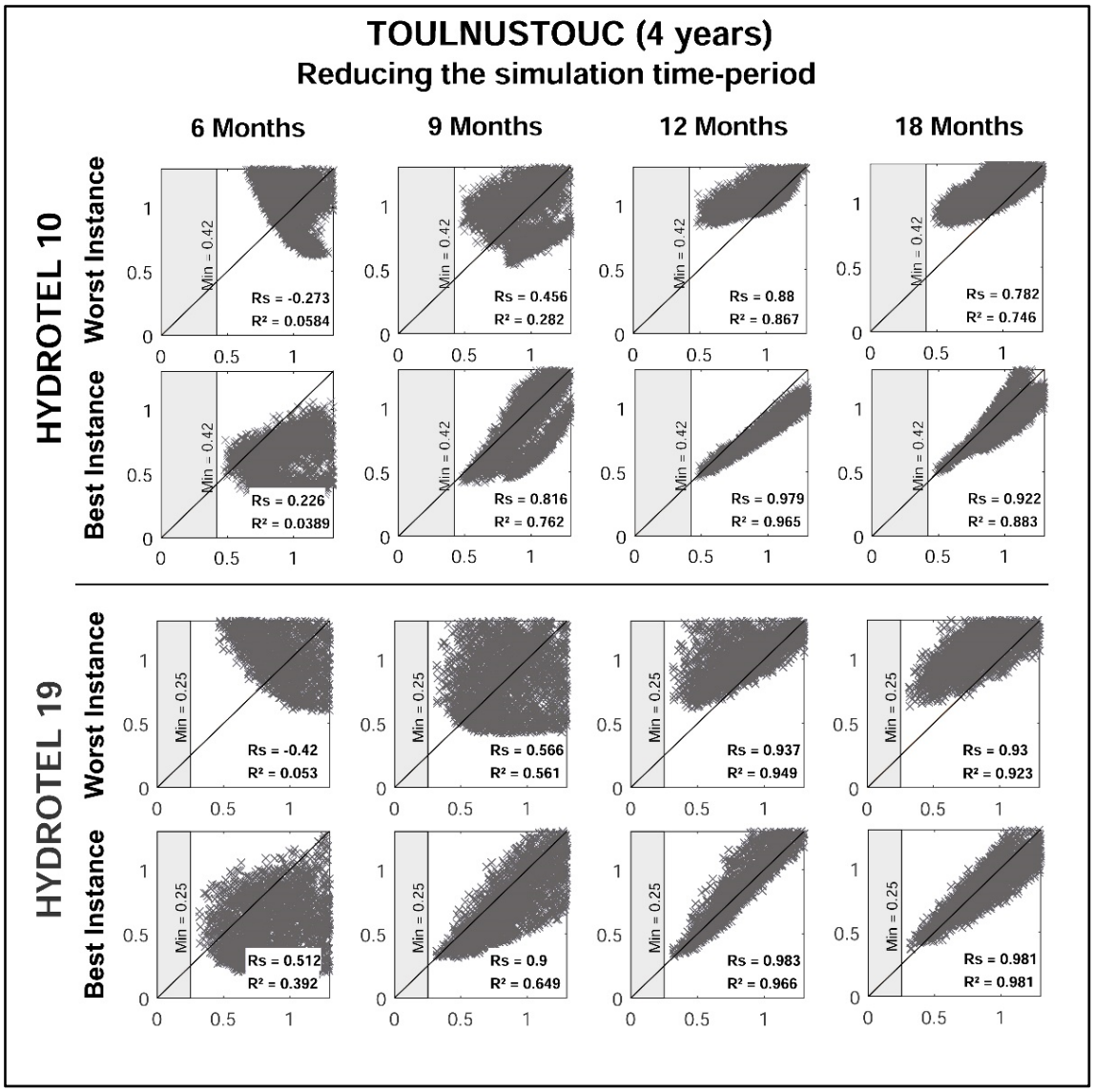


Figure 5.3. $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Toulnustouc watershed comparing the surrogates having a reduced calibration time-period with the original models. Rs and R² coefficients are calculated for the representativeness evaluation

Figure 5.3 shows that the best representativeness is reached with the 12 months time-period where the Rs and R² coefficients are very close or superior to the 0.9 limit value recommended by Toal (2015). This high level of representativeness obtained is related to the coverage of all hydrological events within the calibration time-period. Thus, it is easier to represent the average behaviour of the original calibration time-period consisting of 4 annual hydrological cycles in this case. It is important not to conclude that a time-period of 12

months is sufficient to well represent an original calibration period, but rather that the ratio between the original and surrogate calibration time-periods is short enough to maintain a high representativeness. Representativeness on the time-period of 12 months for the Cowansville watershed are slightly worst because the original calibration time-period is composed of 5 annual hydrological cycles versus 4 years for the other two watersheds.

Finally, Figure 5.3 shows that the calibration time-period of 18 months provides lower representativeness coefficients than the time-period of 12 months. As with the 6 or 9 months time-periods, an incomplete annual hydrological cycle in the calibration time-period of 18 months slightly reduces the level of representativeness.

Tables 5.5 and 5.6 show poor to superior CPU time ratios for all instances on all "Watershed-HYDROTEL problem" combinations because reducing the calibration time-period decreases the computational time of all submodels routed inside HYDROTEL sequence. On HYDROTEL 10, the CPU time ratios obtained are related to the ratio between the original and surrogate calibration time-period, i.e. a reduction by 4 times on the calibration time-period is translated by approximately a CPU time ratio of 4. Although there is not a perfect accuracy in this relation, it would be easy to evaluate approximately a CPU time ratio for any calibration time-period of this type of surrogate on HYDROTEL 10.

Table 5.5. Average CPU time for one single simulation and CPU time ratio of surrogates having a reduced calibration time-period for all "Watershed-HYDROTEL 10" combinations

		HYDROTEL - 10 Parameters		
		CEIZUR	**COWANSVILLE**	**TOULNUSTOUC**
Benchmark Models		120.84 sec	26.81 sec	49.43 sec
Surrogates Simulation Time-Period	18 months	46.13 sec	8.12 sec	16.38 sec
		poor	poor	poor
	12 months	34.81 sec	5.96 sec	12.71 sec
		poor	medium	poor
	9 months	27.64 sec	5.00 sec	8.14 sec
		medium	medium	medium
	6 months	15.96 sec	3.54 sec	5.38 sec
		medium	medium	superior
CPU Time Ratio:		-∞, 1.5] = unusable,	]1.5, 4] = poor,	]4, 8] = medium,
		]8, 15] = superior,	]15, +∞ = ideal	

Table 5.6. Average CPU time for one single simulation and CPU time ratio of surrogates having a reduced calibration time-period for all "Watershed-HYDROTEL 19" combinations

		HYDROTEL - 19 Parameters		
		CEIZUR	**COWANSVILLE**	**TOULNUSTOUC**
Benchmark Models		116.18 sec	40.96 sec	89.05 sec
Surrogates Simulation Time-Period	18 months	66.93 sec	21.72 sec	55.15 sec
		poor	poor	poor
	12 months	59.46 sec	19.57 sec	51.48 sec
		poor	poor	poor
	9 months	54.06 sec	18.64 sec	47.08 sec
		poor	poor	poor
	6 months	47.83 sec	17.28 sec	43.84 sec
		poor	poor	poor
CPU Time Ratio:		-∞, 1.5] = unusable,	]1.5, 4] = poor,	]4, 8] = medium,
		]8, 15] = superior,	]15, +∞ = ideal	

On HYDROTEL 19, this relation is not apparent because the CPU time ratios obtained are always poor (with a ratio close to 2) and a slight decrease in computational time is observed as the surrogate calibration time-period is reduced. The production of the geomorphological hydrographs on HYDROTEL 19 consumes an additional computational time causing an important deterioration of the CPU time ratios.

5.4.3 Reducing the number of RHHUs

The automatic process for the RHHUs division of the territory is based on the overland flow directions structure. Therefore, the preparation of this type of surrogates requires a lot of trials to balance the size of each RHHU versus the number of RHHUs. In this particular case, it is more difficult to anticipate the impact of RHHUs reduction on representativeness of the surrogates and CPU time ratios. Figure 5.4 shows results for the Toulnustouc watershed on HYDROTEL 10 and 19 at different percentages of reduction of the number of RHHUs. Unlike the previous figures, Figure 5.4 does not show the worst and best instances, because a single instance is considered.

The figure reveals that the representativeness decreases gradually as the number of RHHUs decreases. However, the representativeness on all instances is not so impacted by the RHHUs reduction process. Figure 5.4 presents an overall Rs greater than 0.961 and R^2 greater than 0.945 even with a high percentage of reduction of the number of RHHUs (87.5% of RHHUs removed on HYDROTEL 19). Higher representativeness is obtained on all instances of the Ceizur and Cowansville watersheds for both HYDROTEL problems.

A second observation is that the extent of discrepancy between the original and surrogate models on HYDROTEL 19 is more important than on HYDROTEL 10. Considering this, six instances have been made for HYDROTEL 19 with a gradual reduction on the number of RHHUs for better representation and only three instances with more aggressive percentages of reduction for HYDROTEL 10. According to our analysis, this difference is assigned to the production of the geomorphologic hydrographs (step 2 in Table 5.1) inside the HYDROTEL

routine when the calibration parameters relating with them are modified. On HYDROTEL 10, the calibration parameters related to the production of geomorphological hydrographs are not modified on the 5,000 comparison simulations, but the hydrographs are still produced once because the RHHUs subdivision of the territory have changed following the reduction process. So, on this HYDROTEL problem, the representativeness is evaluated between two different RHHUs subdivisions with their own hydrographs patterns. On HYDROTEL 19, the calibration parameters are modified many times on the 5,000 comparison simulations, which leads to the production of new hydrographs for the original and surrogate models every time the calibration parameters related with them are modified. The representativeness is evaluated between the original RHHUs subdivision having a multitude of hydrographs patterns depending on the values of the calibration parameters and the new RHHUs subdivision (surrogate model) having also a multitude of different calculated hydrographs patterns. This may explain why the impact on representativeness is higher on the HYDROTEL 19 problem.

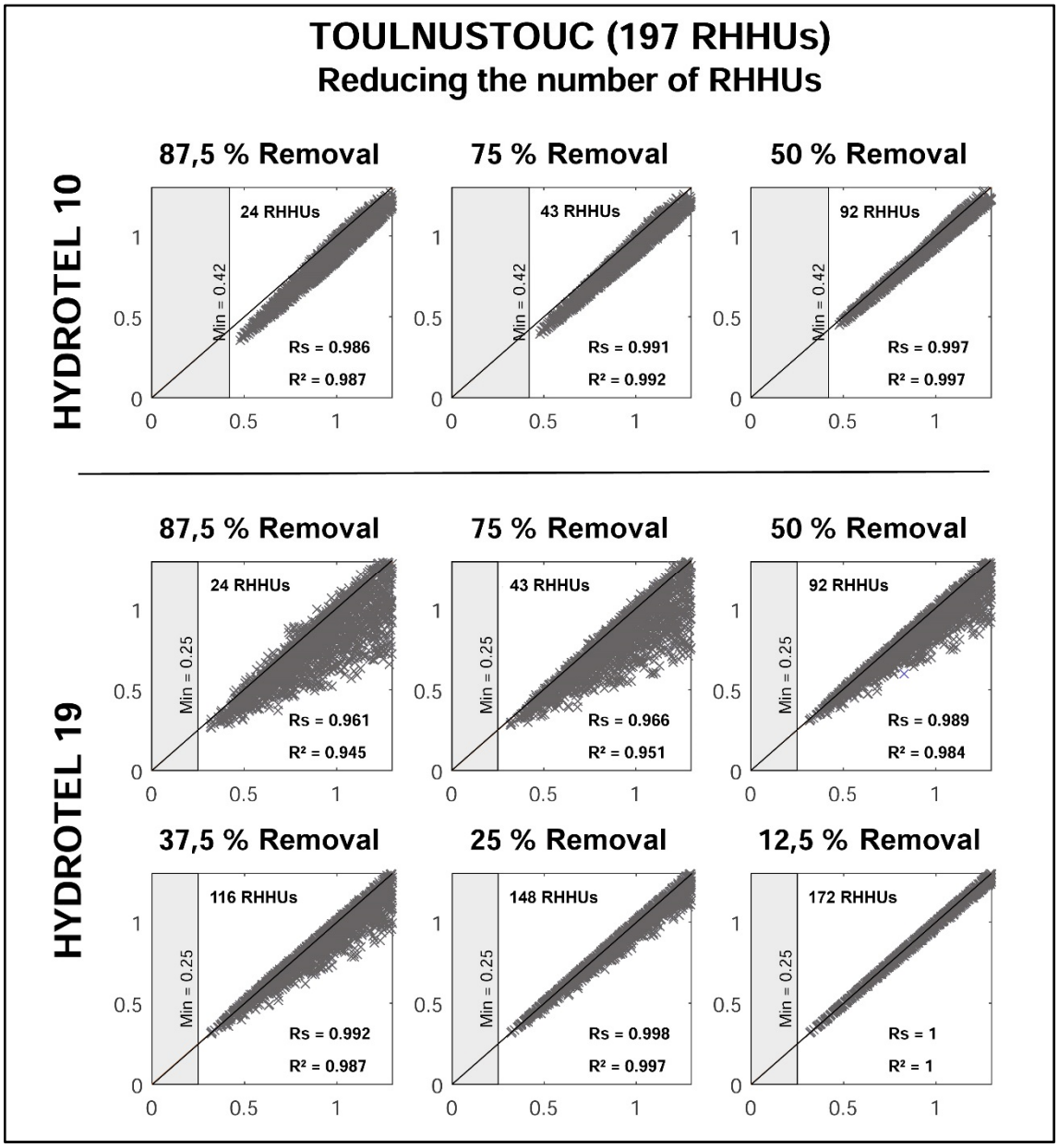


Figure 5.4. $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Toulnustouc watershed comparing the surrogates having a reduced number of RHHUs with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation

Tables 5.7 and 5.8 present two opposite scenarios for HYDROTEL 10 versus 19. Results for HYDROTEL 10 suggest that reducing the number of RHHUs decreases the computational time and improves the CPU time ratios. Medium CPU time ratio is obtained for the

Cowansville watershed and poor ratios are obtained for Ceizur and Toulnustouc with 87,5% RHHUs removal. A high representativeness and better CPU time ratios may indicate that this surrogate may provide acceptable assets to use it inside efficient calibration processes for HYDROTEL 10.

Results for HYDROTEL 19 show that reducing the number of RHHUs is problematic. As mentioned in Section 5.3.3, the main question is about the possibility of reaching a low-cost surrogate simulation with the particular routine of HYDROTEL 19, i.e. that introduces less computing of hydrological processes due to a smaller number of RHHUs, but that also involves the production of longer geomorphological hydrographs. Tables 5.7 and 5.8 give the answer to this question with unusable CPU time ratios for each instance. Slight gains are achieved through wide reductions in RHHUs, but the production of longer hydrographs almost balances the computational time consumed. It possible to observe a slightly peak gain in computational time at 75% RHHUs removal for Ceizur and Cowansville watersheds and at 37,5% for Toulnustouc. However, some instances on Toulnustouc watershed are more time-consuming than the original models. High representativeness is obtained on HYDROTEL 19, but gains in computational time are too small to consider using this type of surrogate inside an efficient calibration process.

Table 5.7. Average CPU time for one single simulation and CPU time ratio of surrogates having a reduced number of RHHUs for all "Watershed-HYDROTEL 10" combinations

		HYDROTEL - 10 Parameters		
		CEIZUR	**COWANSVILLE**	**TOULNUSTOUC**
Benchmark Models		120.84 sec	26.81 sec	49.43 sec
Surrogates Number of RHHUs	12.5% less	- - -	- - -	- - -
	25% less	- - -	- - -	- - -
	37.5% less	- - -	- - -	- - -
	50% less	91.58 sec unusable	15.61 sec poor	31.72 sec poor
	75% less	67.76 sec poor	9.91 sec poor	22.77 sec poor
	87.5% less	58.92 sec poor	6.37 sec medium	17.72 sec poor
CPU Time Ratio:		-∞, 1.5] = unusable,]8, 15] = superior,	]1.5, 4] = poor,]15, +∞ = ideal	]4, 8] = medium,

Table 5.8. Average CPU time for one single simulation and CPU time ratio of surrogates having a reduced number of RHHUs for all "Watershed-HYDROTEL 19" combinations

		HYDROTEL - 19 Parameters		
		CEIZUR	**COWANSVILLE**	**TOULNUSTOUC**
Benchmark Models		116.18 sec	40.96 sec	89.05 sec
Surrogates Number of RHHUs	12.5% less	105.99 sec unusable	37.53 sec unusable	80.31 sec unusable
	25% less	103.69 sec unusable	34.90 sec unusable	78.68 sec unusable
	37.5% less	103.41 sec unusable	31.26 sec unusable	75.44 sec unusable
	50% less	108.61 sec unusable	31.51 sec unusable	85.54 sec unusable
	75% less	100.27 sec unusable	29.70 sec unusable	112.42 sec unusable
	87.5% less	104.34 sec unusable	34.11 sec unusable	124.74 sec unusable
CPU Time Ratio:		-∞, 1.5] = unusable,	]1.5, 4] = poor,	]4, 8] = medium,
		]8, 15] = superior,	]15, +∞ = ideal	

5.4.4 Final combined surrogate models

Based on the previous computational experiments, the best level of meteorological stations removal is 87.5% and the best calibration time-period is 12 months for both HYDROTEL problems. The best level of RHHUs removal is 87.5% for HYDROTEL 10 versus 75% for the Ceizur watershed and 37.5% for the Cowansville and Toulnustouc watersheds in the case of HYDROTEL 19.

Individually, each type of surrogate model has shown that reduced-fidelity models can be produced with high representativeness, but the CPU time ratios obtained are mostly less than that would be expected if to be used within an efficient calibration process. In order to amplify the CPU time ratio, combinations of the three different types of surrogates is tested on each original "Watershed-HYDROTEL problem" to produce final combined surrogate

models. For all watersheds on HYDROTEL 10, final combined surrogate models are set at 87.5% of meteorological stations removal, a 12 months' calibration time-period and at 87.5% of RHHUs removal. For HYDROTEL 19, the variability on the best level of RHHUs removal depends on modeled watershed. This situation generates two different experiments. The first experiments tested the final combined surrogate models at 87.5% of meteorological stations removal, a 12 months' calibration time-period and at the best level of RHHUs removal depending on modeled watershed (75% for Ceizur and 37.5% for Cowansville and Toulnustouc). The second tested the final combined surrogate models at 87.5% of meteorological stations removal, a 12 months' calibration time-period and at the original number of RHHUs; i.e. that the final combined surrogate models are set without reducing the original number of RHHUs.

Table 5.9. Average CPU time for one single simulation and CPU time ratio of final combined reduced-fidelity models for all "Watershed-HYDROTEL 10" combinations

		HYDROTEL - 10 Parameters		
		CEIZUR	**COWANSVILLE**	**TOULNUSTOUC**
Benchmark Models		120.84 sec	26.81 sec	49.43 sec
Final Combined Surrogate Models	Best Level of Reduced-Fidelity Models	2.69 sec	1.61 sec	1.54 sec
		ideal	ideal	ideal
	Without RHHUs Reduction	- - -	- - -	- - -
CPU Time Ratio:		-∞, 1.5] = unusable,	]1.5, 4] = poor,	]4, 8] = medium,
		]8, 15] = superior,	]15, +∞ = ideal	

Table 5.10. Average CPU time for one single simulation and CPU time ratio of final combined reduced-fidelity models for all "Watershed-HYDROTEL 19" combinations

		HYDROTEL - 19 Parameters		
		CEIZUR	**COWANSVILLE**	**TOULNUSTOUC**
Benchmark Models		116.18 sec	40.96 sec	89.05 sec
Final Combined Surrogate Models	Best Level of Reduced-Fidelity Models	46.64 sec	22.54 sec	51.22 sec
		poor	poor	poor
	Without RHHUs Reduction	47.46 sec	19.79 sec	45.82 sec
		poor	poor	poor
	CPU Time Ratio:	-∞, 1.5] = unusable,	]1.5, 4] = poor,	]4, 8] = medium,
		]8, 15] = superior,	]15, +∞ = ideal	

Tables 5.9 presents that the CPU time ratios obtained for HYDROTEL 10 are all qualified as ideal when the best level of reduced-fidelity models is chosen, meaning that the final combined surrogate models are 15 times less (or more) time-consuming than the original hydrological models. More precisely, the CPU time ratios are 44 for the Ceizur watershed, 16 for the Cowansville watershed and 32 for the Toulnustouc watershed. At the opposite, the CPU time ratios obtained for HYDROTEL 19 in Table 5.10 are qualified as poor for both experiments, where the CPU time ratios are around 2. Table 5.10 also illustrates that it is more interesting, in terms of the CPU time ratios, to set the final combined surrogate models without reducing the original number of RHHUs. As reported previously, the production of the geomorphological hydrographs inside the HYDROTEL 19 routine brings uncertainty and variability in the computational time. Consequently, the experiments with final combined surrogate models set without reducing the original number of RHHUs are selected for HYDROTEL 19.

Figure 5.5 presents the representativeness of the final combined surrogate models for each combination "Watershed-HYDROTEL problem" and the summary of its characteristics regarding each type of surrogate models.

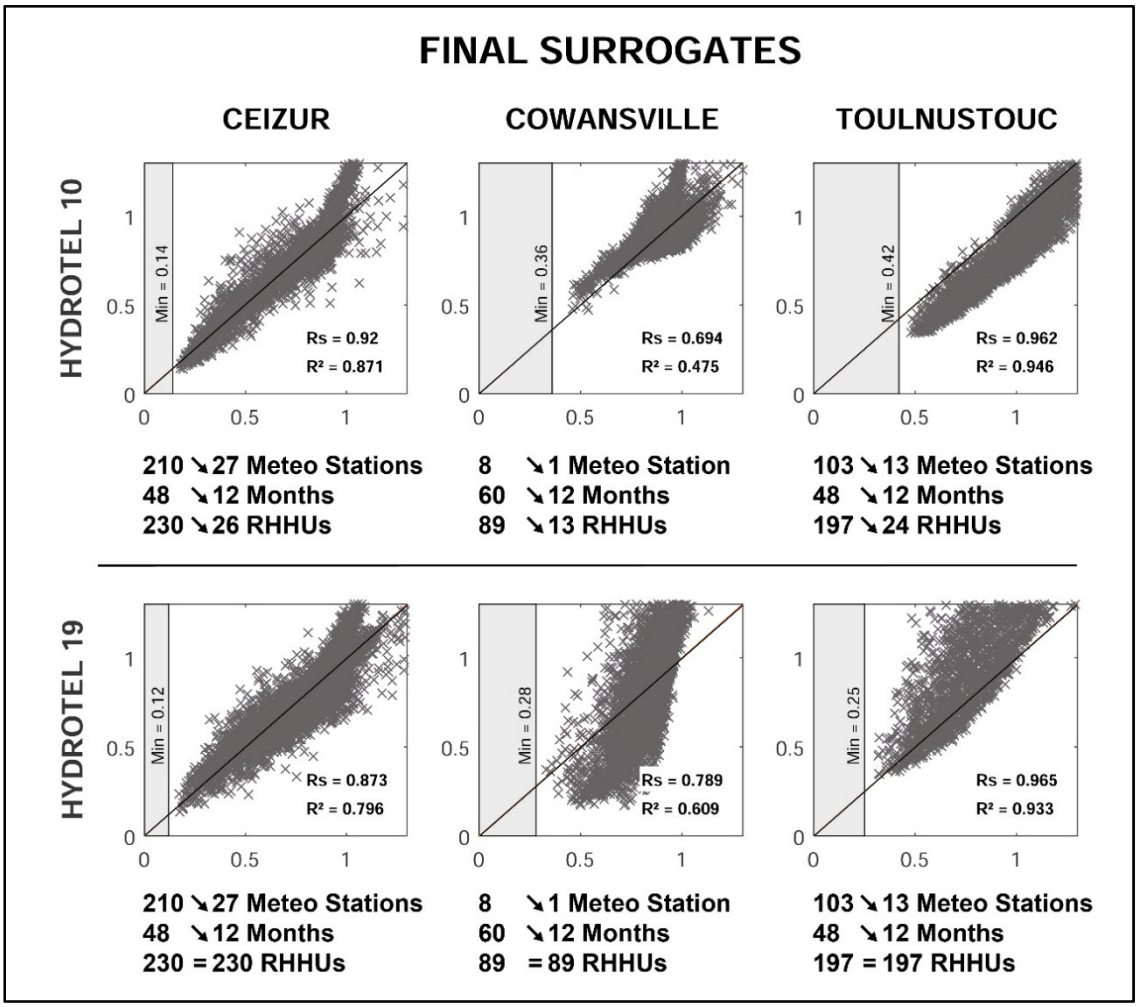


Figure 5.5. $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Ceizur, Cowansville and Toulnustouc watersheds comparing the final selected surrogates with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation. The reduction levels are also indicated below each case presented in the figure

The combination of these three types of surrogates (final combined surrogates) provides higher CPU time ratios in comparison with only one type of surrogate model as presented in Tables 5.3 to 5.10. However, representativeness coefficients of the final combined surrogates illustrated in Figure 5.5 are at a lower-quality and 3 out of 6 surrogate models on this same figure yield representativeness coefficient less than 0.9.

5.5 Discussion

This paper aimed to explore and develop three different types of reduced-fidelity surrogate models and the combination of all them from the original hydrological model HYDROTEL. These three types of surrogates are the reduction of the number of meteorological stations inputting the modelling, the reduction of calibration time-periods and the reduction of the watersheds spatial discretization. The representativeness and the CPU time ratios between original and surrogate models have been compared and evaluated to use it within calibration processes.

The potential to use the final combined surrogate models for HYDROTEL 10 within an efficient calibration process is enviable due to the ideal CPU time ratios obtained. Moreover, the representativeness coefficients are superior or very close to 0.9 for the Ceizur and Toulnustouc watersheds, which should help to exploit properly the surrogate models in calibration. Lower representativeness coefficients are nevertheless obtained on the Cowansville watershed. For the final combined surrogate models of HYDROTEL 19, the CPU time ratios are rather qualified as poor. A priori, the final surrogates developed on this problem do not provide high enough CPU time ratios to implement them within calibration processes, especially since the representativeness coefficients are over 0.9 for the Toulnustouc watershed and under for the other two watersheds.

The second paper of this research will experiment the implementation of the final combined reduced-fidelity models within an efficient calibration process. The optimization framework used for the implementation will be the hybrid optimization approach DDS-MADS developed by Huot et al. (2017). This hybrid method merges the convergence analysis and robust local refinement from the Mesh Adaptive Direct Search (MADS; Audet & Dennis 2006) algorithm with the global exploration capabilities from the heuristic strategies used by the Dynamically Dimensioned Search algorithm (DDS; Tolson & Shoemaker 2007). This two-step based approach will present some advantages to easily exploit efficiently the final reduced-fidelity models.

Three interesting opportunities emerge from the final combined reduced-fidelity models. Firstly, the variability in representativeness coefficients quality allows the opportunity to evaluate if the recommendation from Toal (2015) is also applicable to reduced-fidelity models. The hypothesis that high representativeness coefficients (greater than 0.9) are absolutely necessary to obtain a reduction in computational intensive calibration processes will be addressed in the second paper of this research. Part II will revise this recommendation for reduced-fidelity models.

Secondly, this present paper lists five ranges of CPU time ratio and sets the ideal value of r to over 15. According to these five ranges, the CPU time ratios obtained on HYDROTEL 10 (ideal ratio) and those obtained on HYDROTEL 19 (poor ratios) are at opposite ends of the spectrum. Thus, the importance of an ideal ratio between original and surrogate models inside computationally-intensive calibration processes will also be reviewed in next paper.

Thirdly, the reduced-fidelity models developed in this paper generate an interesting questioning about the original form of the HYDROTEL problems. Final reduced-fidelity models suggest high representativeness of the original models, which means that this reduced-modelling level could also perform well in hydrological modelling. Future works will explore the opportunities of using the reduced-fidelity models as original hydrological models in various hydrological modelling contexts.

5.6 Acknowledgments

The meteorological gridded datasets and daily observed streamflows for studied watersheds have been provided by the Direction de l'Expertise Hydrique (DEH) and Hydro-Québec. Funding for this study has been provided by a PhD scholarship from National Sciences and Engineering Research Council (NSERC) of Canada.

CHAPITRE 6

LOW-COST AND REPRESENTATIVE SURROGATE HYDROLOGICAL MODELS, PART II: USE WITHIN CALIBRATION FRAMEWORKS

Pierre-Luc Huot[1], Annie Poulin[1], Charles Audet[2] et Stéphane Alarie[3]

[1] Département de la Génie de la Construction, École de technologie supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3.
[2] Département de Mathématiques et Génie Industriel, Polytechnique de Montréal et GERAD,
2900 boul. Édouard-Montpetit, Montréal, Québec, Canada H3T 1J4.
[3] Institut de Recherche d'Hydro-Québec,
1800 boul. Lionel Boulet, Varennes, Québec, Canada J3X 1S1.

Article soumis pour publication dans la revue « Environmental Modelling & Software »,
décembre 2018

Abstract

This is a two-part work. In Part I, low-cost and representative reduced-fidelity models of two versions of the HYDROTEL hydrological model are constructed, using three types of surrogate models and their combination. Level of representativeness and CPU time ratios between original and surrogate models are evaluated to construct final reduced-fidelity models with the high potential use within a calibration process. Part II of this study focuses on the use of these models within an existing efficient calibration method: the hybrid DDS-MADS optimization approach. Based on this hybrid approach, this paper proposes a range of calibration frameworks exploiting reduced-fidelity models. The calibration frameworks are assessed and compared with one another and results demonstrate that exploiting reduced-fidelity models within the hybrid DDS-MADS optimization approach decreases the overall computational time while maintaining the quality of the final solutions. The tested frameworks provide a range of tradeoffs between computational time and objective function value, which offers the users the possibility to select the appropriate framework according to their calibration objectives and optimization constraints.

Keywords: Distributed hydrological model, Computationally-intensive simulation model, Optimization algorithm, Efficient calibration, Calibration framework, Hybrid DDS-MADS approach

6.1 Introduction

The objective of this research project is to develop efficient tools to calibrate the parameters of computationally-intensive hydrological models. Two avenues proposed by Razavi et al. (2010) are combined to achieve this: the development of efficient optimization strategies and the use of low-cost surrogate models. The former has been previously studied by the authors. Huot et al. (2017) proposed the hybrid optimization approach DDS-MADS, which combines the "Dynamically Dimensioned Search" algorithm (DDS) of Tolson and Shoemaker (2007) with the "Mesh Adaptive Direct Search" algorithm (MADS) of Audet and Dennis (2006). In comparison with existing optimization algorithms, results show that the hybrid DDS-MADS approach reduces the total number of model evaluations without compromising the quality of the final solution. The use of low-cost and representative surrogate models within efficient calibration frameworks is studied in the present work.

Part I (Chapter 5) of this research studies and analyses the construction of various low-cost and representative surrogate models, from the family of reduced-fidelity physical-based models. The hydrological model used for the exploration of different approaches of reduced-fidelity models is the HYDROTEL model (Fortin et al., 2001a-b), a distributed and computationally-intensive model. Two versions of the HYDROTEL model with 10 and 19 parameters are used, called hereafter HYDROTEL 10 and HYDROTEL 19, respectively. Three watersheds from the province of Québec are modeled with these two HYDROTEL versions, namely Cowansville, Ceizur and Toulnustouc watersheds (see Section 5.3). The six combinations of "Watershed-HYDROTEL problem" define the calibration problems used as computational experiments in this study. Part I demonstrates the potential of using reduced-fidelity models by combining three approaches (further described in Section 6.3.1. Results show significant reduction of computational time and high-levels of representativeness

between final combined reduced-fidelity surrogates and original models. These two characteristics suggest the potential to efficiently use them within hydrological model calibration processes.

Part I was dedicated to the construction of reduced-fidelity models with the hydrological model HYDROTEL. Part II focuses on the use of the reduced-fidelity models within various calibration frameworks. A range of optimization methods exploiting surrogate models is proposed to let users select an appropriate calibration framework depending on their own calibration objectives and optimization constraints.

6.1.1 Literature review

Response surface functions, or adaptive functions (Le Digabel, 2011; Talgorn et al., 2017), are based on the emulation of an ensemble of past evaluated solution points and then applying empirical interpolation techniques to reproduce as accurately as possible the "surface" of the objective function of a simulation model. Within an optimization process, all evaluations of the simulation model may provide a new beneficial contribution to the adaptive function. The quality of the emulation may be improved by adding or removing some points in the sampling set (Booker et al., 1999; Regis & Shoemaker, 2007; Razavi et al., 2012a; Wang et al., 2014). Moreover, response surface surrogates can adapt to various models depending on users' needs. These characteristics allow a wide range of possible interactions between an efficient calibration approach and the response surface surrogates. Razavi et al. (2012a) exposed four different interaction frameworks to efficiently exploit surrogates in order to calibrate hydrological models. In the simplest framework, the response surface surrogates are constructed a priori in the optimization process and are updated with new information generated during the optimization process, or are entirely constructed during this process (Marsden et al., 2004; Razavi et al., 2012a-b).

The literature on response surface surrogates is rich and encompasses a wide range of application fields. Popular response surface surrogates include Gaussian kriging techniques

as developed by Krige in 1951 (Krige, 1951; Lophaven et al., 2002), polynomial expressions (Hussain et al., 2002; Lophaven et al., 2002; Fen et al., 2009), Radial Basis Functions (RBF, Hussain et al., 2002; Mugunthan et al., 2005; Regis & Shoemaker, 2007), Artificial Neural Networks (ANNs, Papadrakakis et al., 1998) and Support Vector Machines (SVMs, Zhang et al., 2009). Some papers present more exhaustive overviews of different types of surrogate models (e.g. Queipo et al., 2005; Forrester & Keane, 2009; Razavi et al., 2012a).

There are several uses of adaptive functions that deal with computationally intensive problems including hydrological or environmental modelling. Booker et al. (1999) propose a surrogate management framework using kriging or polynomial models for optimization processes solving design problems for helicopter rotor blades. Recently, Audet et al. (2018) proposed an algorithmic framework to combine various types of surrogates. Closer to the scope of this study, Regis and Shoemaker (2004, 2007, 2009, 2013) make extensive use of response surface surrogates such as RBFs in hydrological modelling and underground water restoration processes. They improve the final values of the objective functions while reducing the computational time. Khu et al. (2004) demonstrate that the use of ANNs inside the optimization process applied to hydrological models reduces the computational time by up to 40% in comparison with a standard metaheuristics optimization algorithm. Razavi et al. (2012b) published a comparative study of three different adaptive functions; RBFs, ANNs and Kriging, which indicates that the use of adaptive functions may decrease the CPU time compared to existing optimization algorithms. Wang et al. (2014) identified Gaussian kriging as the best adaptive function to significantly reduce the computational time of a conceptual rainfall-runoff model. Finally, some existing algorithms integrate adaptive functions as sequencing strategy in order to provide better optimization efficiency. The MADS algorithm (Audet & Dennis, 2006) is adapted to exploit a quadratic polynomial model to locally represent the parametric space by estimating the objective function values of future candidates and sequencing those that will be evaluated in the optimization process according to the improvement potential (Conn & Le Digabel, 2013). All studies cited above prove that CPU time gains may be obtained by using response surface surrogates within an efficient calibration framework or in sensitivity analysis applications. Many other studies also support

this fact (Marsden et al., 2004; Neelin et al., 2010; Mousavi & Shourian, 2010; Castelleti et al., 2010; Yazdi & Salehi Neyshabouri, 2014; Brunetti et al., 2017; Dan Lu et al., 2017; Willers Moore et al., 2016).

Reduced-fidelity models are a priori design models which cannot be updated and continually improved during the optimization process. They represent a simplification of the original simulation model while maintaining representativeness and reducing the computational time as much as possible (Regis & Shoemaker, 2007; Castelletti et al., 2012; Razavi et al., 2012a, Razavi & Tolson, 2013; Wang et al., 2014). This kind of surrogate is commonly used to pre-evaluate the potential of some optimized parameter sets before running a simulation with computationally-intensive model. Reduced-fidelity models with high level of representativeness may be used during the entire calibration process, or partially. Razavi and Tolson (2013) illustrated the case where early in the optimization process improvements of the objective function value can be achieved by using the surrogate model only, which generally leads to very similar results as when the original model is used. But after a while, no further improvement is obtained with the surrogate model due to a non-perfect representation of the original model. Therefore, an important challenge is to correctly manage the use of the reduced-fidelity surrogate versus the original model during optimization.

Some papers explore reduced-fidelity models in the specific context of hydrological modelling. Razavi and Tolson (2013) demonstrated that a reduced-fidelity hydrological model, characterized by a shorter but carefully selected simulation time-period, used within an efficient calibration framework, results in interesting gains on objective function values relative to a standard optimization algorithm, both limited to the same computational time budget. Haghnegahdar et al. (2015) studied various levels of spatial discretization (modelling scales) on different computational time budgets to evaluate the impact on final values of the objective function, but do not focus on reduced-fidelity models within calibration processes. Other studies proposed strategies for the implementation of reduced-fidelity surrogates within efficient calibration frameworks outside of the hydrological modelling domain (Polak

& Wetter, 2006; Castelleti et al., 2012; Koziel & Leifsson, 2013; Leifsson et al., 2014; Huang et al., 2015).

6.1.2 Paper organization

This paper is organized as follows. Section 6.2 exposes the two benchmark algorithms and the calibration framework used to exploit surrogate models. Section 6.3 presents five calibration frameworks exploiting surrogate models and they are applied to six different "Watershed-HYDROTEL problem" combinations. Performance of calibration frameworks is evaluated in terms of computational time and final objective function values. Discussion and future works conclude this paper.

6.2 Benchmark algorithms and calibration framework

Previous works of Huot et al. (2017) study the evaluation of existing calibration algorithms and develop a more efficient calibration approach for computationally-intensive hydrological problems. Two different calibration algorithms serve as benchmark algorithms in this present study: the DDS algorithm (Tolson & Shoemaker, 2007) and the hybrid DDS-MADS calibration approach (Huot et al., 2017). The hybrid DDS-MADS approach exploits surrogate models developed in this two-part research study. Brief descriptions of these two algorithms are presented below.

6.2.1 The Dynamically Dimensioned Search (DDS) algorithm

DDS is a global search heuristic optimization algorithm that evolves the current best single solution. All DDS internal parameters have default values, except for the allowable number of calls to the simulation model representing the total (maximum) simulation budget. Given a user-specified simulation budget, this algorithm dynamically adjusts the search from global to local by gradually reducing the number of perturbed model parameters (search dimensions of the problem). DDS starts by exploring the entire space of variables then gradually focuses on a series of subspaces, until it reaches a one-dimensional subspace. DDS always exploits

the totality of the user-specified simulation budget (the stopping criteria), i.e. it does not base the optimization search strategies on quality of current best solution. For calibrations of computationally-intensive hydrological models, the user needs to carefully select the simulation budget to control the overall computational time. Studies on hydrological modelling (Razavi et al. 2010, Arsenault et al. 2014, Huang et al., 2014; Huot et al., 2017) report that the global search strategies used by DDS converge more rapidly to good-quality solutions in comparison with some other existing algorithms.

The DDS release 1.2 with updates from February 2015 is used in the present work. DDS is used only as a benchmark algorithm for the calibration experiments, and is named hereafter DDS-Bench.

6.2.2 The hybrid DDS-MADS optimization approach

The hybrid optimization approach DDS-MADS was developed in the context of computationally-intensive calibration of hydrological models. This hybrid optimization approach merges the DDS algorithm (as presented previously) with the MADS algorithm. This two-step hybrid optimization approach (a DDS step and a MADS step) benefits from the advantages of both algorithms: the global exploration capabilities of the DDS algorithm, and the local refinement process and automatic stopping criteria (based on the quality of solutions) from the MADS algorithm.

The MADS algorithm is a direct search method that discretizes the solution space into a grid mesh. It performs an adaptive search on this mesh; i.e. the mesh becomes coarser when the current best solution is improved and is refined when the local search around the current best solution fails to identify a better solution. The MADS algorithm is designed to terminate automatically the calibration as soon as some local optimality conditions are satisfied, guaranteeing that the final solution satisfies on local optimality conditions. Moreover, the MADS algorithm is implemented with many functionalities including a response surface function integrated to the direct search method. This response surface function is a quadratic

model and is used to order the $2N$ (where N is the number of calibrated parameters) trial points on the mesh according to orthogonal directions from the current best solution in a new iteration. The quadratic model is an adaptive function constructed with past-evaluated trial points that influence the local domain around the current best solution point. Best trial point performances in the quadratic model are simulated first to avoid several computationally-intensive simulations. Conn & Le Digabel (2013) demonstrated that the use of quadratic models in the trial points ordering improves the performance of the MADS algorithm significantly by reducing the total number of objective function evaluations. Although this paper does not specifically examine response surface surrogate, the MADS algorithm within the DDS-MADS approach uses this functionality in its optimization process. For more details about the optimization strategies and/or functionalities of the MADS algorithm, please see the relevant papers (Audet & Dennis, 2006; Abramson et al., 2009; Conn & Le Digabel, 2013; Audet et al., 2016).

To adequately merge the two algorithms, Huot et al. (2017) propose five simple transition features: (1) introducing a cache file system, (2) implementing a stagnation parameter to switch from DDS step to MADS step automatically, (3) fine-tuning the initial mesh size of the MADS step, (4) managing the total simulation budget and (5) rounding the objective function value. The cache file system registers all simulated solutions and their respective objective function value for two purposes: to avoid repeating the same simulation by both algorithms and to collect information about parameters sensitivity. When a good-quality solution is reached in the global exploration step (DDS step), local optimization strategies used by DDS may be exhausted and stagnation in the improvement of the objective function value occurs. A stagnation parameter interrupts the DDS step at the appropriate timing to continue the optimization with the MADS step and requires no adjustment. The initial mesh size parameter of the MADS step is set according to a spatial variability analysis from good-quality points registered in the cache file of the DDS step; in other words, the position in the solution space of all good-quality points serve to adjust independently each search dimension of the mesh. The total user-specified simulation budget is first attributed to the DDS step. When the stagnation parameter interrupts the DDS step, the remaining simulation budget is

transferred to the MADS step. However, the MADS step does not necessarily consume all the remaining simulation budget depending on the necessary number of simulations to identify a nearby local optimum. Finally, numerical noise in the objective function surface may cause a continuous adaptation of the mesh, delaying the local convergence process to a local optimum from the MADS step. Rounding the objective function value to an appropriate number of digits (depending on the objectives pursued by the calibration process) attenuates this numerical noise and stabilizes the performance of the algorithm.

These five transition features of the hybrid DDS-MADS approach provide a calibration process that reduces the total number of hydrological model simulations necessary to reach good-quality local optimum. Huot et al. (2017) report that this hybrid approach reduces on average by 70% and 40% the computational time for the HYDROTEL 10 and HYDROTEL 19 problems respectively in comparison with the DDS-Bench algorithm without compromising the quality of final solutions. The ability of this hybrid approach to escape from poor quality-zones in the parametric space has been also demonstrated. Moreover, the DDS-MADS approach does not require any tuning or sensitivity analysis of algorithm parameters which would involve consuming high computational times.

The hybrid DDS-MADS approach merges the DDS release 1.2 with the NOMAD 3.5 implementation of the MADS algorithm (Le Digabel, 2011). DDS-MADS is used as benchmark algorithm for the calibration experiments as well as in the calibration framework which implements the surrogate models. The method will be referred to as DDS-MADS-Bench.

6.3 Use of surrogate models

This section opens with a review of surrogate highlights obtained with the reduced-fidelity models developed in Part I of this study (Chapter 5). The objective of this review is to guide the implementation of potential calibration frameworks. Next, five calibration frameworks exploiting surrogate models are defined and tested. All combinations of “Watershed-

HYDROTEL problem" are used for numerical testing. Comparisons between calibration frameworks and/or benchmark algorithms according to the computational time reductions and quality of final objective function values conclude this section.

6.3.1 Review of reduced-fidelity models highlights

Part I of this study focused on the development of reduced-fidelity models which present interesting characteristics for using them within calibration frameworks: low-cost and high representativeness. The CPU time ratio (ratio between the CPU times of the original and surrogate models) was used to evaluate the potential of computational time reductions of reduced-fidelity models. A terminology regarding the levels of CPU time ratio has been developed and the representativeness was evaluated according to the Spearman's rank correlation coefficient (R_s) and R-Square coefficient (R^2) calculated from the objective function $1 - NSE$ obtained with original and surrogate models. The reader is invited to consult Section 5.2.2 for more details. Three types of reduced-fidelity models and their combination have been evaluated: (1) reduction of the number of "pseudo-meteorological" stations on the territory, (2) the reduction of the calibration time-period and (3) the reduction of the level of spatial discretization by decreasing the number of simulation units, referring to RHHUs. Two main results from Part I are highlighted herein.

First, all three types of reduced-fidelity models provide generally high R_s and R^2 coefficients. However, their individual use is not recommended either because of the poor CPU time ratios that were obtained (refer to Section 5.2.2), or because of the very low representativeness coefficients that were obtained in some instances (although interesting time ratios were observed in some of these poorly representative cases).

Second, the combination of all three types of reduced-fidelity models were evaluated in Part I. The final reduced-fidelity models retained were the combination of all three types of reduced-fidelity models for the HYDROTEL 10 problem. For the HYDROTEL 19 problem, it was the merge of the reduction of the number of meteorological stations and the reduction

of the calibration time-period (no modification to spatial discretization). These final reduced-fidelity models are potentially usable within calibration frameworks since they provide the best compromise between high representativeness coefficients and a decrease in CPU time. For the HYDROTEL 10 problem, the final reduced-fidelity models provided representativeness coefficients over or close to the 0.9 value for the Ceizur and Toulnustouc watersheds as recommended by Toal (2015). The reduced-fidelity models for the Cowansville watershed appeared to be less representative with R_s of 0.694 and R^2 of 0.475. However, all three modelled watersheds with HYDROTEL 10 provided ideal CPU time ratios, being 16 to 44 times less computationally-intensive than the original models. For the HYDROTEL 19 problem, similar results are obtained for the levels of representativeness (over or close to 0.9 for Ceizur and Toulnustouc watersheds with smaller values for the Cowansville watershed). Nonetheless, the final reduced-fidelity models developed for HYDROTEL 19 are all approximately 2 times less computationally-intensive; i.e. resulting as poor CPU time ratios.

6.3.2 Calibration framework experiments

Five different calibration frameworks exploiting final reduced-fidelity models and response surface functions within the hybrid DDS-MADS optimization approach are tested. As mentioned in Section 6.2.2, the DDS-MADS approach is a two-step method; DDS and MADS steps focusing on global and local searches respectively. This approach provides many possibilities to manage surrogate models for calibration problems due to this clear division of global/local searches and intrinsic characteristics of each algorithm. Tables 6.1 and 6.2 summarize the three (**A**, **B** and **C**) and five (**A** to **E**) calibration frameworks tested on the original HYDROTEL 10 and 19 problems, respectively. A description of each calibration framework is presented next.

Calibration framework A

The calibration framework **A** allies the global search strategy of the DDS step performing on the final reduced-fidelity models and the local search strategy of the MADS step performing

on the original models. The objective is to save computational time during the global search, using low-cost and highly representative reduced-fidelity models, and then to identify optimal solutions on the original models. Given that DDS step performs on reduced-fidelity models and the MADS step on original models; the initial mesh size adjustment cannot be used in calibration framework **A**. As mentioned in Section 6.2.2, the MADS algorithm use the quadratic model to order the $2N$ trial points on the mesh. Considering the local search role of the MADS algorithm, the use of quadratic models would result in reducing the total number of computationally-intensive simulations. In the worst situation, all trials points are simulated as would have been the case if the quadratic models had not been exploited. The calibration framework **A** is performed on the original HYDROTEL 10 and 19 problems and is summarized in Tables 6.1 and 6.2 under the column (A).

Table 6.1. Three calibration frameworks exploiting reduced-fidelity models and response surface functions tested on the original HYDROTEL 10 problems

		Experiments on HYDROTEL 10		
Calibration Frameworks		**(A)**	**(B)**	**(C)**
Hybrid DDS-MADS Approach	**DDS Step**	**HYDROTEL 10** Reduced-Fidelity Models	**HYDROTEL 10** Reduced-Fidelity Models	**HYDROTEL 10** Reduced-Fidelity Models
	Initial Mesh Size Adjustement	**NO (default value)**	**YES**	**YES**
	MADS Step [with Quadratic Models]	**HYDROTEL 10** Original Models	**HYDROTEL 10** Reduced-Fidelity Models	**HYDROTEL 10** Reduced-Fidelity Models
MADS Algorithm [with Quadratic Models]				**HYDROTEL 10** Original Models

Calibration framework B

Calibration framework **B** aims to fully exploit the low-cost and the high representativeness of the final reduced-fidelity models (especially for HYDROTEL 10). Thus, the calibration framework **B** is the hybrid DDS-MADS optimization approach performing strictly on final reduced-fidelity models without considering the original models in the calibration process. The initial mesh size adjustment and the quadratic models within the MADS step are used. At the end of the optimization process, only one last simulation is evaluated on the original models to obtain the final objective function value. Moreover, results of the calibration framework **B** may provide more information about the objective function gap between local optima in the final reduced-fidelity models and their counterpart in the original models. The calibration framework **B** is performed on both HYDROTEL versions and is summarized in Tables 6.1 and 6.2.

Calibration framework C

The first step of the calibration framework **C** is the calibration framework **B**; i.e. the hybrid DDS-MADS approach is performed strictly on the final reduced-fidelity models. The solution obtained from this first step is then used as a starting point for the second step, which launches a local search with the MADS algorithm (using quadratic models for trial points ordering) on the original models. The objective is to exploit the calibration framework **B** as the global search strategy and then use the local search features of the MADS algorithm to fill the representativeness gap between local optimal solutions in the final reduced-fidelity models and their counterpart in the original models. If this representativeness gap is low, the required CPU time to compensate for it with the launch of the MADS algorithm is expected to be low. Conversely, if the representativeness gap is high, the local search is expected to be time-consuming and/or to lead to low-quality solutions. The calibration framework **C** is performed on the HYDROTEL 10 and 19 versions and is summarized in Tables 6.1 and 6.2.

Calibration framework D

The calibration framework **D** is performed only on the HYDROTEL 19 version and is summarized in Table 6.2. This framework exploits a fourth type of reduced-fidelity model;

i.e. a reduction of the number of calibration parameters. As mentioned in Section 6.3.1, the ideal CPU times ratios obtained by final reduced-fidelity models on HYDROTEL 10 problems are much more interesting than the poor CPU time ratios obtained on HYDROTEL 19 problems. Recall that this important gap in terms of CPU time ratios is mainly caused by the production of computationally-intensive geomorphological hydrogaphs. These hydrographs are produced on every new combination of 2 calibration parameters included in the overland routing of the HYDROTEL 19 version (see Section 5.2.1). The calibration framework **D** is designed to take advantage of this gap in terms of CPU time ratios using the final reduced-fidelity models obtained on HYDROTEL 10 rather than those on HYDROTEL 19. The first step of the calibration framework **D** uses the calibration framework **B** as global search strategy, but performs on the HYDROTEL 10 final reduced-fidelity models (instead of HYDROTEL 19 final reduced-fidelity models). The initial mesh size adjustment and the quadratic models are both used in this DDS-MADS first step. The second step of the framework **D** uses the solution from the first step as a starting point, and then launches a local search with the MADS algorithm (using quadratic models for trial points ordering) on HYDROTEL 19.

Table 6.2. Five calibration frameworks exploiting reduced-fidelity models and response surface functions tested on the original HYDROTEL 19 problems

	Calibration Frameworks	Experiments on HYDROTEL 19				
		(A)	(B)	(C)	(D)	(E)
Hybrid DDS-MADS Approach	**DDS Step**	**HYDROTEL 19** Reduced-Fidelity Models	**HYDROTEL 19** Reduced-Fidelity Models	**HYDROTEL 19** Reduced-Fidelity Models	**HYDROTEL 10** Reduced-Fidelity Models	**HYDROTEL 17** Reduced-Fidelity Models
	Initial Mesh Size Adjustement	**NO (default value)**	**YES**	**YES**	**YES**	**YES**
	MADS Step [with Quadratic Models]	**HYDROTEL 19** Original Models	**HYDROTEL 19** Reduced-Fidelity Models	**HYDROTEL 19** Reduced-Fidelity Models	**HYDROTEL 10** Reduced-Fidelity Models	**HYDROTEL 17** Reduced-Fidelity Models
	MADS Algorithm [with Quadratic Models]			**HYDROTEL 19** Original Models	**HYDROTEL 19** Original Models	**HYDROTEL 19** Original Models

Calibration framework E

The calibration framework **E** is based on same principle as previous framework **D** and is performed only using the HYDROTEL 19 version. As mentioned in the calibration framework **D**, two calibration parameters are mainly responsible of the poor CPU time ratios. To isolate these two time-consuming calibration parameters, the final reduced-fidelity models on the HYDROTEL 19 have been modified to create a 17 calibration parameters version, named hereafter HYDROTEL 17. No experiment has been completed to evaluate the level of representativeness and the CPU time ratios of the reduced-fidelity HYDROTEL 17 models, but this version is expected to yield good performance on CPU time ratios and representativeness. The calibration framework **E** is identical to the calibration framework **D** except that the final reduced-fidelity models on HYDROTEL 10 is replaced by the HYDROTEL 17 version. As for the framework **D**, the main objective is to exploit the final reduced-fidelity models on HYDROTEL 17 which are less computationally-intensive (similar to the HYDROTEL 10 version), but more representative than the HYDROTEL 10 version. Table 6.2 summarizes the calibration framework **E**.

6.3.3 Results from the calibration frameworks on HYDROTEL 10

Figure 6.1 presents the final $1 - NSE$ results from 32 calibration trials using calibration frameworks **A** to **C** for the three modeled watersheds with HYDROTEL 10. Calibration framework results are also compared with the two benchmark optimization algorithms DDS-MADS-Bench and DDS-Bench as presented in Section 6.2. The DDS-MADS-Bench algorithm is used for the computational time comparisons while the DDS-Bench is rather for the final objective function value comparisons. Table 6.3 presents minima, maxima and averages of computational times (hours) related to Figure 6.1 for each of calibration frameworks **A** to **C**. The computational time reductions (%), according to the computational time of DDS-MADS-Bench, are also presented in the neighbouring columns. Results for the final objective function values (Figure 6.1) and related computational times (Table 6.3) are discussed in parallel for each calibration framework.

Calibration framework A

Based on medians and interquartile ranges, Figure 6.1 shows that calibration framework **A** produces slightly better objective function values than the DDS-MADS-Bench (-0.005), and slightly worst values than those of DDS-Bench ($+0.005$) for all three modelled watersheds. Table 6.3 suggests that on average, framework **A** decreases the computational times compared to those of DDS-MADS-Bench.; i.e. ranging from 2% for the Cowansville watershed to 14% for the Ceizur watershed reductions in the computational times. These computational time reductions by the calibration framework **A** are probably associated to the global search strategy of DDS step performing on final reduced-fidelity models. Nonetheless, calibration framework **A** provides similar performances to both benchmark algorithms regarding the final $1 - NSE$ values, with small computational time reductions.

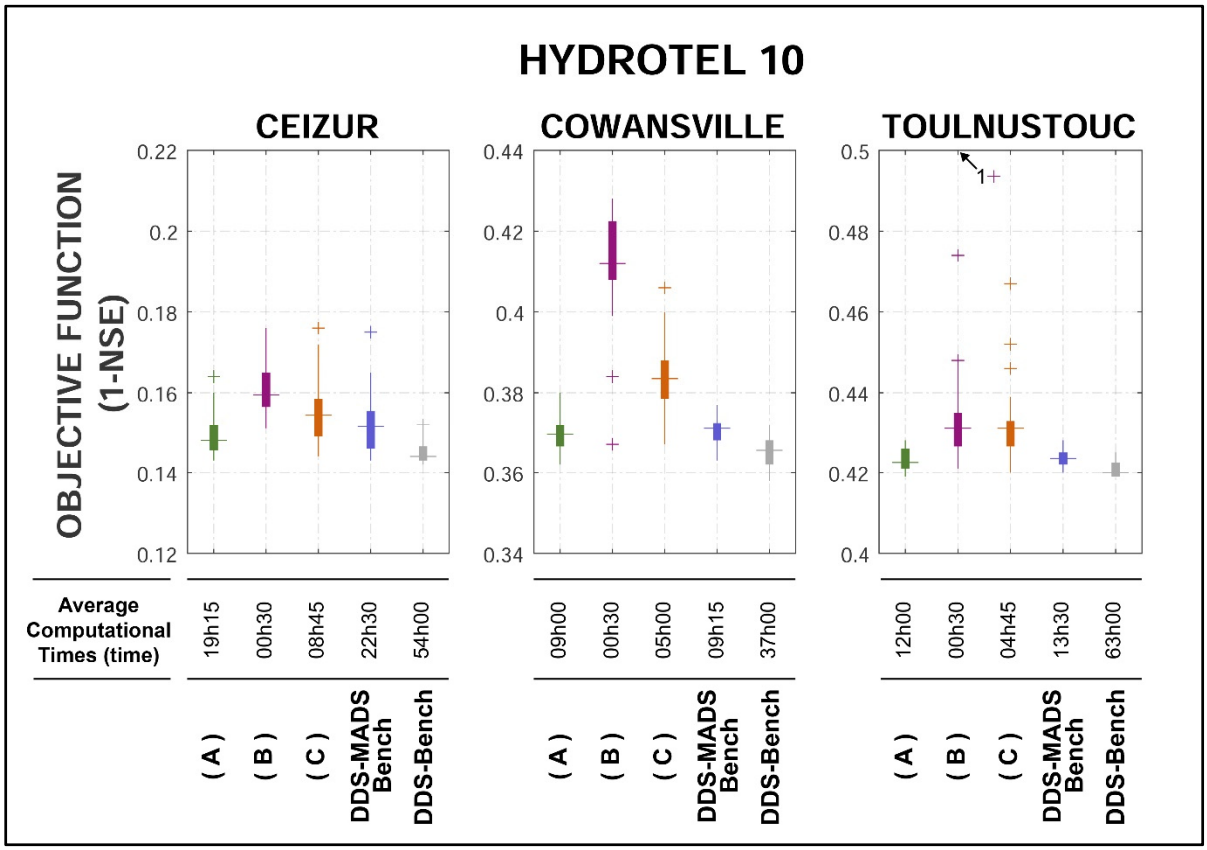


Figure 6.1. Final $1 - NSE$ results from calibration frameworks **A**, **B** and **C** (boxplots), and for DDS-MADS-Bench and DDS-Bench. Average computational times are listed for all optimization methods. Calibrations were performed on all three modelled watersheds: Ceizur, Cowansville and Toulnustouc with HYDROTEL 10

Calibration framework B

Calibration framework **B** offers the worst performance of the three frameworks **A** to **C** regarding the final objective functions values, but the difference between the medians of framework **B** and DDS-Bench is less than $+0.02$ for the Ceizur and Toulnustouc watersheds and $+0.04$ for the Cowansville watershed (Figure 6.1). Considering that this framework only performed on reduced-fidelity models, it may be confirmed that local optimal solutions of reduced-fidelity models are still good-quality solutions (even if not optimal) in the original models. Although the calibration framework **B** does not present the best final objective function values, Table 6.3 shows a significant decrease in computational times for all three watersheds. Approximatively 30 minutes are needed to complete all calibration trials with the framework **B** on any modelled watershed, representing reductions of computational time of 94% to 98% in comparison with DDS-MADS-Bench. As the calibration framework **B** only performs on reduced-fidelity models (except for the very last evaluation), these high computational time reductions are explained by the ideal CPU time ratios for all three modelled watersheds (see Section 5.4.4). It is important to keep in mind that framework **B** does not provide a local optimal solution for the original model.

Table 6.3. Minima, maxima and averages of computational times (hours) and computational time reductions (%) for each calibration framework **A**, **B** and **C** in comparison with DDS-MADS-Bench and DDS-Bench. Ceizur, Cowansville and Toulnustouc watersheds are modelled on HYDROTEL 10

		HYDROTEL 10							
	Calibration Frameworks	**(A)**		**(B)**		**(C)**		**DDS-MADS Bench ***	**DDS-Bench**
Computational Times (Hours) / **Computational Times Reductions (%)**	**Ceizur**								
	Minimum	12.11	- 19	0.34	- 98	4.56	- 69	14.88	---
	Average	**19.27**	**- 14**	**0.52**	**- 98**	**8.79**	**- 61**	**22.50**	**54.00**
	Maximum	33.13	- 8	0.90	- 98	13.61	- 62	36.20	---
	Cowansville								
	Minimum	5.16	- 5	0.30	- 95	1.67	- 69	5.46	---
	Average	**9.09**	**- 2**	**0.52**	**- 94**	**5.07**	**- 45**	**9.23**	**37.00**
	Maximum	14.18	+ 1	0.81	- 94	7.09	- 50	14.10	---
	Toulnustouc								
	Minimum	7.51	- 26	0.29	- 97	2.58	- 75	10.18	---
	Average	**11.96**	**- 11**	**0.48**	**- 96**	**4.80**	**- 64**	**13.44**	**63.00**
	Maximum	16.35	- 8	1.04	- 94	6.68	- 63	17.83	---

* Computational time reductions are calculated according to the computational times of DDS-MADS Bench

Calibration framework C

Figure 6.1 shows that the medians of framework **C** are at a distance of $+0.01$ of the DDS-Bench $1 - NSE$ values for the Ceizur and Toulnustouc watersheds and at a distance of $+0.02$ for the Cowansville watershed. These $1 - NSE$ gaps are greater with calibration framework **A**, and interquartile ranges of framework **C** are also wider. These differences in the results between frameworks A and C still remain very small from a hydrological perspective. As described, calibration framework **C** uses the framework **B** in a first step and then, the MADS algorithm performs a local descent to optimal solutions on the original models. Gaps between final results from frameworks **B** and **C** illustrate the additional labor of the MADS algorithm to reduce the $1 - NSE$ gaps between framework **B** and benchmark algorithms at the expense of higher computational times. Table 6.3 shows that framework **C** is 45% (Cowansville watershed) to 64% (Toulnustouc watershed) less time-consuming in comparison with the DDS-MADS-Bench. Computational times associated to the MADS

algorithm are illustrated in Table 6.3 by the difference between the values of framework **B** and framework **C**. Overall, calibration framework **C** offers an intermediate compromise between frameworks **A** and **B**; i.e. higher computational time reductions than framework **A** but slightly lower performance on final $1 - NSE$ values, and lower computational time reductions than framework **B** but better performance in terms of objective function values.

6.3.4 Results from the calibration frameworks on HYDROTEL 19

Calibration results of the frameworks **A** to **E** for the three modelled watersheds with HYDROTEL 19 are presented in Figure 6.2. Comparison with the DDS-MADS-Bench for the computational times and with the DDS-Bench for the final objective function values are discussed in this section. Boxplots in Figure 6.2 present the results from 32 calibration trials for each combination of "Watershed-Framework" for the HYDROTEL 19 version. Computational times and percentage of computational time reductions related to Figure 6.2 for each of the frameworks **A** to **E** are presented in Table 6.4.

Calibration framework A

Figure 6.2 shows that calibration framework **A** performs differently on each modelled watershed. For the Ceizur watershed, the median and interquartile range are similar to those from DDS-MADS-Bench and slightly bigger than those from DDS-Bench ($+0.01$). A gap of $+0.02$ of $1 - NSE$ values between medians of framework **A** and DDS-Bench is shown for the Cowansville watershed, along with similar interquartile range magnitudes. For the Toulnustouc watershed, the interquartile range of framework **A** is much wider than that of DDS-Bench and the median is at a distance of $+0.04$ of $1 - NSE$ value from the DDS-Bench median. Table 6.4 shows that framework **A** is slightly less time-consuming than the DDS-MADS-Bench with averages of 3% (Cowansville watershed) to 9% (Ceizur watershed) of computational time reductions. These reductions are lower than those obtained for the same framework on HYDROTEL 10 (Table 6.1). This is probably due to poor CPU time ratios obtained by the final reduced-fidelity models on HYDROTEL 19 in comparison with the ideal CPU time ratios on HYDROTEL 10 (as shown in Section 5.4.4).

Calibration framework B

Calibration framework **B** presents a decrease in performance in terms of final objective function values in comparison with all other frameworks (Figure 6.2). For all three modelled watersheds, interquartile ranges are the widest of all calibration frameworks and medians are different by at least $+0.05$ of $1 - NSE$ value (Cowansville watershed) from the DDS-Bench median value. These final $1 - NSE$ values are less interesting that those obtained using HYDROTEL 10 for calibration framework **B**. This is probably due to a lower level of representativeness obtained by reduced-fidelity models for HYDROTEL 19. Yet computational times related to framework **B** are very favorable and provide on average 54% to 58% reductions for all modelled watersheds (Table 6.4), although they are smaller than those obtained on HYDROTEL 10 (Table 6.3). This difference may be explained by the poor CPU times ratios (around 2) for all reduced-fidelity models on HYDROTEL 19. Despite the wider gaps of $1 - NSE$ in comparison with other calibration frameworks and benchmark algorithms, framework **B** is still the less time-consuming framework for HYDROTEL 19.

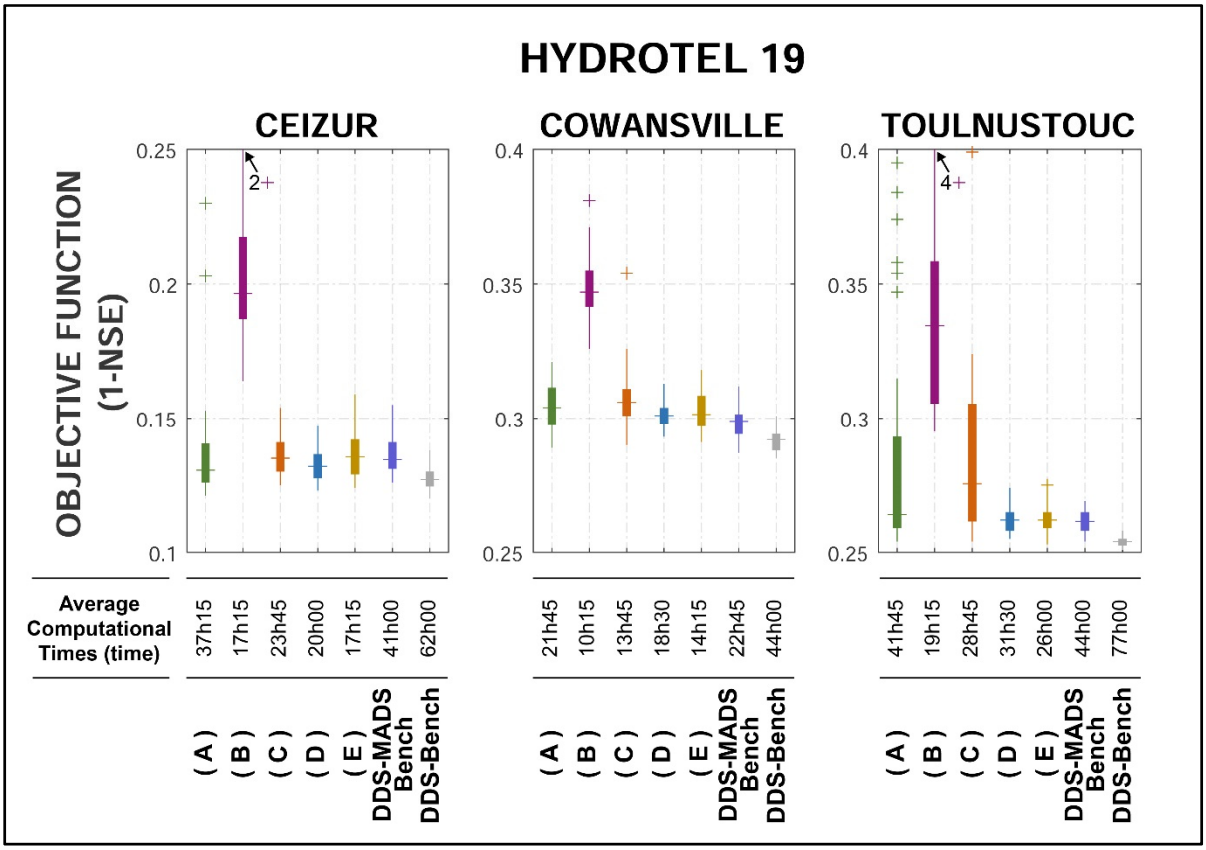


Figure 6.2. Final $1 - NSE$ results from calibration frameworks **A** to **E** (boxplots), and for the DDS-MADS-Bench and DDS-Bench. Average computational times are listed for all optimization methods. Calibrations were performed on all three modelled watersheds: Ceizur, Cowansville and Toulnustouc with HYDROTEL 19

Calibration framework C

Calibration framework **C** presents an uneven final performance from one watershed to the other regarding the final objective function values. For the Ceizur watershed, Figure 6.2 shows similar median and interquartile range between framework **C** and DDS-MADS-Bench while a small difference ($+0.01$ of $1 - NSE$ value) is observed between medians for the Cowansville watershed. The boxplot for the Toulnustouc watershed shows a much wider interquartile range. Moreover, a gap of $+0.08$ of $1 - NSE$ value is obtained between the medians of calibration framework **C** and DDS-Bench for this watershed. Launching the MADS algorithm on the original models (as additional step with respect to framework B) work well for the Ceizur and Cowansville watersheds and results in smaller interquartile ranges and larger medians of the objective function values (first step of framework **C**), but

not for the Toulnustouc watershed. It seems that this last watershed poses more difficulties to the local strategies of the MADS algorithm, which ends up being trapped in lower-quality zones of the parametric space. Table 6.4 reports that average computational time reductions are however fairly homogenous, ranging from 35% for Toulnustouc watershed to 42% for the Ceizur watershed. Despite these interesting computational time reductions, calibration framework **C** provides an ambiguous performance in terms of the objective function values.

Table 6.4. Minima, maxima and averages of computational times (hours) and computational time reductions (%) for the calibration frameworks **A** to **E** in comparison with DDS-MADS-Bench and DDS-Bench. Ceizur, Cowansville and Toulnustouc watersheds are modelled on HYDROTEL 19

Computational Times (Hours) / Computational Times Reductions (%)	HYDROTEL 19											
Calibration Frameworks	**(A)**		**(B)**		**(C)**		**(D)**		**(E)**		**DDS-MADS Bench ***	**DDS-Bench**
Ceizur												
Minimum	23.77	- 4	7.91	- 68	17.40	- 30	13.31	- 46	10.85	- 56	24.76	---
Average	**37.33**	**- 9**	**17.27**	**- 58**	**23.69**	**- 42**	**20.05**	**- 51**	**17.24**	**- 58**	**40.88**	**62.00**
Maximum	49.22	- 28	29.10	- 58	35.46	- 48	26.47	- 61	25.70	- 63	68.70	---
Cowansville												
Minimum	11.73	- 23	3.29	- 78	8.04	- 47	9.93	- 35	7.45	- 51	15.29	---
Average	**21.86**	**- 3**	**10.37**	**- 54**	**13.68**	**- 40**	**18.53**	**- 18**	**14.13**	**- 38**	**22.64**	**44.00**
Maximum	39.79	- 9	22.25	- 49	21.53	- 51	27.73	- 36	22.46	- 48	43.58	---
Toulnustouc												
Minimum	27.64	+ 46	6.52	- 66	14.09	- 26	22.66	+ 19	12.49	+ 34	18.97	---
Average	**41.70**	**- 5**	**19.15**	**- 56**	**28.71**	**- 35**	**31.44**	**- 28**	**26.10**	**- 41**	**43.92**	**77.00**
Maximum	66.81	- 13	33.85	- 56	47.22	- 38	44.27	- 42	35.71	- 53	76.50	---

* Computational time reductions are calculated according to the computational times of DDS-MADS Bench

Calibration framework D

Figure 6.2 shows that calibration framework **D** is slightly less performant than DDS-Bench with objective function values of $+0.01$ of $1 - NSE$, but leads to similar results as those obtained with DDS-MADS-Bench for all three modelled watersheds. Given that reduced-fidelity models on HYDROTEL 10 are used instead of HYDROTEL 19, interquartile ranges and medians are very close to DDS-MADS-Bench results performing strictly on original HYDROTEL models. This performance indicates that localization of good-quality zones in the parametric space is probably similar in both HYDROTEL problems for a same watershed. Table 6.4 shows a reduction in computation times of 18% for the Cowansville watershed and 51% for the Ceizur watershed in comparison with the average computational time of DDS-MADS-Bench. As results suggest, the ideal CPU time ratios obtained with the reduced-fidelity models on HYDROTEL 10 allow for a larger decrease in computational time without sacrificing the final $1 - NSE$ values. Thus, the MADS algorithm (second step of framework **D**) is able to fill the representativeness gaps between the reduced-fidelity models on HYDROTEL 10 and the original models on HYDROTEL 19.

Calibration framework E

Calibration framework **E** performs similarly to calibration framework **D**, as seen in Figure 6.2. Boxplots of final $1 - NSE$ values present similar interquartile ranges and median values, but are positioned slightly higher, in comparison with DDS-MADS-Bench. DDS-Bench results remain marginally better. As in the case of framework **D**, the MADS algorithm is able to fill the representativeness gaps between the reduced-fidelity HYDROTEL 17 models and the original HYDROTEL 19 models. Calibration framework **E** stands out from framework **D** in the computational time reductions. Table 6.4 presents average computational time decreases ranging from 38% for the Cowansville watershed to 58% for the Ceizur watershed. Moreover, it is observed that the average computational time from framework **E** for the Ceizur watershed (17.24 hours) is equivalent to that from the framework **B** (17.27 hours), which strictly uses the reduced-fidelity HYDROTEL 19 models (except for the last simulation which is conducted on original HYDROTEL 19 models).

6.4 Discussion and conclusion

All previous results demonstrate that surrogate models provide an important improvement of the computational times involved in the calibration of computationally-intensive hydrological models while still yielding final good-quality solutions (objective function values). Huot et al. (2017) have shown the efficiency of the hybrid DDS-MADS optimization approach in reducing the computational times of hydrological model calibration. However, the use of low-cost and representative surrogate models within the hybrid DDS-MADS approach contributes to even greater reductions of computational times. Both approaches therefore act synergistically in dealing with computationally-intensive optimization problems.

On all HYDROTEL problems, the tested frameworks offer a range of alternatives for the calibration of computationally-intensive hydrological models. Depending on the user's calibration objectives and optimization constraints, the three calibration frameworks (**A** to **C**) tested on HYDROTEL 10 problems offer different trade-offs between CPU time and quality of final solutions. If the calibration objectives are focused on the final objective function value, framework **A** is more appropriate with, in addition, slight computational time reductions. Conversely, calibration framework **B** is the dominant method according to the computational time reduction, but a slight decrease in final solutions quality must be accepted. Calibration framework **C** is in between frameworks **A** and **B** with a more balanced compromise.

In a perspective to decrease as much as possible the computational times, calibration framework **B** remains the less time-consuming framework on average for all three watersheds when tested on HYDROTEL 19 problems. However, computational time reductions are not as high as those obtained for HYDROTEL 10 problems, and final objective function values are much more degraded. Calibration frameworks **D** and **E** represent both the best compromise between final objective function value and computational time reduction. Framework **D** presents slightly better performance in terms of the final objective function values than framework **E**, but this latter framework offers a higher

computational time reduction. One or the other may be chosen depending on user' calibration objectives and optimization constraints. The main strategy behind these frameworks is to use the HYDROTEL 10 and HYDROTEL 17 versions as reduced-fidelity models which both exclude the 2 particularly time-consuming calibration parameters responsible of the production of geomorphological hydrographs. With these 2 isolated calibration parameters, reduced-fidelity models used in these two frameworks combine the advantage of high CPU time ratios and high representativeness. These performances of frameworks **D** and **E** in comparison with frameworks **A** to **C** raise the importance of using reduced-fidelity models with CPU time ratios qualified as ideal (over a factor of 15). Moreover, for all calibration frameworks tested on HYDROTEL 10 and 19 problems, all ideal CPU time ratios lead to significant decreases in the computational times. Conversely, poor CPU time ratios obtained with the reduced-fidelity HYDROTEL 19 models are not interesting enough to justify their implementation even though they can provide high representativeness. This paper proposes, as in Part I, a terminology to qualify the CPU time ratios between reduced-fidelity and original models but further work could study more precisely how different levels of CPU time ratio can impact the computational time reductions.

This study questions the requirement of a 0.9 or higher representativeness coefficients. Toal (2015) suggests to use only surrogate models that have representativeness coefficients greater than 0.9, but this recommendation was specific to response surface surrogates, not to reduced-fidelity models. In this present paper, few final reduced-fidelity models developed in Part I reached this 0.9 limit, but calibration frameworks were able to generate computational time reductions and good-quality final solutions. Further work could focus on the required level of representativeness according to the reduced-fidelity models.

The HYDROTEL model shows some particularities in its internal structure which lead to the construction of the final reduced-fidelity models used in this paper. Nevertheless, other computationally-intensive hydrological models present different challenges to achieve the construction of low-cost and representative reduced-fidelity models. They could be studied in the same perspective to exploit them within efficient calibration frameworks to decrease the

computational times. Moreover, a wider range of modelled watersheds or objective functions bring also the opportunity to extend the conclusions of this study.

6.5 Acknowledgments

The authors would like to sincerely thank provider of the sources codes of the DDS algorithm, Bryan A. Tolson, and the precious help from Christophe Tribes with the NOMAD software.

The meteorological gridded datasets and daily observed streamflows for studied watersheds have been provided by the *Direction de l'Expertise Hydrique* (DEH) and *Hydro-Québec*. Funding for this study has been provided by a PhD scholarship from National Sciences and Engineering Research Council (NSERC) of Canada.

CHAPITRE 7

FONCTIONS DE SURFACE

Ce chapitre présente l'étude des fonctions de surface en lien avec le deuxième objectif énoncé au Chapitre 3. Cette section expose la capacité des fonctions polynomiales et des modèles de krigeage à représenter la surface de réponse des modèles hydrologiques originaux selon les trois bassins versants Ceizur, Cowansville et Toulsnustouc.

7.1 Formulation des fonctions de surface

Comme mentionné dans la revue de la littérature, de nombreux articles exploitent des fonctions de surface existantes au sein de processus d'optimisation en raison de l'amélioration qu'elles peuvent y apporter et de leur grande polyvalence d'adaptation à presque tous les problèmes d'optimisation. Dans ce chapitre, deux fonctions de surface sont explorées : (1) les fonctions polynomiales (Hussain et al. 2002, Lophaven et al. 2002, Fen et al. 2009) et (2) les modèles de krigeage (Krige 1951, Lophaven et al. 2002). Les fonctions polynomiales sont une technique de représentation inexacte ; c'est-à-dire que les fonctions polynomiales sont une approximation de la valeur de la fonction objectif réelle. Les fonctions polynomiales tentent d'adapter leur surface de réponse à l'ensemble des points évalués antérieurement. Cela occasionne donc un biais systématique entre la valeur réelle de l'ensemble points évalués et la fonction polynomiale. À l'inverse, les modèles de krigeage sont plutôt une technique d'interpolation exacte, ce qui signifie qu'aucun biais n'existe entre l'ensemble des points évalués et le modèle de krigeage (Razavi et al. 2012a). Autrement dit, si l'on évalue un point au sein du modèle de krigeage qui a servi à créer ce modèle, la valeur réelle de la fonction objectif sera obtenue.

L'objectif de cette section est d'évaluer la représentativité de ces deux types de fonctions de surface sur les trois bassins versants Ceizur, Cowansville et Toulnustouc modélisés sur HYDROTEL 10 et HYDROTEL 19 en vue de les utiliser dans des processus d'optimisation. La représentativité obtenue par ces fonctions de surface est étudiée selon la variabilité de deux facteurs : (1) la taille d'échantillonnage (nombre de points évalués) qui servira à la

confection de la fonction de surface, notée S_k, où $k = 25, 50, 100$ et 200 et représente le nombre de points évalués (points échantillonnés par un « Latin Hypercube Search » (LHS)) et (2) le taille du domaine modélisé D_p, où $p = 10, 25, 50$ et 100 et représente le pourcentage par dimension du problème modélisé (paramètres du modèle hydrologique) sur l'écart entre les bornes supérieure et inférieure du paramètre constituant le sous-domaine modélisé. D_p est exprimé comme suit :

$$D_{10} = \left(0{,}10 * (Max_i - Min_i)\right)^N$$

$$D_{25} = \left(0{,}25 * (Max_i - Min_i)\right)^N$$

$$D_{50} = \left(0{,}50 * (Max_i - Min_i)\right)^N$$

$$Domaine\ Global \ \rightarrow \ D_{100} = (Max_i - Min_i)^N$$

où $i = 1,2,3, \dots, N$ est le nombre de dimensions du problème modélisé. Pour toutes les expériences réalisées, les fonctions de surface sont a priori construites à partir d'un premier LHS uniformément distribué de taille S_k et la représentativité est évaluée à partir d'un deuxième LHS de même taille et toujours sur le même sous-domaine D_p.

Les deux fonctions de surface sont utilisées et mises en place à l'aide du « Matlab Kriging Toolbox » version 2.0, aussi appelé le « Design and Analysis of Computer Experiments (DACE) Toolbox », rendu disponible par Lophaven et al. (2002). DACE est un outil de représentation multi-modèles construit à partir d'échantillon de points évaluées provenant de problèmes d'optimisation. Cet outil peut produire des fonctions polynomiales selon trois niveaux de régression : linéaire, quadratique et cubique. Étant donné les tailles de S_k utilisées selon le nombre de dimensions des problèmes hydrologiques originaux, il est possible d'utiliser seulement les niveaux de régression linéaire (dans tous les cas) et quadratique (dans certains cas). Pour évaluer un jeu de paramètres x constitué de N paramètres à modéliser dans un sous-domaine D_p, la formulation de $P(x)$ est exprimée comme suit :

$$P(x) = \left\{ \begin{array}{l|c} \sum_{i=1}^{N} a_i * x_i + b_i \quad (linéaire) & si\ S_k \leq N^2 \\ \sum_{i=1}^{N} \sum_{j=i+1}^{N} a_i * {x_i}^2 + b_i * x_i * x_j + c_i * x_i + d_i \quad (quadratique) & sinon; \\ \quad tel\ que \quad x \in D_p \quad et \quad a, b, c, d \in \mathbb{R} & \end{array} \right\}$$

Les modèles de krigeage peuvent être produits en intégrant l'un des trois niveaux de régression : constant, linéaire et quadratique, à l'un des sept modèles de corrélation (semi-variogrammes) suivants : gaussien, exponentiel, exponentiel généralisé, cubique, splines, sphérique et linéaire. Dans le cadre d'expérimentations antérieures, les sept semi-variogrammes ont été appliqués à la représentation de la surface de réponse du modèle hydrologique GR4J, qui est conceptuel et global, et comporte quatre paramètres de modélisation. D'abord, ces sept modèles de krigeage ont été construits à partir d'un premier groupe de 100 solutions échantillonnées par un LHS suivant une distribution uniforme. Un second groupe, différent du premier, mais construit selon la même méthode, est ensuite évalué au sein des sept modèles de krigeage et du modèle hydrologique réel. L'écart entre les valeurs de la fonction objectif $1 - NSE$ obtenu avec les sept modèles de krigeage et le modèle hydrologique GR4J est alors calculé et est représenté sous forme de boîte à moustache à la Figure 7.1 pour deux bassins versants testés (Oklahoma et Texas). Selon cette figure, les processus de corrélation gaussien combiné à une régression linéaire ou exponentiel apparaissent comme étant ceux qui représente le mieux la surface de réponse du modèle hydrologique GR4J étant donné les plus faibles écarts interquartiles, pour les deux bassins.

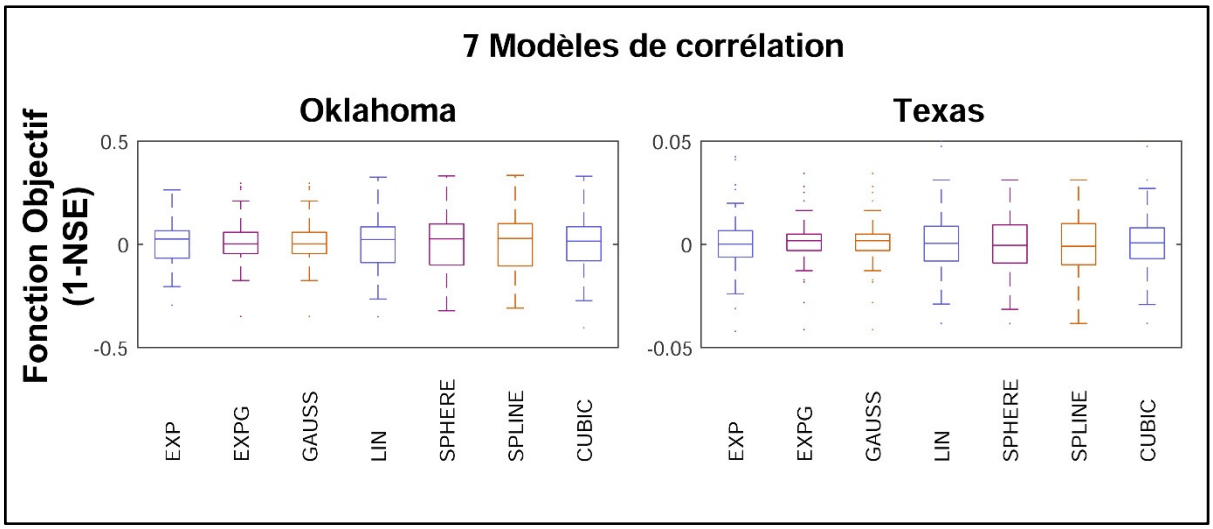


Figure 7.1. Écarts entre les valeurs de la fonction objectif $1 - NSE$ obtenues avec les sept modèles de corrélation et celles obtenues sur le modèle hydrologique GR4J pour les bassins versant Oklahoma et Texas

Pour la suite, la combinaison modèle gaussien-régression linéaire a été sélectionnée et forme le modèle de krigeage $K(x)$ qui est construit à partir des historiques de points évalués S_k. La formulation permettant à $K(x)$ d'évaluer un jeu de paramètres x constitué de N paramètres à modéliser dans le sous-domaine D_p est exprimée comme suit :

$$K(x) = \left(\prod_{i=1}^{N} e^{-(w_i - x_i)^2} \, (Gaussien) + \sum_{i=1}^{N} a_i * x_i + b_i \, (Linéaire) \right) \quad tel\ que \quad w \in S_k \quad et \quad x \in D_p \quad et \quad a, b \in \mathbb{R}$$

où $(w_i - x_i)^2$ représente l'erreur quadratique moyenne entre le jeu paramètres x et tous les autres points évalués dans S_k.

7.2 Expérimentations

Les Figures 7.2 à 7.5 présentent la représentativité obtenue par les fonctions polynomiales et les modèles de krigeage sur les problèmes d'optimisation HYDROTEL 10 et HYDROTEL 19 pour le bassin versant Ceizur uniquement, bien que les expérimentations ont été conduites sur les trois bassins versants. Des résultats similaires sont obtenus sur les deux autres bassins versants et se trouvent à l'Annexe V. À l'exception du domaine global D_{100}, deux scénarios d'échantillonnage pour les LHS sont illustrés pour chaque combinaison de S_k et de D_p : les croix bleues représentent les expérimentations pour lesquelles le sous-domaine D_p est positionné de manière aléatoire dans le domaine global et les croix pourpres représentent le sous-domaine D_p positionné autour de la meilleure solution connue (qui est un optimum local ou global). La comparaison s'effectue entre la fonction objectif 1 moins le critère objectif de Nash-Sutcliffe ($1 - NSE$) obtenu par les fonctions de surface et les problèmes hydrologiques originaux. Le coefficient de corrélation de rang de Spearman (R_s) et le coefficient de détermination (R^2) sont évalués pour chaque combinaison de S_k et de D_p. Toal (2015) a étudié l'utilisation de modèles de krigeage au sein d'un processus d'optimisation lorsqu'appliqué à certains problèmes analytiques et conceptuels. Cette étude conclut que les fonctions de surface dont le coefficient de détermination (R^2) est supérieur à 0,9 sont suffisamment représentatives pour bénéficier au processus d'optimisation. Dans cette section,

la valeur de 0,9 est donc considérée comme une valeur de référence visant à atteindre un niveau de représentativité élevé pour les coefficients R_s et R^2.

Fonctions polynomiales

Tout d'abord, la Figure 7.2 montre que les fonctions polynomiales n'offrent pas une représentation très juste du domaine global (D_{100} - S_{100} et D_{100} - S_{200}) pour le problème original HYDROTEL 10 puisque les coefficients R_s et R^2 sont tous deux éloignés de la valeur recommandée de 0,9. Sur ces mêmes tailles d'échantillons (S_{100} et S_{200}) pour HYDROTEL 19, des coefficients similaires ou moins représentatifs sont illustrés à la Figure 7.3 pour le domaine global D_{100}. Cela signifie alors que l'utilisation de fonctions polynomiales pour représenter le domaine global dans un processus d'optimisation est difficilement envisageable compte tenu que même avec 200 points évalués, la représentativité reste tout de même assez faible.

Puis, les Figures 7.2 et 7.3 présentent une représentativité qui diffère entre les deux scénarios d'échantillonnage (zone aléatoire versus autour du meilleur point connu) pour les tailles d'échantillons S_{25} et S_{50}. Entre autres, les combinaisons S_{25} - D_{10} et S_{50}- D_{10} de la Figure 7.2 montrent des coefficients de représentativité très élevés pour les points échantillonnés dans une zone aléatoire, mais très faibles pour les points échantillonnés autour du meilleur point connu. La combinaison S_{25} - D_{25} de la Figure 7.3 présente plutôt des résultats opposés entre les deux scénarios d'échantillonnage. Cela peut être causé par à une distribution non optimisée du LHS dans le sous-domaine qui a servi à construire la fonction polynomiale et ainsi conduire à une mauvaise représentation. Cependant, la possibilité plus probable est celle que la fonction polynomiale soit incapable de représenter adéquatement (par un polynôme linéaire) une surface de réponse bruitée et/ou plus difficile à représenter qui pourrait être particulièrement le cas dans les zones autour du meilleur point connu.

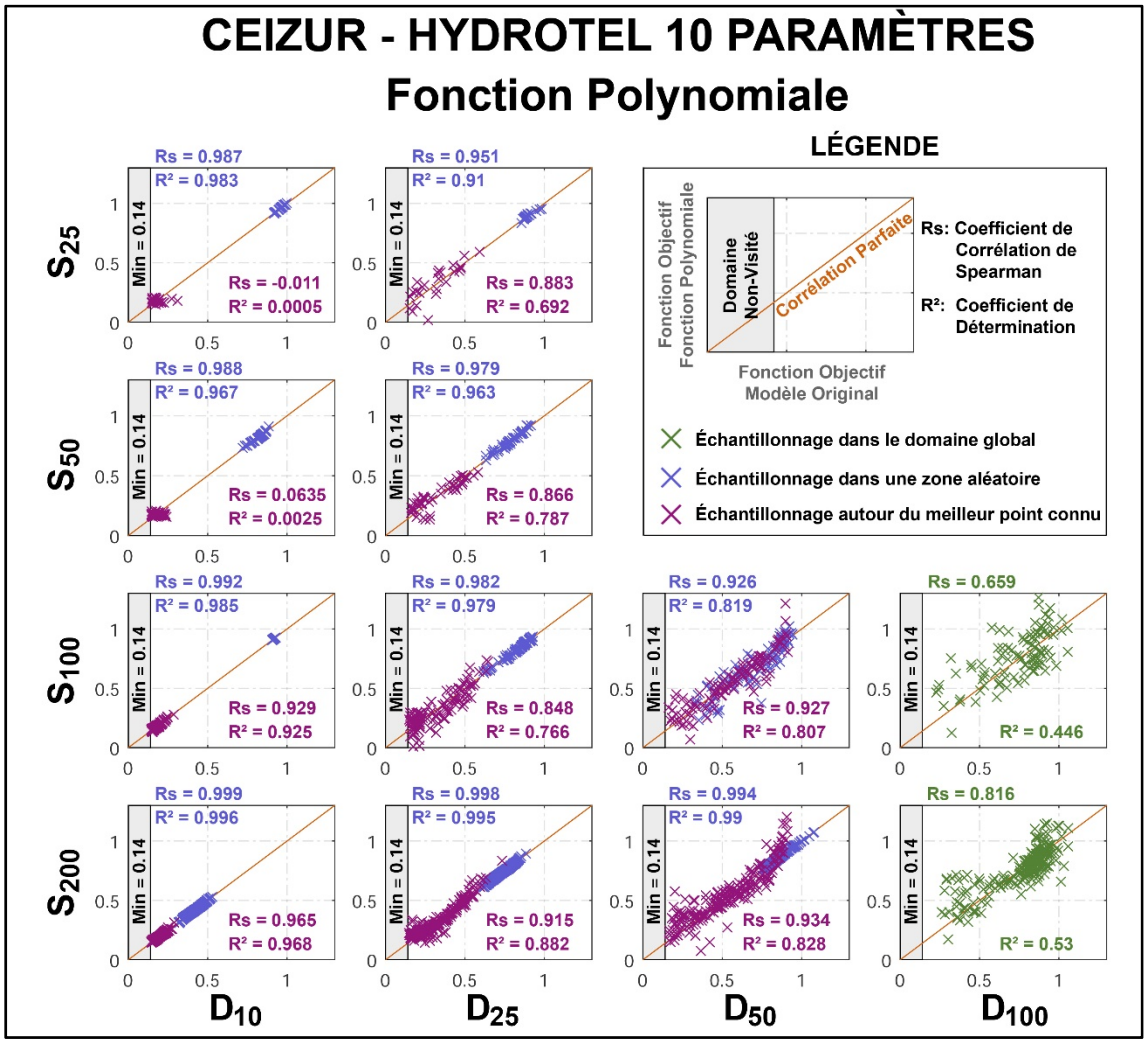


Figure 7.2. Évaluation de la représentativité des fonctions polynomiales selon les valeurs de la fonction objectif $1 - NSE$ obtenues sur HYDROTEL 10 pour le bassin versant Ceizur. Les coefficients R_s et R^2 sont présentés pour chacun des deux scénarios d'échantillonnage

La Figure 7.2 présente généralement de meilleurs coefficients de représentativité pour le scénario d'échantillonnage dans la zone aléatoire par rapport à celui autour du meilleur point connu. Cependant, les meilleurs coefficients de représentativité sont partagés entre les deux scénarios d'échantillonnage sur la Figure 7.3. Cela peut signifier que les fonctions polynomiales ne sont pas en mesure de représenter les problèmes originaux HYDROTEL 10 et HYDROTEL 19 de manière similaire. De plus, il est possible d'observer que les coefficients R_s et R^2 ont tendance à diminuer plus le domaine de l'espace paramétrique

s'agrandit ; c'est-à-dire que les Figures 7.2 et 7.3 présentent de façon générale la meilleure performance sur le domaine D_{10} et la moindre sur le domaine D_{100}. Comme on pouvait s'y attendre, les sous-domaines de petite taille sont plus facilement représentés par des fonctions polynomiales que les domaines globaux. Enfin, toutes les combinaisons de D_{10} ou D_{25} avec S_{100} ou S_{200} semblent fournir les meilleurs coefficients de représentativité, quels que soient le problème HYDROTEL et le scénario d'échantillonnage ; c'est-à-dire que les coefficients R_s et R^2 sont supérieurs ou très près de la valeur recommandé de 0,9.

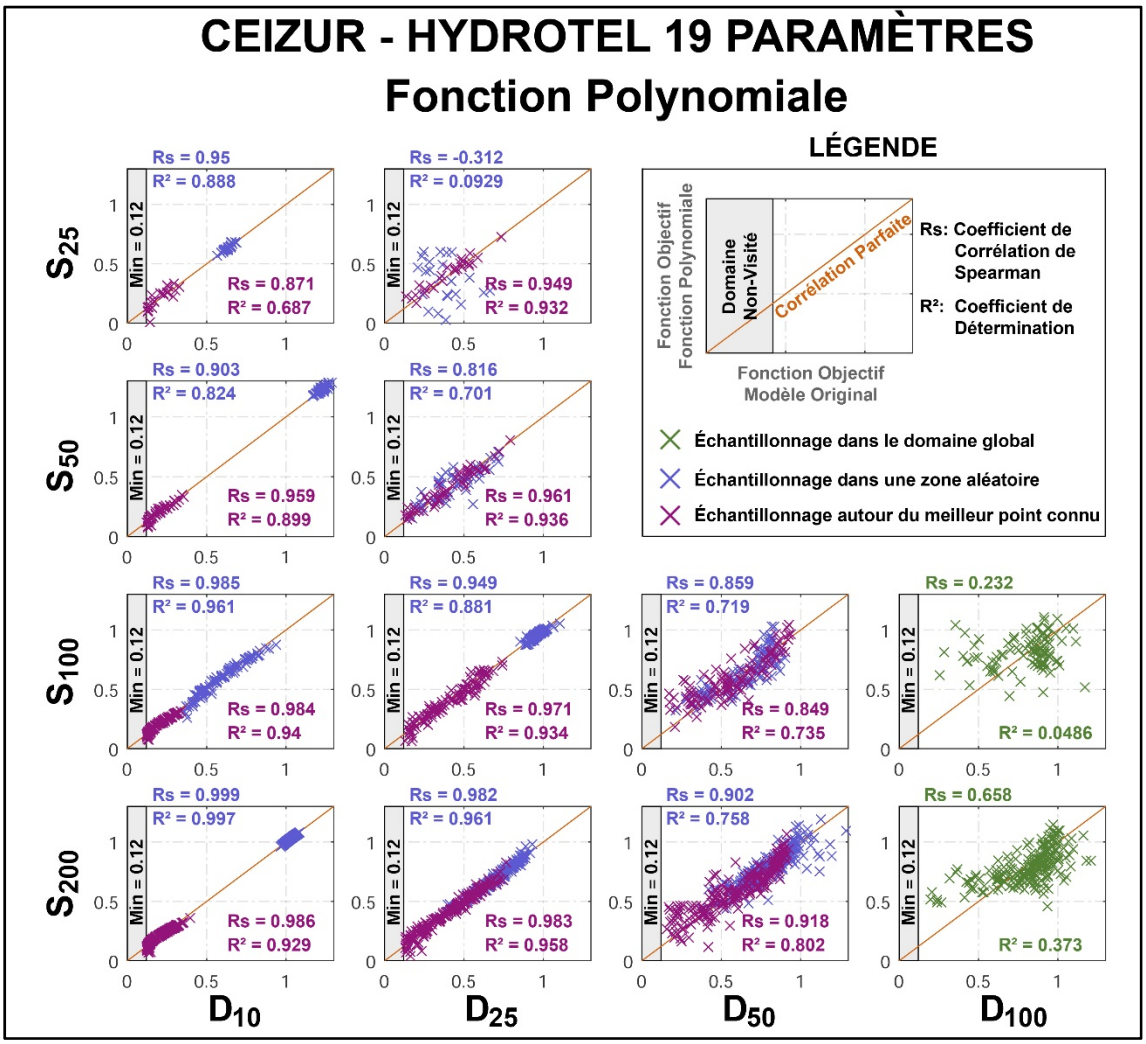


Figure 7.3. Évaluation de la représentativité des fonctions polynomiales selon les valeurs de la fonction objectif $1 - NSE$ obtenues sur HYDROTEL 19 pour le bassin versant Ceizur. Les coefficients R_s et R^2 sont présentés pour chacun des deux scénarios d'échantillonnage

Modèles de krigeage

Les modèles de krigeage offrent un même niveau assez faible de représentativité pour le domaine global (D_{100} - S_{100} et D_{100} - S_{200}). Il est donc difficile de concevoir une exploitation efficace des modèles de krigeage dans un processus de calibration étant donné une lacune importante dans la représentation globale des problèmes originaux. Sur de plus petit sous-domaines, les modèles de krigeage représentant HYDROTEL 10 et HYDROTEL 19 obtiennent généralement les meilleurs coefficients de représentativité lorsque le nombre de points échantillonnés est élevé (D_{10} - S_{100}, D_{10} - S_{200}, D_{25} - S_{100} et D_{25} - S_{200}). De plus, la combinaison de S_{50} - D_{10} sur la Figure 7.4 montre des résultats extrêmement divergents entre les coefficients R_s et R^2 obtenus sur le scénario aléatoire (dépassant la valeur de 0,9) et ceux obtenus sur le scénario autour du meilleur point connu (près d'une valeur nulle). Les raisons derrière ces résultats sont probablement les mêmes que celles énoncées pour les fonctions polynomiales. Aussi, les coefficients R_s et R^2 des Figures 7.4 et 7.5 tendent à décroître légèrement au fur et à mesure que le sous-domaine paramétrique modélisé s'agrandit, et ce pour toutes les tailles d'échantillons. Toutefois, les valeurs de ces coefficients restent au-dessus ou près de la valeur de référence de 0,9. Plus le domaine paramétrique est de petite taille, plus les modèles de krigeage semblent être en capacité de bien représenter la surface de réponse des problèmes originaux. Il est important de mentionner que ces commentaires sont basés sur les comportements moyens et que des exceptions sont remarquées ici et là.

7.3 Discussion comparative

Les résultats présentés aux Figures 7.4 et 7.5 pour les modèles de krigeage sont relativement similaires à ceux obtenus avec les fonctions polynomiales présentés aux Figures 7.2 et 7.3. Une légère différence entre les deux fonctions de réponse de surface réside dans le fait que les modèles de krigeage présentent des coefficients de représentativité légèrement supérieurs pour les mêmes expérimentations que les fonctions polynomiales.

D'un point de vue général face aux résultats obtenus par les deux fonctions de surface, des points intéressants émergent des expérimentations. D'abord, les deux fonctions de surface

sont certainement en capacité de représenter les problèmes originaux sur des sous-domaines de petite taille avec des tailles d'échantillons élevées. Toutes les combinaisons de D_{10} ou D_{25} avec S_{100} ou S_{200} pour les deux problèmes HYDROTEL et les deux scénarios d'échantillonnage fournissent des R_s et R^2 supérieurs ou proches de la valeur recommandée de 0,9. Également, avec de plus petites tailles d'échantillons (S_{25} et S_{50}), les résultats démontrent que les deux fonctions de surface peuvent éprouver des difficultés à représenter dans certains cas les sous-domaines. Ceci peut être expliqué par un manque de points évalués ou par un positionnement non-optimisé du LHS pour permettre de bien informer la fonction de surface de toutes les irrégularités dans la surface de réponse des problèmes hydrologiques originaux. Le bruit dans la surface de réponse des problèmes peut aussi entraîner une diminution parfois radicale de la performance des fonctions de surface. Finalement, les coefficients de représentativité obtenus démontrent clairement que ni les fonctions polynomiales, ni les modèles de krigeage peuvent représenter efficacement les surfaces de réponse des problèmes hydrologiques originaux lorsque les domaines globaux sont modélisés.

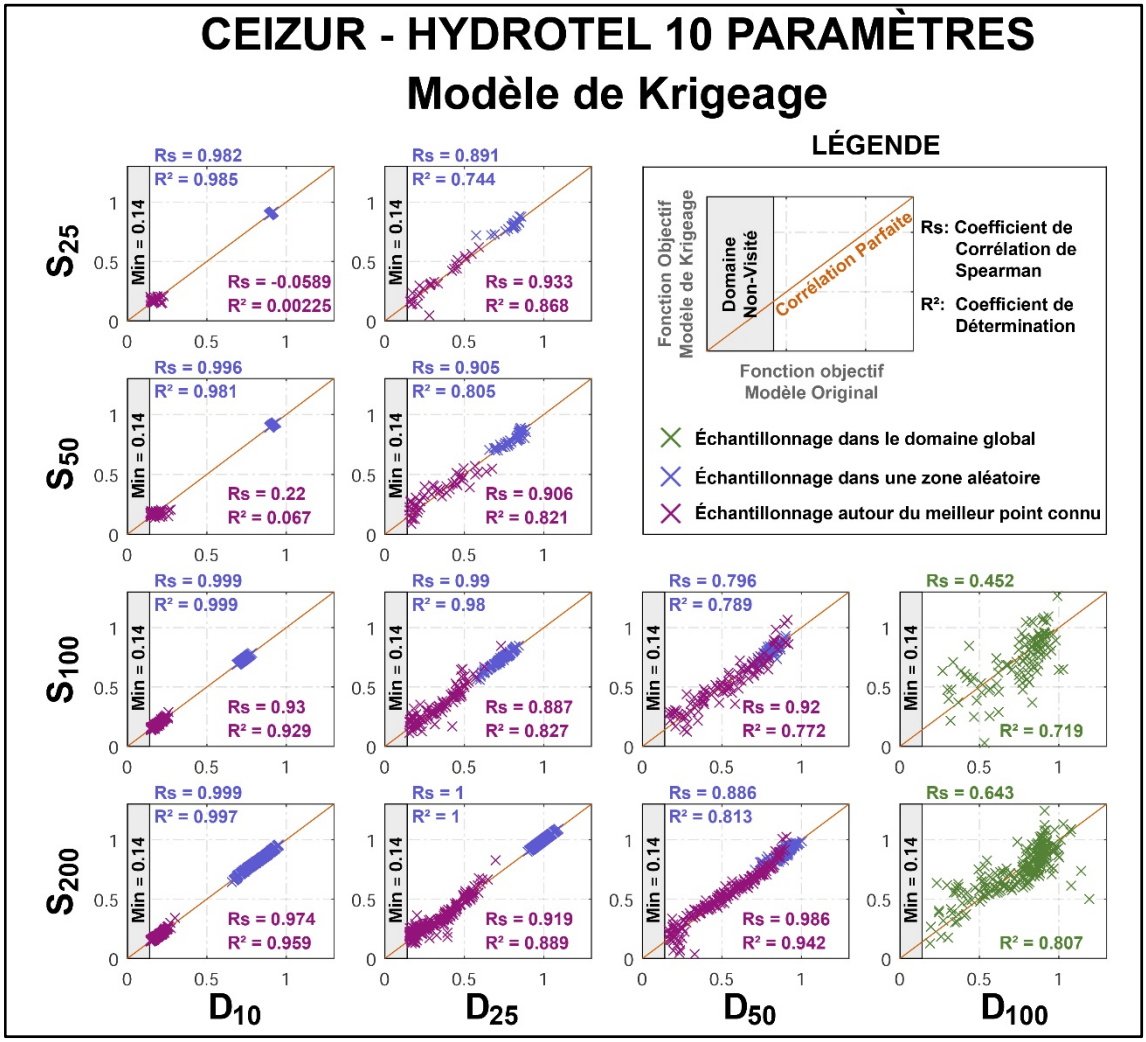


Figure 7.4. Évaluation de la représentativité des modèles de krigeage selon les valeurs de la fonction objectif $1 - NSE$ obtenues sur HYDROTEL 10 pour le bassin versant Ceizur. Les coefficients R_s et R^2 sont présentés pour chacun des deux scénarios d'échantillonnage

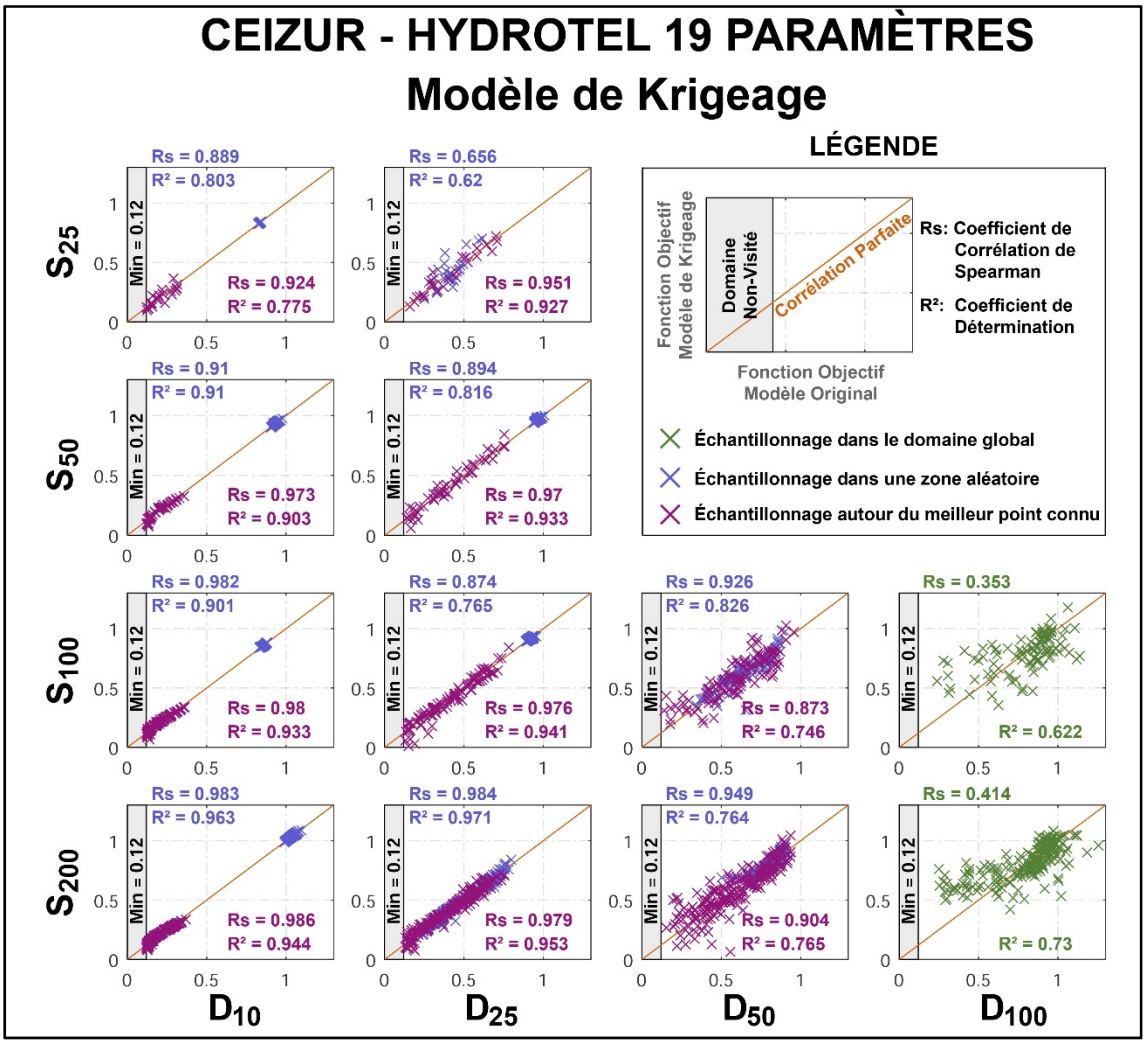


Figure 7.5. Évaluation de la représentativité des modèles de krigeage selon les valeurs de la fonction objectif $1 - NSE$ obtenues sur HYDROTEL 19 pour le bassin versant Ceizur. Les coefficients R_s et R^2 sont présentés pour chacun des deux scénarios d'échantillonnage

CHAPITRE 8

DISCUSSION GÉNÉRALE

Ce chapitre présente une discussion générale qui s'étend sur l'ensemble des résultats obtenus dans chacun des trois articles ainsi que ceux du Chapitre 7.

8.1 Approche de calage plus efficace

Le développement de nouvelles approches de calage exige un examen précis des caractéristiques intrinsèques du problème d'optimisation pour arriver à fournir une approche efficace et adaptée au problème optimisé. Comme il a été démontré par la mise en place de l'approche d'optimisation hybride DDS-MADS, cet examen ainsi que les différentes étapes de développement demandent de nombreuses expérimentations et analyses. Cependant, la calibration de modèles hydrologiques couvre une petite partie de la recherche ou des pratiques opérationnelles liées au domaine de l'hydrologie. C'est le modèle hydrologique une fois calé qui est au final employé pour la prévision hydrologique et qui alimente la modélisation hydraulique de rivières, l'étude des zones inondables ou des impacts des changements climatiques sur le réseau hydrique. Pour ces raisons, le développement d'approches d'optimisation efficace devient donc nécessaire lorsque certaines problématiques comme celle des modèles hydrologiques coûteux en temps de calcul surviennent. L'utilisateur doit alors employer une approche de calage simple et qui répond adéquatement à ses besoins.

L'approche d'optimisation hybride DDS-MADS développée au Chapitre 4 présente quatre avantages importants en comparaison avec les autres algorithmes d'optimisation existants. Premièrement, cette approche est construite de manière à n'exiger aucun travail préalable à son emploi de la part de l'utilisateur. Cela signifie qu'un faible niveau de connaissance de la réponse du modèle hydrologique à optimiser est requis et qu'aucune analyse de sensibilité n'est exigée pour l'ajustement des paramètres internes à l'approche d'optimisation. En plus, cette approche exploite une stratégie d'arrêt automatique basée sur la qualité de la fonction objectif où le processus d'optimisation se terminera lorsque les conditions d'optimalité sur le

jeu de paramètres seront rencontrées. Le budget maximal de simulations toléré par l'utilisateur n'est alors pas obligatoire, mais peut être spécifié si celui-ci représente une contrainte d'optimisation importante. Cela aura pour impact de terminer le processus d'optimisation une fois le budget maximal de simulations atteint. Si tel n'est pas le cas, l'approche hybride DDS-MADS prend en charge le processus d'arrêt automatique (convergence vers une solution optimale) et permet alors d'éviter la prise en charge par l'utilisateur du moment où le processus de calage se doit d'être interrompu.

Deuxièmement, l'approche DDS-MADS est conçue pour être exécutée sur une ressource informatique à processeur simple, ce qui peut être fortement bénéfique lorsque les ressources informatiques sont limitées. L'utilisation est alors simplifiée pour un contexte opérationnel ou de recherche scientifique, devenant aussi attrayante pour de nombreux étudiants qui disposent d'un seul ordinateur comme outil de recherche. Dans un contexte où les ressources informatiques sont plus vastes (plusieurs processeurs de calcul), il est aussi possible de transposer l'approche DDS-MADS à une version qui exploite le parallélisme ; c'est-à-dire où plusieurs simulations du modèle sont exécutées de façon simultanées. Les auteurs des deux algorithmes DDS et MADS proposent leurs versions équivalentes utilisant le parallélisme. Le « Parallele Dynamically Dimensioned Search » (PDDS; Tolson et al. 2007) et le logiciel NOMAD (Le Digabel, 2001, Audet et al. 2008) fournissent des versions pour l'utilisation des deux algorithmes dans un environnement informatique parallèle. Bien que les travaux de cette étude n'aient pas couvert le parallélisme, il devient envisageable d'exploiter l'approche DDS-MADS en mode parallèle tout en adaptant les cinq fonctionnalités de transition à un environnement de simulations en parallèle.

Troisièmement, l'approche hybride combine les stratégies avantageuses et complémentaires des deux algorithmes DDS et MADS de manière à fournir une méthode pleinement adaptée au calage du modèle hydrologique HYDROTEL. La capacité d'exploration de l'espace paramétrique de l'algorithme DDS permet de conduire rapidement le processus d'optimisation vers des zones de l'espace paramétrique de bonne qualité. Ensuite, lorsque la phase d'exploration s'épuise, l'algorithme MADS assure un raffinement local de la valeur de

la fonction objectif en convergeant vers un optimum local. Toutefois, il arrive que l'algorithme DDS reste emprisonné dans des zones de mauvaise qualité lors de son exploration globale (voir Figures 4.5 et 4.9) puisque celui-ci est principalement basé sur une approche d'exploration globale heuristique et n'est pas dotée de mécanisme d'évasion. Si un tel cas survient, le processus d'ajustement initial du « mesh » de l'algorithme MADS permet de récupérer une exploration globale défaillante tout en utilisant le processus d'expansion et de raffinement du « mesh » pour échapper à ces zones de l'espace paramétrique de moindre qualité.

Finalement, le principal avantage de l'approche d'optimisation hybride DDS-MADS est sa capacité à réduire de façon significative le nombre d'évaluations du modèle hydrologique, et par conséquent, le temps de calcul. Pour les trois bassins versants modélisés, un gain en temps de calcul moyen de 70 % est obtenu pour la version HYDROTEL à 10 paramètres et 40 % pour la version HYDROTEL à 19 paramètres en comparaison avec un processus d'optimisation de 2000 évaluations du modèle hydrologique (voir les Figures 4.7, 4.8 et 4.10). De plus, l'approche DDS-MADS a démontré une stabilité quant aux valeurs finales de la fonction objectif puisqu'un faible écart interquartile et aucune valeur aberrante pour les problèmes HYDROTEL à 10 et 19 paramètres ont été obtenus. Les résultats ont également démontré que, malgré la diminution significative du temps de calcul, les processus de calage complétés à l'aide de l'approche d'optimisation hybride DDS-MADS génèrent des valeurs finales de la fonction objectif qui sont comparables à celles obtenues avec un algorithme existant tel que l'algorithme DDS. En guise de comparaison, l'algorithme DDS a été lancé avec des budgets de simulations équivalents, soit de 600 et 1200 simulations pour HYDROTEL à 10 et 19 paramètres respectivement. Les résultats ont montré que l'algorithme DDS pouvait rester emprisonné dans des zones de l'espace paramétrique de moins bonne qualité et ainsi produire certaines valeurs finales aberrantes de la fonction objectif.

8.2 Modèles substituts représentatifs et moins coûteux en temps de calcul

L'exploitation de modèles substituts à même le processus de calage exige une compréhension adéquate des caractéristiques propres à chacune des deux familles de modèles substituts évaluées dans le cadre de cette thèse : les modèles à fidélité réduite et les fonctions de surface. La famille à laquelle appartient le substitut peut déjà orienter les choix méthodologiques liés à la manière dont il sera exploité dans le processus d'optimisation. En plus de leurs caractéristiques propres qui sont d'abord décrites ci-après, les résultats et analyses des modèles à fidélité réduite et des fonctions de surface obtenus dans le cadre des expérimentations de cette thèse sont ensuite discutés. Les problèmes originaux employés pour les expérimentations des deux familles de modèles substituts étudiées sont les mêmes, soient les deux versions HYDROTEL (10 et 19 paramètres) modélisant les trois bassins versants : Ceizur, Cowansville et Toulnustouc.

Modèles à fidélité réduite

En ce qui a trait aux modèles à fidélité réduite, leur nature oblige une construction a priori au processus de calage et il n'est en aucun cas possible, au cours du processus d'optimisation, d'améliorer la qualité de la représentation à l'aide de nouveaux jeux de paramètres évalués. Ceci met alors un poids considérable sur l'importance d'assurer une représentativité élevée entre le modèle original et le modèle à fidélité réduite dès sa construction. Comme il représente une simplification du problème d'optimisation original, le temps de calcul lié à l'évaluation d'un jeu de paramètres au sein du substitut n'est pas négligeable comme cela est souvent le cas avec les fonctions de surface. Il est donc primordial d'obtenir un ratio entre le temps de calcul du problème original et celui du modèle substitut le plus élevé possible.

Cette thèse a exploré un éventail de possibilités quant à la mise en place de modèles substituts. L'Article 2 présenté au Chapitre 5 a élaboré plusieurs voies de simplification des deux versions paramétriques du modèle hydrologique HYDROTEL. L'objectif était d'obtenir des modèles à fidélité réduite qui présentent des niveaux de représentativité élevés et des ratios de temps de calcul importants entre les problèmes originaux et les modèles substituts,

en référence dans l'Article 2 au « CPU time ratio ». Trois types de simplifications qui apparaissaient comme les plus prometteuses ont été explorées : (1) la réduction du nombre de pseudo-stations météorologiques qui répartit sur le territoire modélisé les conditions météorologiques (précipitations et températures), (2) la réduction de la longueur des périodes de calage et (3) la réduction du nombre d'UHRH discrétisant spatialement le territoire des bassins versants. Suite à l'analyse de chacun des types de simplification en fonction des représentativités obtenues et des ratios de temps de calcul, les résultats (Figures 5.2 à 5.4 et Tableaux 5.3 à 5.8) ont démontré que l'un ou l'autre de ces types, lorsque considéré individuellement, ne présentait pas les caractéristiques souhaitées en vue de les exploiter efficacement dans des processus de calage. Pour la plupart, les ratios obtenus entre le temps de calcul des modèles originaux et celui des modèles à fidélité réduite n'étaient pas assez prononcés, bien que les niveaux de représentativité étaient élevés.

Suite à cette analyse des résultats, la combinaison des trois types de simplifications des modèles originaux a été envisagée dans le but d'augmenter le ratio de temps de calcul entre les modèles originaux et les modèles à fidélité réduite, tout en sachant que les niveaux de représentativité allaient être affectés. Les résultats (Figure 5.5, Tableaux 5.9 et 5.10) ont démontré que la combinaison des avenues (1), (2) et (3) pour la version HYDROTEL à 10 paramètres et la combinaison des avenues (1) et (2) pour la version HYDROTEL à 19 paramètres fournissaient les meilleurs compromis entre la représentativité et le ratio de temps de calcul. En somme, des ratios de temps de calcul entre les modèles originaux et substituts s'étendaient de 16 à 44 pour la version HYDROTEL à 10 paramètres alors qu'un ratio de 2 était plutôt la valeur moyenne pour la version HYDROTEL à 19 paramètres. Cette différence de qualité des ratios entre les versions HYDROTEL à 10 et 19 paramètres démontre que la production des hydrogrammes géomorphologiques dans la version à 19 paramètres est responsable d'une part importante du temps de calcul de la simulation du modèle HYDROTEL. L'augmentation des ratios de temps de calcul a eu pour effet de diminuer les niveaux de représentativité des modèles à fidélité réduite finaux pour les deux versions HYDROTEL sur les trois bassins versants modélisés, générant pour la plupart des coefficients de représentativité qui n'atteignent pas la valeur de 0,9 recommandée par Toal

(2015). Ces modèles à fidélité réduite finaux représentent tout de même les meilleurs substituts potentiels développés dans ces travaux.

Fonctions de surfaces

Comme il a été spécifié dans la revue de littérature, les fonctions de surface nécessitent obligatoirement l'évaluation d'un certain nombre de jeux de paramètres de manière à créer un ensemble de points qui permet à la fonction substitut de représenter au maximum de ses capacités la fonction objectif originale. De plus, au fur et à mesure que le processus d'optimisation progresse, le nombre de points évalués au sein de la fonction objectif originale augmente. Il devient donc possible de construire en cours de route (mise à jour) la fonction de surface à partir de cet ensemble de solutions connues qui ne cesse de grandir. Cette mise à jour continuelle de la fonction de surface permet d'augmenter la représentativité du modèle substitut par rapport au problème d'optimisation original puisque l'historique de connaissance ne cesse de s'accroître et la surface de réponse de l'espace paramétrique se précise toujours davantage. Les fonctions de surface ont également une qualité fortement pertinente pour les processus de calage coûteux en temps de calcul ; c'est-à-dire que le temps lié à la construction de la fonction de surface et à l'évaluation d'un jeu de paramètres au sein de celle-ci est très peu coûteux par rapport aux temps de calcul consommé par la simulation. De ces faits, l'exploitation des fonctions de surface peut alors être bénéfique durant tout le processus d'optimisation et ainsi apporter un appui considérable au calage d'un problème coûteux en temps de calcul.

Le Chapitre 7 expose les résultats et l'analyse de deux fonctions de surface : les fonctions polynomiales et les modèles de krigeage où chacun a été évalué en fonction de la représentativité des modèles originaux qu'ils pouvaient offrir. Le ratio de temps de calcul entre les modèles originaux et les modèles substituts n'est pas ici évalué compte-tenu de sa charge négligeable dans le temps de calcul total du processus d'optimisation. Toutefois, deux facteurs ont été pris en compte quant à leur capacité de représenter les problèmes originaux, soient le nombre de points évalués servant à la construction de la fonction de surface ainsi que la taille du domaine paramétrique modélisé par le substitut. Ces deux facteurs peuvent

apporter un impact considérable sur la capacité de la fonction de surface à représenter le problème original.

Les résultats obtenus aux Figures 7.1 à 7.4 (ainsi que celles placées à l'Annexe V) ont apporté des points intéressants quant aux potentiels d'exploitation des fonctions de surface et ce, pour les deux versions du modèle hydrologique HYDROTEL. D'abord, les fonctions polynomiales et les modèles de krigeage utilisés ont démontré de façon générale des niveaux similaires de performance. Tous deux ont la capacité de bien représenter les problèmes originaux sur des petites tailles de domaines paramétriques, à condition d'avoir un nombre de points évalués égal ou supérieur à 100. À l'inverse, avec un nombre de points évalués inférieur à 100, la qualité de représentation semble être moins constante, variant de très basse à très élevée et ce, même sur les petites tailles de domaine paramétrique. Finalement, hors de tout doute, ni les fonctions polynomiales, ni les modèles de krigeage sont en mesure de fournir une bonne qualité de représentation pour le domaine global des problèmes originaux. Il est donc peu concevable d'exploiter ces deux fonctions de surface dans une représentation globale des problèmes originaux du modèle hydrologique HYDROTEL, même avec un historique de 200 solutions connues. Bien entendu, plus le nombre de points évalués augmente, plus la qualité de la représentation du domaine global tendra à s'améliorer. Toutefois, dans le contexte de calage du modèle hydrologique HYDROTEL où une seule simulation consomme plusieurs minutes, le nombre de simulations des problèmes originaux vise à être limité pour réduire le temps de calcul. Rappelons qu'avec l'approche de calage DDS-MADS, environ 600 simulations sont nécessaires au calage efficace de HYDROTEL 10. Ceci nuit alors à l'exploitation de ces fonctions de surface dans un rôle de représentation du domaine global puisqu'il est préférable de concentrer le processus de calibration sur les solutions évaluées dans les problèmes originaux. La représentation de domaines paramétriques plus restreint est toutefois certainement envisageable par les deux fonctions de surface étudiées au Chapitre 7.

8.3 Modèles substituts exploités au sein de méthodes d'optimisation efficaces

La méthode d'optimisation DDS-MADS a démontré qu'elle pouvait apporter des réductions significatives en terme de temps de calcul par rapport à des algorithmes d'optimisation existants en diminuant le nombre de simulations effectuées au cours du processus de calage. Razavi et al. (2010) identifient d'ailleurs le développement d'algorithmes d'optimisation plus efficaces comme l'une des pistes de solution à la problématique de calage des modèles hydrologiques coûteux en temps de calcul. Une autre façon d'aborder cette même problématique est d'exploiter des modèles substituts représentatifs et moins coûteux en temps de calcul au sein du processus d'optimisation. Les travaux présentés au Chapitre 6 démontrent que l'exploitation de ces modèles au sein de l'approche DDS-MADS contribue à une réduction encore plus importante des temps de calcul des processus d'optimisation tout en ne sacrifiant pas la qualité des solutions finales.

Les expérimentations conduites permettent d'offrir un éventail de cadres de calibration qui profitent de la synergie de l'approche d'optimisation hybride DDS-MADS et de l'exploitation des modèles substituts en fonction des divers objectifs de calibration et des contraintes d'optimisation de l'utilisateur. À titre d'exemple pour le calage de la version HYDROTEL à 10 paramètres, l'approche DDS-MADS performant strictement sur les modèles à fidélité réduite finaux permet d'obtenir en 30 minutes des solutions de bonne qualité (cadre de calibration **B**). Cependant, ce cadre de calibration répond à des objectifs de calage qui favorisent le temps de calcul par rapport à la qualité finale de la fonction objectif. À objectif inverse, pour les mêmes problèmes originaux, il serait préférable d'employer le cadre de calibration **A** qui mise davantage sur la performance de la qualité finale de la fonction objectif. Ce cadre de calibration est constitué en premier lieu d'une exploration globale au sein des modèles à fidélité réduite à l'aide de l'étape de DDS. Puis, en reprenant la meilleure solution trouvée par cette précédente étape, l'algorithme MADS performe ensuite sur les problèmes originaux. Finalement, dans un contexte où le meilleur compromis est souhaité entre la réduction du temps de calcul et la qualité finale de la fonction objectif, le cadre de calibration **C** est privilégié.

En ce qui concerne le calage de la version HYDROTEL à 19 paramètres, il est possible d'établir les mêmes associations entre objectif de calage et cadre de calibration efficace. D'abord, le cadre de calibration **B** reste pour tous les bassins versants celui qui engendre la plus grande réduction du temps de calcul aux dépends d'une valeur finale de la fonction objectif de moins bonne qualité par rapport aux autres cadres. À l'inverse, le cadre de calibration **D** offre la meilleure performance à l'égard de la valeur finale de la fonction objectif, mais consomme davantage de temps de calcul. Si l'on cherche plutôt le meilleur compromis entre la réduction du temps de calcul et la qualité de la solution finale trouvée, le cadre de calibration **E** peut être retenu. De plus, l'un des éléments importants qui a été soulevé dans les cadres de calibration appliqués aux problèmes originaux HYDROTEL à 19 paramètres est l'utilisation d'un nouveau type de modèles à fidélité réduite où le nombre de paramètres à caler est moindre. Deux paramètres de calage dans la version HYDROTEL à 19 paramètres génèrent la production d'hydrogrammes géomorphologiques qui consomme une part importante du temps de calcul. Une version HYDROTEL à 17 paramètres qui exclut ces deux paramètres a donc été développée dans le cadre de calibration **E**. L'objectif derrière cette version à 17 paramètres est de profiter à la fois d'un ratio de temps de calcul élevé entre le problème original et le modèle à fidélité réduite, tel que pour la version HYDROTEL à 10 paramètres, ainsi que d'une représentativité plus élevée comme avec la version HYDROTEL à 19 paramètres. Il est aussi intéressant de rappeler que tous les cadres de calibration performants sur les deux versions HYDROTEL, peu importent les objectifs de calage préconisés, emploient des modèles à fidélité réduite qui ont un ratio de temps de calcul élevé.

Finalement, les fonctions de surface ont aussi joué un rôle au sein des cadres de calibration présentés dans l'Article 3. L'algorithme d'optimisation MADS exécutant le raffinement local dans l'approche hybride emploie une fonction quadratique (fonction polynomiale de deuxième degré) pour représenter localement l'espace paramétrique dans le but d'ordonnancer les points qui seront par la suite évalués au sein du problème original. Tel que Conn et Le Digabel (2013) l'indiquent, cet emploi de la fonction quadratique dans l'ordonnancement des directions orthogonales visitées permet de réduire le nombre de simulations exécutées pour l'identification d'un jeu de paramètres de meilleure qualité.

8.4 Limites de l'étude

Cette thèse comporte des limites quant aux choix méthodologiques effectués et qui ont un impact certain sur les résultats obtenus. En raison du temps de calcul élevé de toutes les expérimentations menées dans cette étude, le développement des approches de calage est basé sur un nombre limité de problèmes d'optimisation. D'autres modèles hydrologiques coûteux en temps de calcul et/ou d'autres bassins versants modélisés auraient pu être testés afin d'étendre les conclusions de l'étude.

Également, le critère $1 - NSE$ a été choisi comme seule fonction objectif pour l'ensemble des expérimentations de ces travaux et constitue une limite de la thèse. Bien que le NSE soit employé comme critère de référence dans de nombreux articles scientifiques, plusieurs autres fonctions objectifs peuvent mieux cibler certains processus hydrologiques souvent analysés par les experts. À titre d'exemple, le critère de « Kling Gupta Efficiency » (KGE; Gupta et al. 2009) aurait pu être la fonction objectif utilisée pour cette thèse puisque contrairement au NSE qui accorde beaucoup d'importance aux crues, le KGE répartit mieux le calcul de son critère à l'ensemble de l'hydrogramme. L'impact du NSE sur les résultats de calage en comparaison avec d'autres critères objectifs pourrait être étudié afin de mieux évaluer la dépendance entre la fonction objectif et la méthode d'optimisation employée.

L'utilisation efficace de modèles à fidélité réduite au sein des processus de calage a suscité quelques questions concernant le niveau de représentativité nécessaire que ces substituts doivent obtenir. La littérature ne semble pas fournir de critère sur le niveau de représentativité requis pour l'utilisation efficace des modèles à fidélité réduite au sein des procédures d'optimisation. En conséquence, cette étude s'est appuyée sur les travaux de Toal (2015) qui définit un seuil de représentativité recommandé pour les fonctions de surface à une valeur de 0,9 du coefficient de détermination. Cette valeur, qui a été transposée aux modèles à fidélité réduite de cette étude, pourrait donc être revue. Il est clair que les résultats de cette thèse démontrent que malgré des coefficients de valeur inférieure à 0,9, il est tout de

même possible de bénéficier des modèles à fidélité réduite pour réduire significativement les temps de calcul des calibrations. Ceci dit, le recours à des tests statistiques constitue une avenue qui pourrait être explorée davantage en ce sens. Par exemple, lorsque la comparaison des valeurs $1 - NSE$ est faite sous la forme d'une droite (Chapitre 5 et 7), un test pourrait être appliqué afin de vérifier si cette droite est significativement différente de 1.

La comparaison des performances entre les diverses approches de calage proposées dans cette étude s'est fait via l'écart entre les résultats médians et l'étendue des écarts interquartiles. Bien que cette façon permette d'identifier les approches les plus efficaces, des tests statistiques auraient pu être employés pour différencier significativement, ou non, la performance des approches de calibration proposées.

Finalement, le Chapitre 7 a présenté une analyse de performance des fonctions polynomiales et des modèles de krigeage pour la représentation des problèmes originaux HYDROTEL. Bien que ces deux fonctions de surface soient très employées, il en existe certainement d'autres qui auraient pu être testées comme la revue de littérature du Chapitre 2 (Section 2.2.2) l'illustre.

8.5 Défis de la thèse

Le défi le plus important est celui de comprendre le problème à optimiser. Au fil de ce projet de recherche et même dans le mémoire de Huot (2014), plusieurs phases exploratoires ont été effectuées dans le but d'accumuler l'information pertinente sur les particularités des problèmes d'optimisation étudiés. L'idée est d'avoir en tête la meilleure représentation possible de ce à quoi la surface de réponse de l'espace paramétrique du problème ressemble, pour ensuite pouvoir construire une stratégie d'optimisation efficace et adaptée. Comme il existe des centaines voire des milliers de façons d'optimiser un même problème, il peut être facile de s'embourber dans des méthodes d'optimisation populaires et généralistes qui ne répondent pas aux besoins de l'utilisateur. La compréhension des particularités propres au problème optimisé devient alors une étape de conception essentielle qui permet d'adapter

davantage la méthode d'optimisation et ainsi fournir une meilleure efficacité d'optimisation. C'est finalement ce gain en efficacité qui peut par la suite se traduire en une réduction du temps de calcul et/ou une augmentation de la qualité finale de la fonction objectif. À titre d'exemple, la structure d'exécution des calculs mathématiques propre au modèle hydrologique HYDROTEL engendre la production d'hydrogrammes géomorphologiques lorsque la modification de certains paramètres internes au modèle est effectuée. Comme il a été présenté dans l'Article 3 du Chapitre 6, cette particularité permet l'exploitation d'une version du modèle hydrologique ayant un nombre inférieur de paramètres à optimiser (HYDROTEL 17) pour la réduction du temps de calcul attribué au calage du modèle hydrologique HYDROTEL à 19 paramètres. C'est donc la compréhension du problème à optimiser qui se traduit en une approche de calibration adaptée et plus efficace que des méthodes d'optimisation existantes.

Un second défi est de ne pas tomber dans le développement d'approches d'optimisation « sur-efficaces ». Si l'on souhaite définir des méthodes d'optimisation qui peuvent être utilisables dans différents domaines, problèmes ou contextes, il faut fonder la construction de l'approche d'optimisation sur des caractéristiques moyennes à un ensemble de problèmes plus ou moins similaires. En recherche et développement, on veut sans aucun doute porter le gain en efficacité comme premier objectif, mais il faut tout de même réfléchir à la polyvalence de la méthode d'optimisation. Comme il existe un très grand nombre de problèmes à optimiser, on ne peut développer une approche d'optimisation propre pour chacun étant donné toutes les étapes de conception et d'analyse que cela requiert. La sur-efficacité peut aussi engendrer une perte d'intérêt de la communauté scientifique envers cette approche. Il faut alors trouver le bon niveau d'efficacité tout en conservant une polyvalence de la méthode d'optimisation.

La gestion du temps de calcul engendré par les expérimentations est un défi à considérer pour tous les travaux liés à la problématique de cette thèse. Le temps de calcul consommé est plus particulièrement important dans le cadre de la construction des modèles à fidélité réduite où un peu plus de la moitié du temps de calcul y est associé étant donné un nombre très élevé

d'essais effectués pour chacun des types de modèles à fidélité réduite, selon chaque version HYDROTEL et bassin versant modélisé. Le défi est donc de bien arrimer les expérimentations qui conduiront aux bonnes analyses avec la capacité des ressources informatiques disponibles pour l'étude. Une sous-évaluation des temps de calcul consommés aurait alors pu conduire à la renonciation de plusieurs éléments importants de l'étude.

L'approche hybride DDS-MADS combine les stratégies avantageuses de deux algorithmes d'optimisation connus et existants dans le but de proposer une manière plus efficace de caler le modèle hydrologique HYDROTEL. Le choix de construire une approche adaptée à partir d'algorithmes existants s'explique d'abord par le lourd travail de programmation informatique que nécessite la mise en place d'une nouvelle méthode d'optimisation. D'un point de vue technique, l'un des défis est de proposer une programmation informatique sans faille et des plus compréhensibles pour les utilisateurs. L'algorithme doit être en mesure non seulement d'optimiser adéquatement le problème, mais aussi de surmonter certaines situations anormales, mais probables, qui peuvent survenir. Par exemple, un arrêt prématuré d'une simulation ne fournissant pas de valeur de fonction objectif pourrait interrompre des programmes déficients. Le choix de proposer une approche hybride à partir d'algorithmes existants en intégrant des fonctionnalités de transition simples diminue considérablement la déficience possible des programmes informatiques. Les fonctionnalités de transition ont été présentées sous forme de pseudo-code informatique (Figure 4.3) de manière à ce que chaque utilisateur s'approprie celles-ci et les transcrive dans son propre language de programmation.

Aussi, un autre défi est certainement la capacité de présenter une méthode d'optimisation qui offre soit, une meilleure performance d'optimisation à l'égard de la qualité finale de la fonction objectif ou une stratégie nouvelle qui procure certains avantages qui peuvent combler les besoins des utilisateurs. La notion de performance peut être variable d'un utilisateur à l'autre, mais certains critères objectifs comme la qualité finale de la fonction objectif ou le temps de calcul nécessaire au processus d'optimisation sont récurrents. Il faut alors présenter des méthodes d'optimisation novatrices qui améliorent significativement les objectifs et diminuent les contraintes d'optimisation. L'intégration des nouvelles approches

d’optimisation dans les habitudes des utilisateurs peut aussi être un défi important qui se relève en démontrant des résultats concrets qui justifient sa mise en place.

CONCLUSION

L'objectif de ce projet de recherche est de développer des approches d'optimisation efficaces pour la problématique de calage des modèles hydrologiques coûteux en temps de calcul. Comme problèmes d'optimisation, deux versions du modèle hydrologique distribué HYDROTEL, l'une à 10 et l'autre à 19 paramètres à optimiser ont été utilisées pour la modélisation de trois bassins versants ayant des caractéristiques physiographiques différentes. À travers les articles présentés aux chapitres 4 à 6 ainsi que le chapitre 7, il a été possible de proposer des méthodes d'optimisation qui réduisent significativement le temps de calcul des processus de calage. Deux principales avenues ont été explorées : le développement d'une approche d'optimisation plus efficace du point de vue des stratégies de recherche employées et l'exploitation de modèles substituts représentatifs et moins coûteux en temps de calcul au sein du processus d'optimisation.

Concernant le développement d'approches de calage plus efficaces, la proposition de jumeler les stratégies d'optimisation avantageuses et complémentaires des deux algorithmes existants DDS et MADS est l'une des contributions importantes de cette thèse. Avec l'approche d'optimisation hybride DDS-MADS, le nombre de simulations nécessaires aux processus de calage peut être diminué d'environ 70 % pour la version HYDROTEL à 10 paramètres et de 40 % pour la version HYDROTEL à 19 paramètres comparativement à une méthode existante exécutant 2000 simulations. De plus, cette approche hybride permet de fournir des jeux de paramètres finaux de bonne qualité qui satisfont les conditions d'optimalité.

Il a également été démontré que la construction et l'exploitation de modèles substituts au sein d'un processus de calage était une voie pertinente pour la réduction du temps de calcul. La combinaison de l'approche hybride DDS-MADS et de l'exploitation efficace des modèles substituts, tant les modèles à fidélité réduite que les fonctions de surface, apporte un effet de synergie qui bonifie la réduction des temps de calcul. De cette combinaison, un éventail de cadres de calibration a alors été testé et certains ont été identifiés comme étant les plus efficaces pour répondre à divers objectifs de calage et contraintes d'optimisation. Peu

d'études exposées dans la littérature présentent des approches d'optimisation qui combinent à la fois le développement de méthodes plus efficaces et l'exploitation des modèles substituts, plus particulièrement les modèles à fidélité réduite. Cette thèse contribue donc originalement à la recherche scientifique liée à la problématique des problèmes coûteux en temps de calcul.

Comme il a été mentionné dans les limites de l'étude (Section 8.4), Toal (2015) recommande que les fonctions de surface obtiennent un niveau de représentativité qui se traduit par un coefficient de détermination supérieur à une valeur de 0,9. Toutefois, aucune étude ne présente le même genre d'analyse pour les modèles à fidélité réduite. De plus amples expérimentations pourraient être réalisées pour analyser la relation entre le niveau de représentativité minimal et l'efficacité au sein des procédures d'optimisation de manière à fournir une valeur de référence pour les modèles à fidélité réduite à travers, par exemple, le recours à des tests statistiques.

Finalement, les caractéristiques des trois diverses avenues de simplifications des modèles hydrologiques originaux ainsi que leur combinaison ont permis de soulever la question : a-t-on besoin d'une modélisation hydrologique distribuée simplifiée dans les pratiques hydrologiques d'aujourd'hui ? Si ces modèles à fidélité réduite sont représentatifs de modèles hydrologiques distribués orignaux, peut-on trouver des applications concrètes et efficaces qui permettraient d'améliorer les pratiques hydrologiques ? De futures réflexions et travaux sur ces questions pourraient amener à des applications concrètes dans les milieux de recherche et opérationnels.

ANNEXE I

ÉTUDES PRÉLIMINAIRES

Cette annexe présente la figure pour le bassin versant de Cowansville qui complète l'étude préliminaire présenté au sein de l'Article 1 (Chapitre 4).

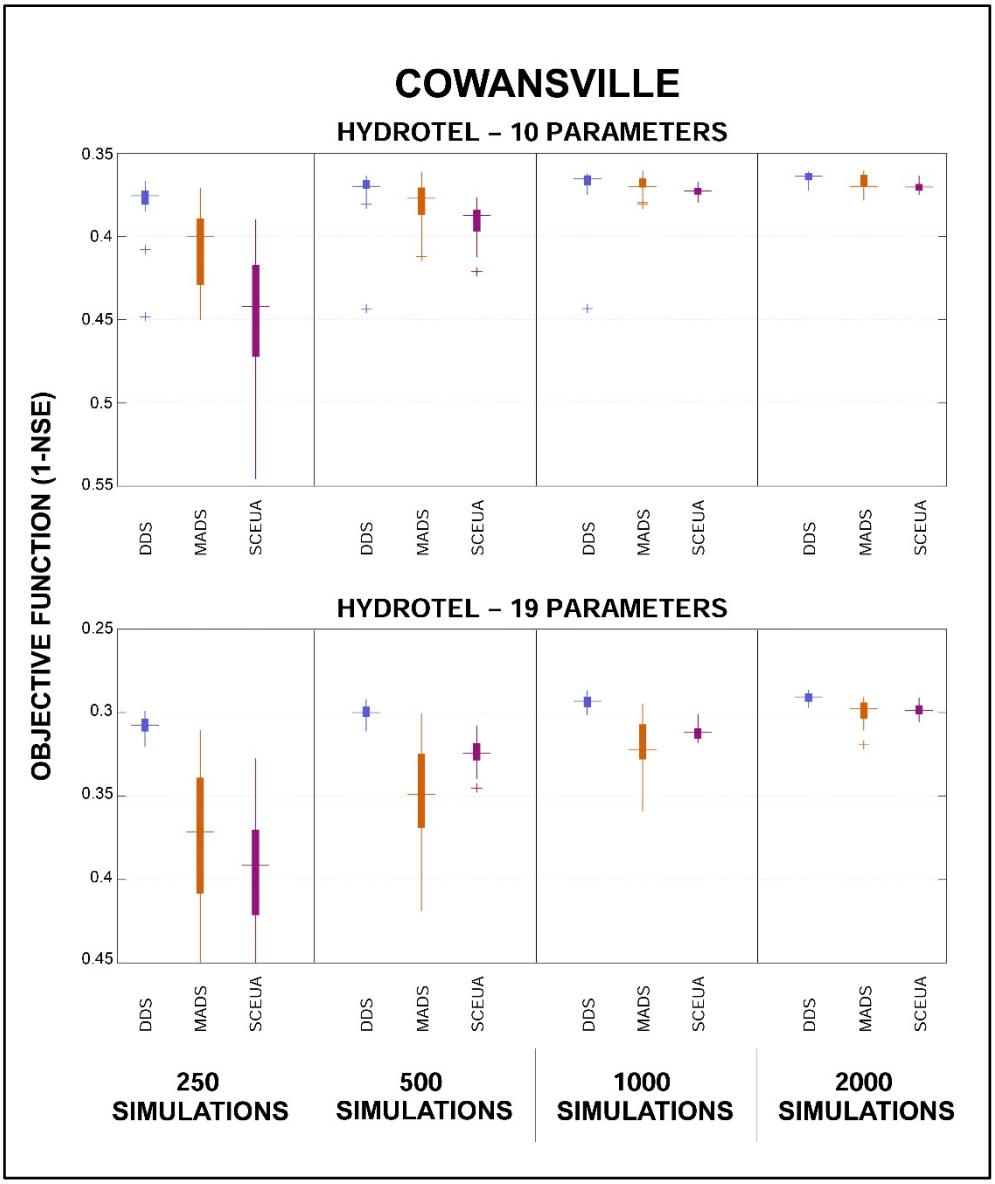


Figure-A I-1. Final $1 - NSE$ results from optimizations by DDS, MADS and SCE-UA for the calibrations of HYDROTEL 10 and HYDROTEL 19 on the Cowansville watershed at 250, 500 and 1,000 simulations, and at the end of the optimization

ANNEXE II

RÉDUCTION DU NOMBRE DE PSEUDO-STATIONS MÉTÉOROLOGIQUES

Cette annexe présente les figures pour les bassins versants de Cowansville et Ceizur pour la mise en place de modèles substituts selon diverses réductions du nombre de pseudo-stations météorologiques, complément de l'Article 2 (Chapitre 5).

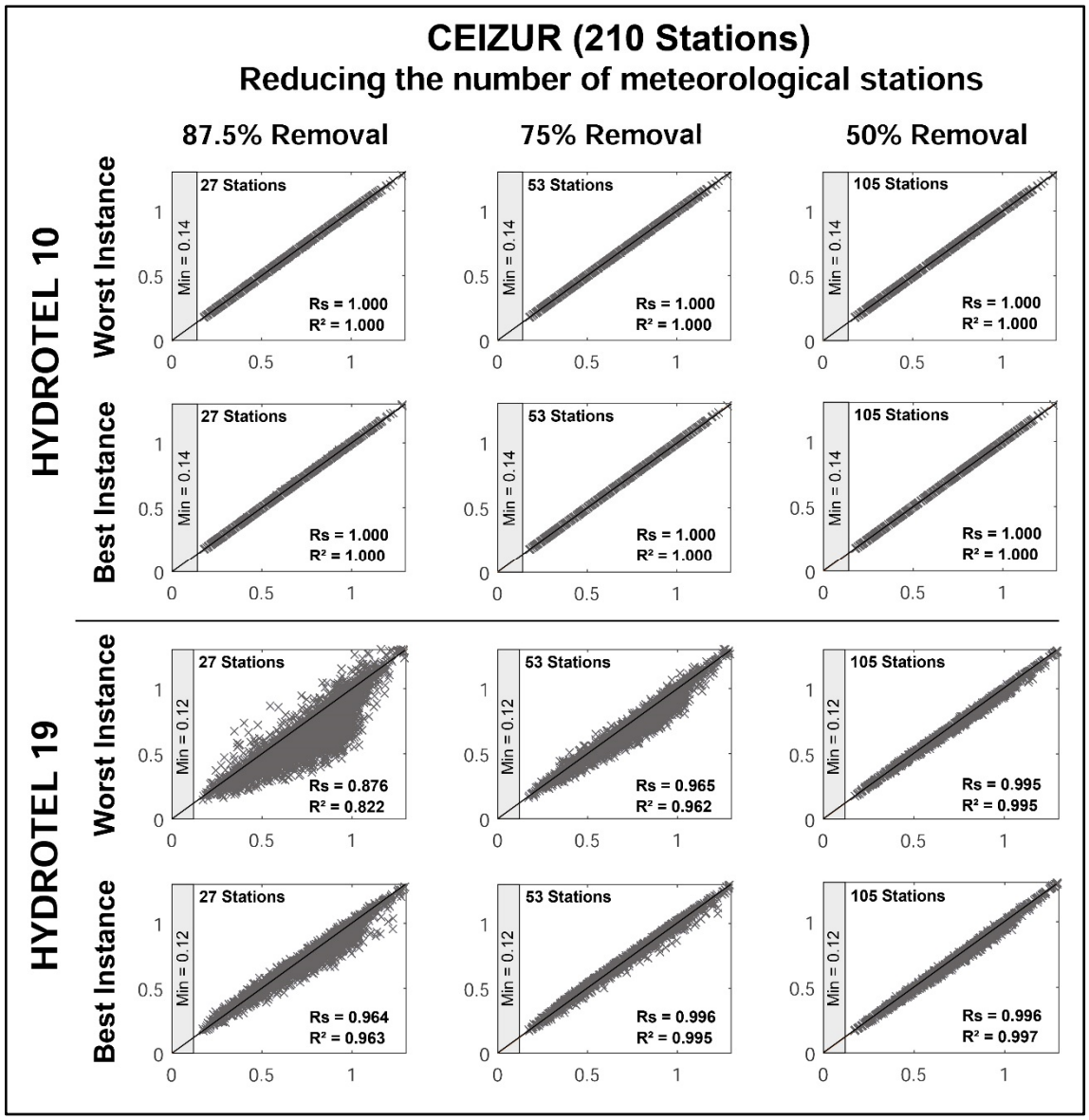


Figure-A II-1. $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Ceizur watershed comparing the surrogates having a reduced number of meteorological stations with the original models. Rs and R² coefficients are calculated for the representativeness evaluation

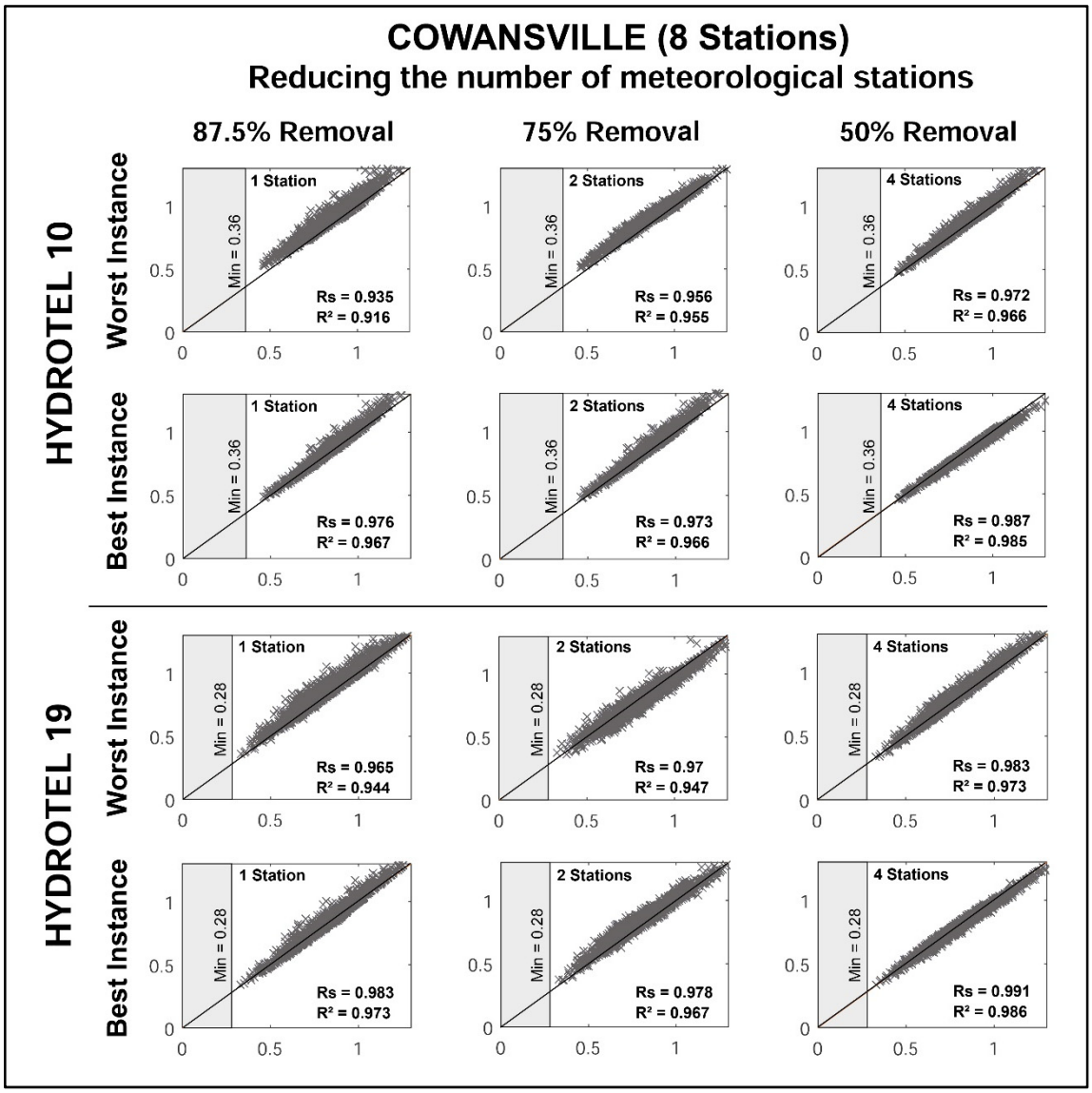


Figure-A II-2. $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Cowansville watershed comparing the surrogates having a reduced number of meteorological stations with the original models. Rs and R² coefficients are calculated for the representativeness evaluation

ANNEXE III

RÉDUCTION DE LA PÉRIODE DE CALIBRATION

Cette annexe présente les figures pour les bassins versants de Cowansville et Ceizur pour la mise en place de modèles substituts selon diverses réductions de la période de calibration, complément de l'Article 2 (Chapitre 5).

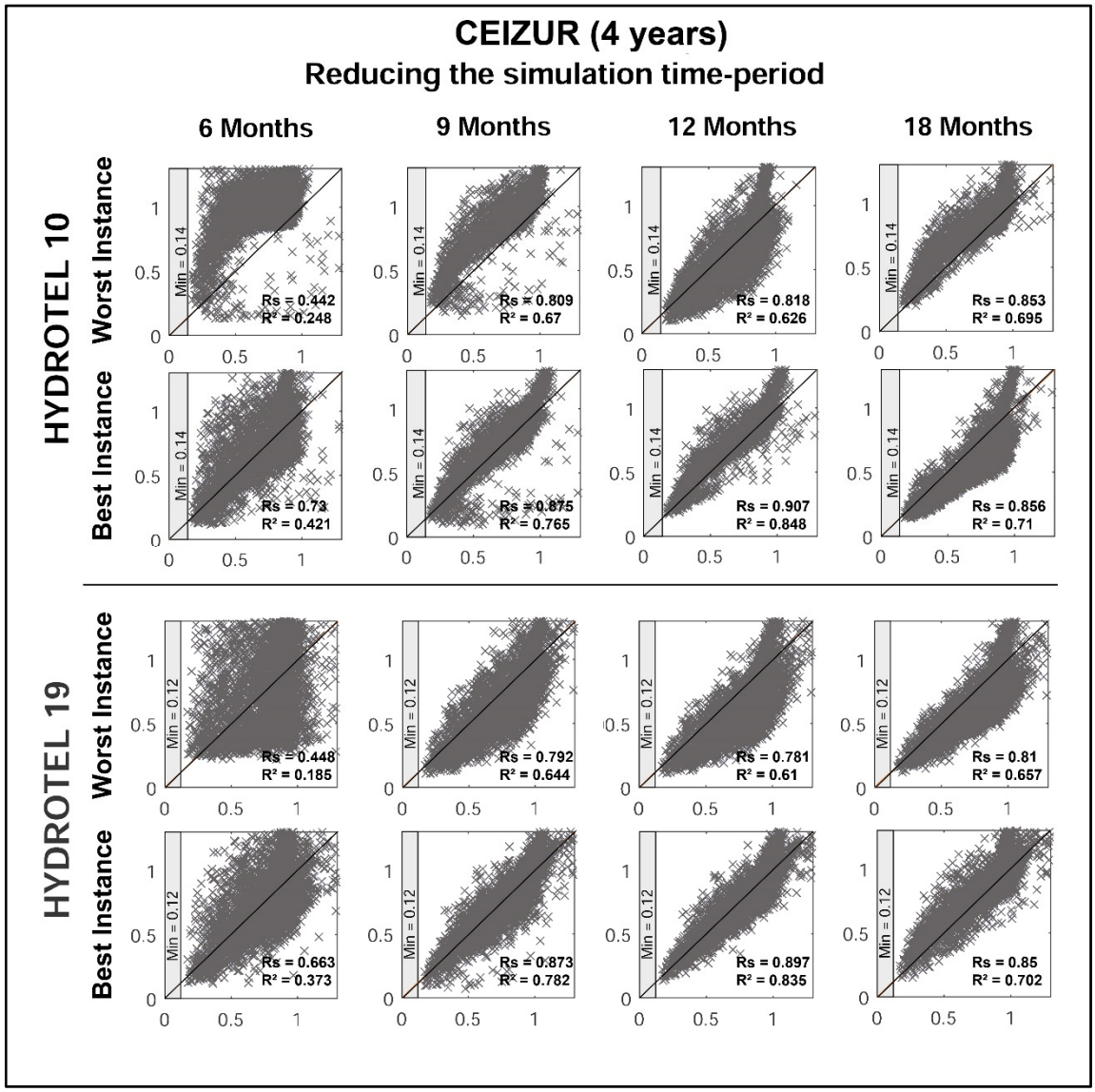


Figure-A III-1. $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Ceizur watershed comparing the surrogates having a reduced calibration time-period with the original models. Rs and R² coefficients are calculated for the representativeness evaluation

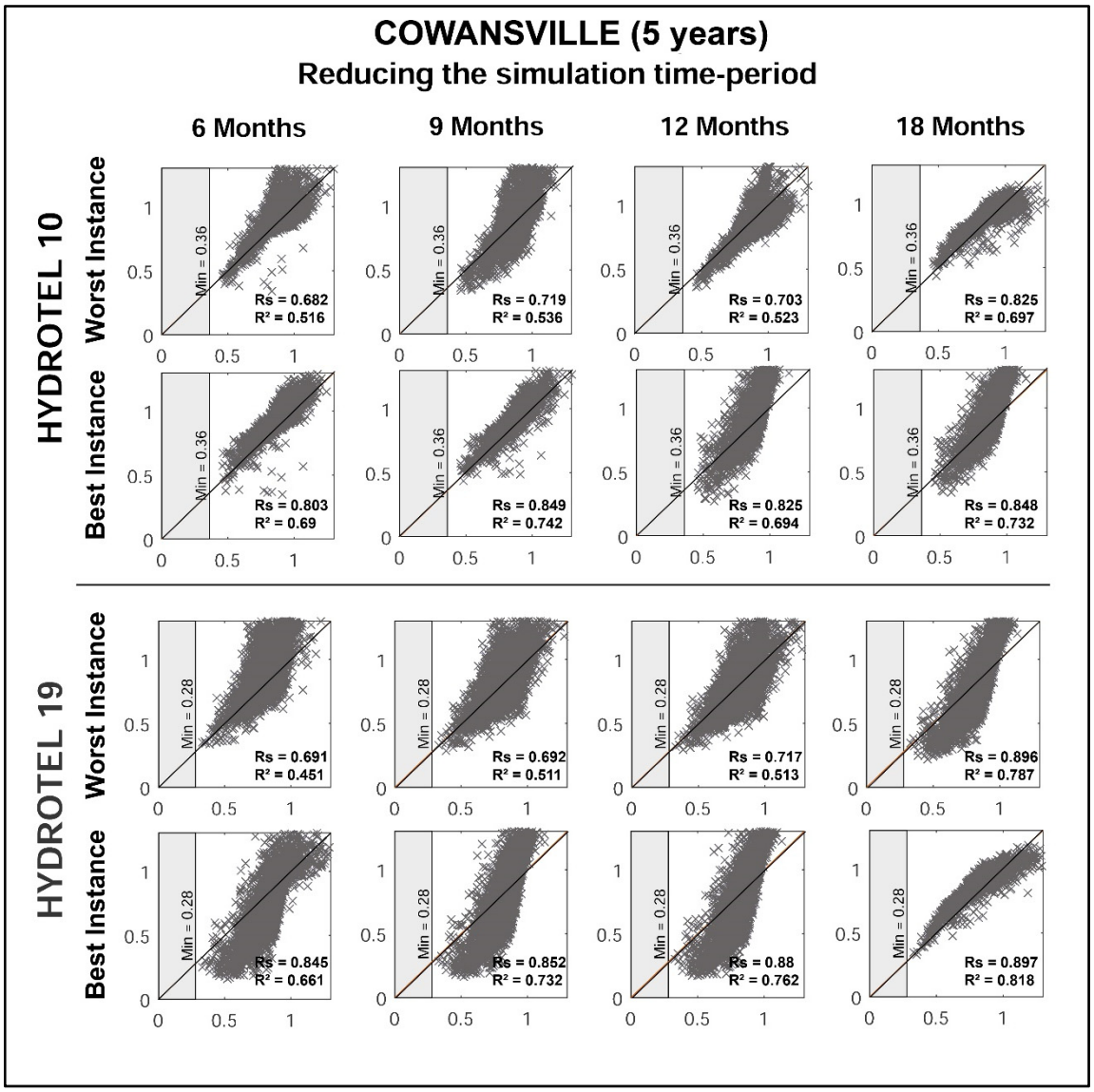


Figure-A III-2. $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Cowansville watershed comparing the surrogates having a reduced calibration time-period with the original models. Rs and R^2 coefficients are calculated for the representativeness evaluation

ANNEXE IV

RÉDUCTION DU NOMBRE DE UHRHS

Cette annexe présente les figures pour les bassins versants de Cowansville et Ceizur pour la mise en place de modèles substituts selon diverses réductions du nombre de UHRHs, complément de l'Article 2 (Chapitre 5).

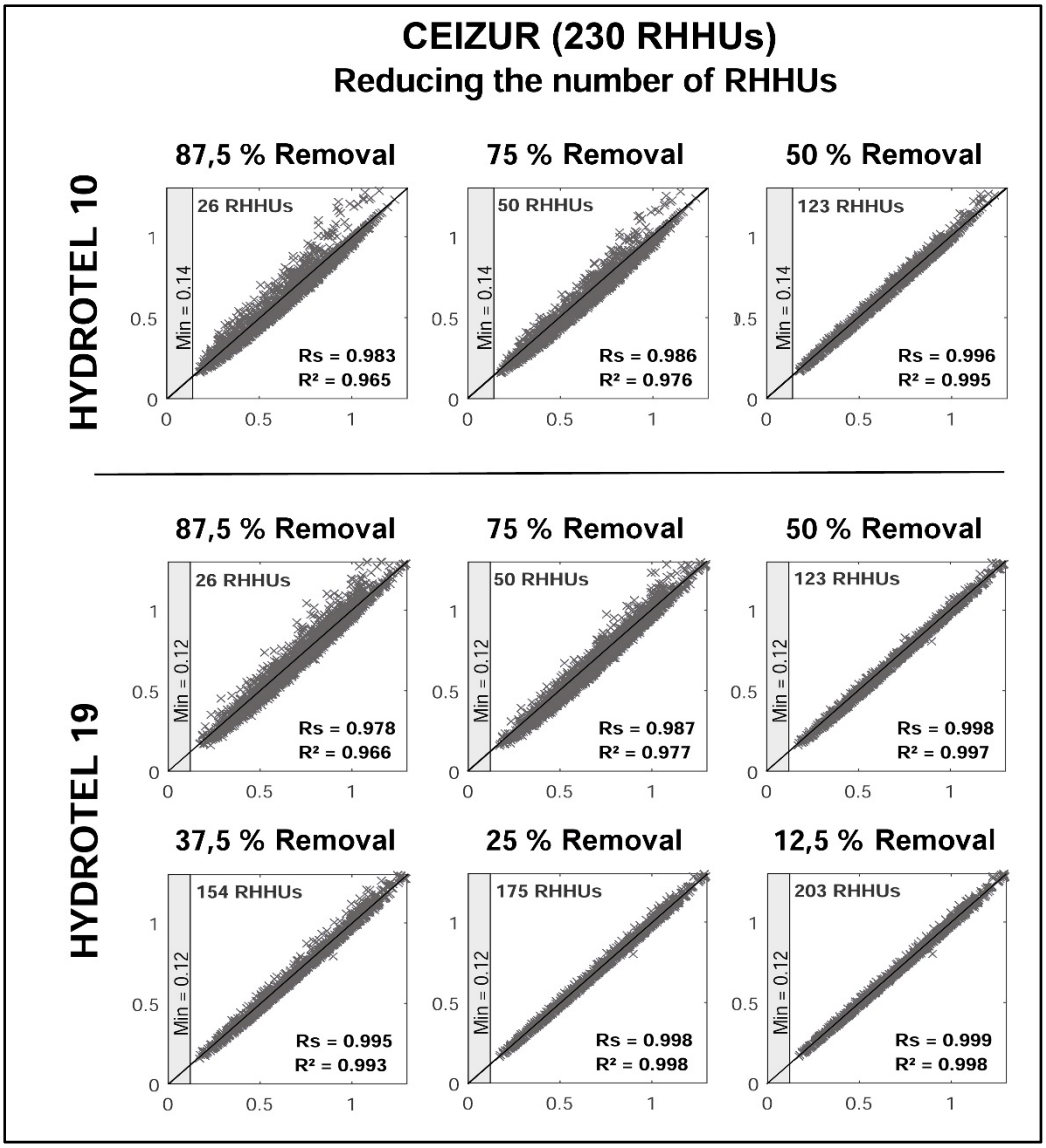


Figure-A IV-1. $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Ceizur watershed comparing the surrogates having a reduced number of RHHUs with the original models. Rs and R² coefficients are calculated for the representativeness evaluation

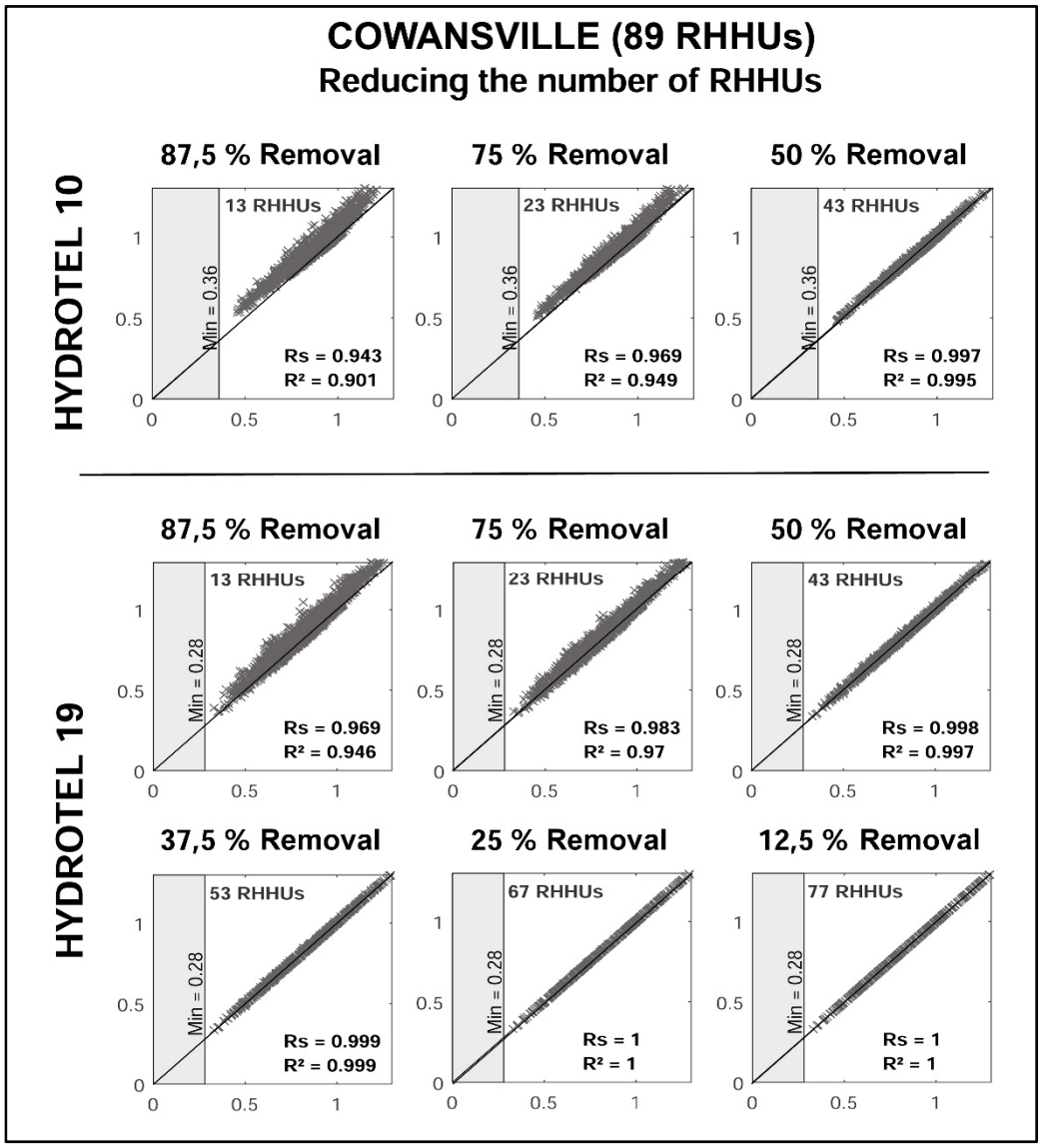


Figure-A IV-2. $1 - NSE$ simulation results obtained for HYDROTEL 10 and 19 problems on the Cowansville watershed comparing the surrogates having a reduced number of RHHUs with the original models. Rs and R² coefficients are calculated for the representativeness evaluation

ANNEXE V

FONCTIONS DE SURFACE (AUTRES FIGURES)

Cette annexe présente les figures pour les bassins versants de Cowansville et Toulnustouc pour les fonctions de surface du Chapitre 7.

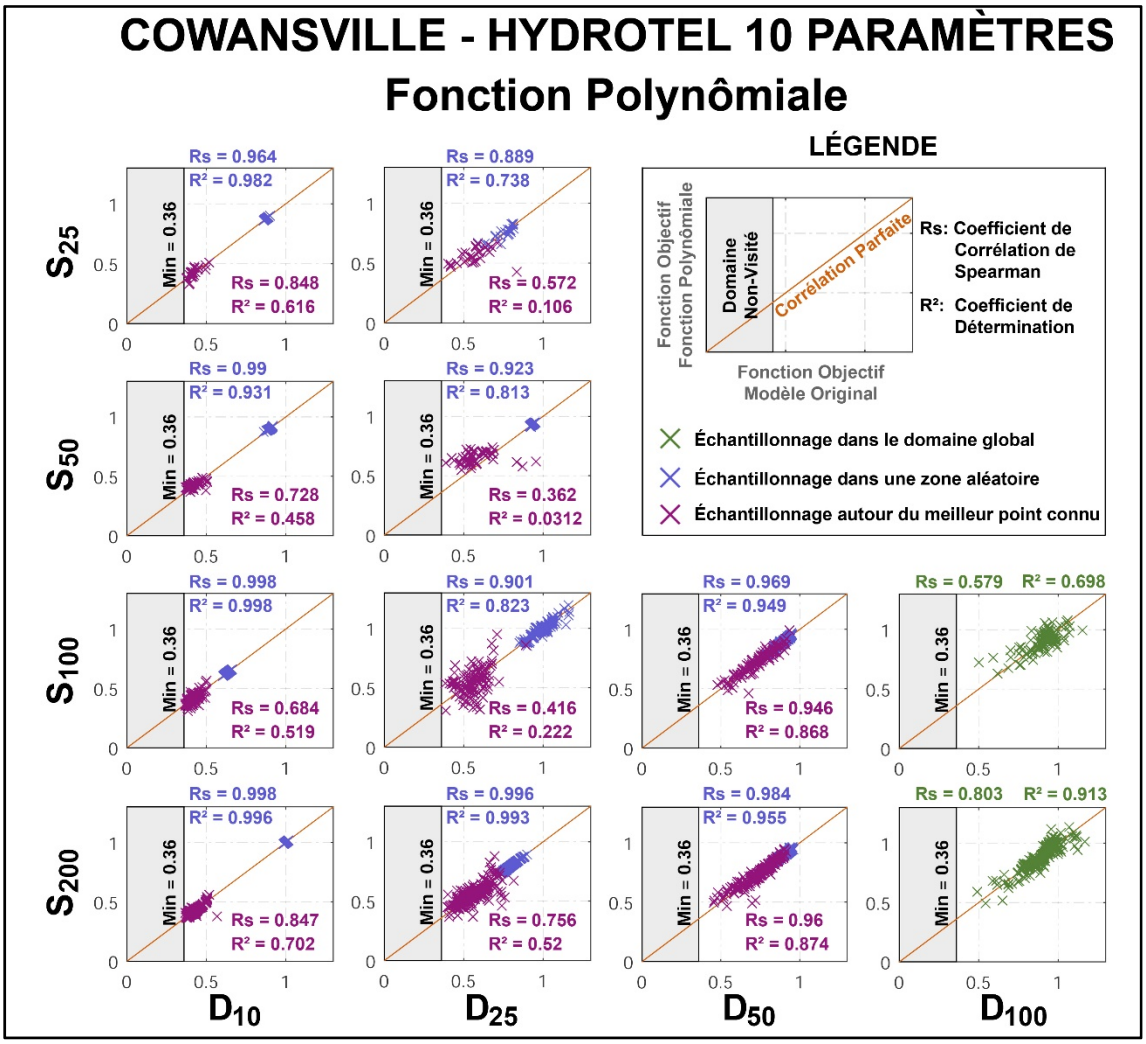


Figure-A V-1 Évaluation de la représentativité des fonctions polynomiales selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 10 pour le bassin versant Cowansville. Les coefficients R_s et R^2 sont présentés pour chacun des deux scénarios d'échantillonnage

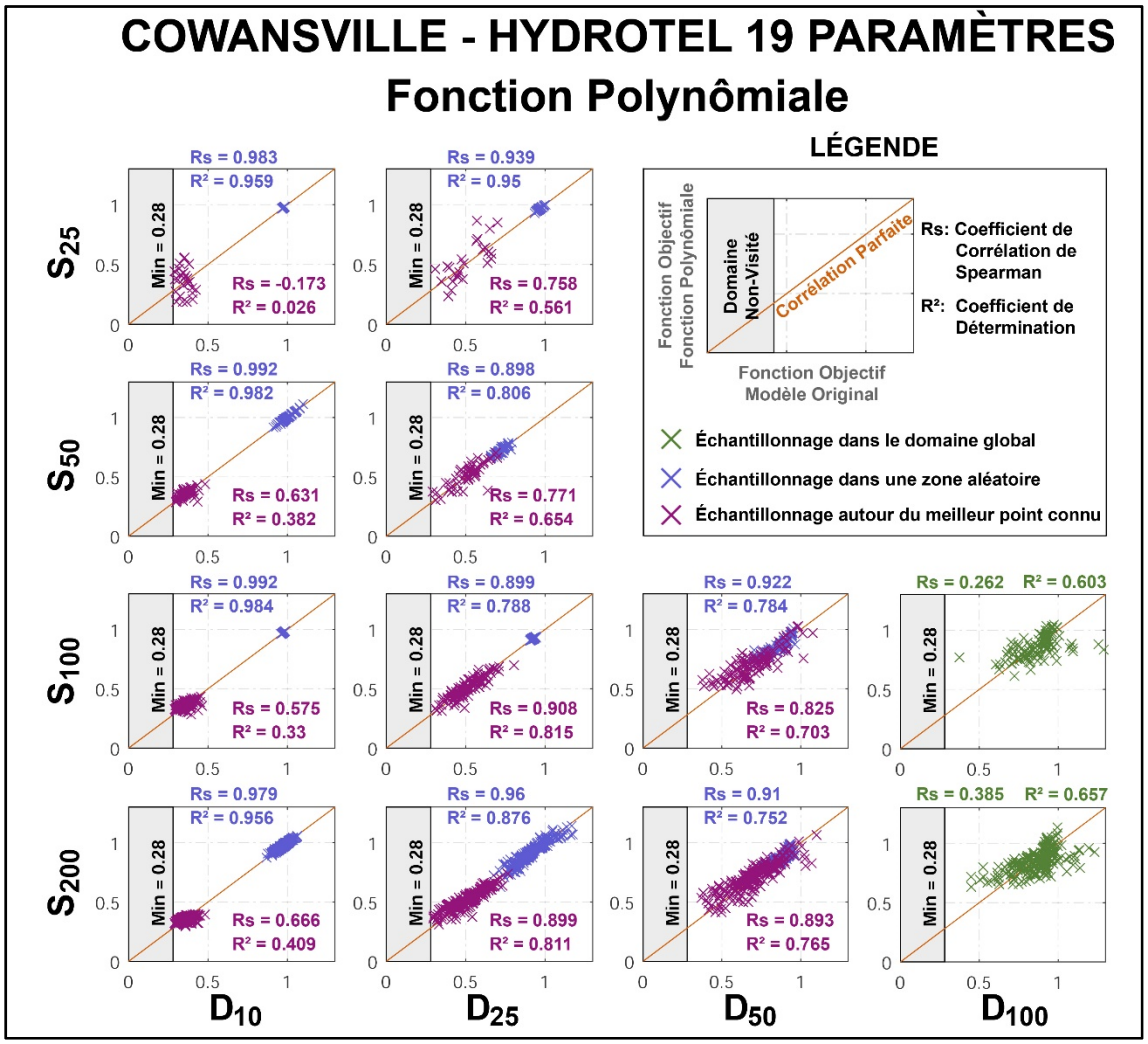


Figure-A V-2 Évaluation de la représentativité des fonctions polynomiales selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 19 pour le bassin versant Cowansville. Les coefficients R_s et R^2 sont présentés pour chacun des deux scénarios d'échantillonnage

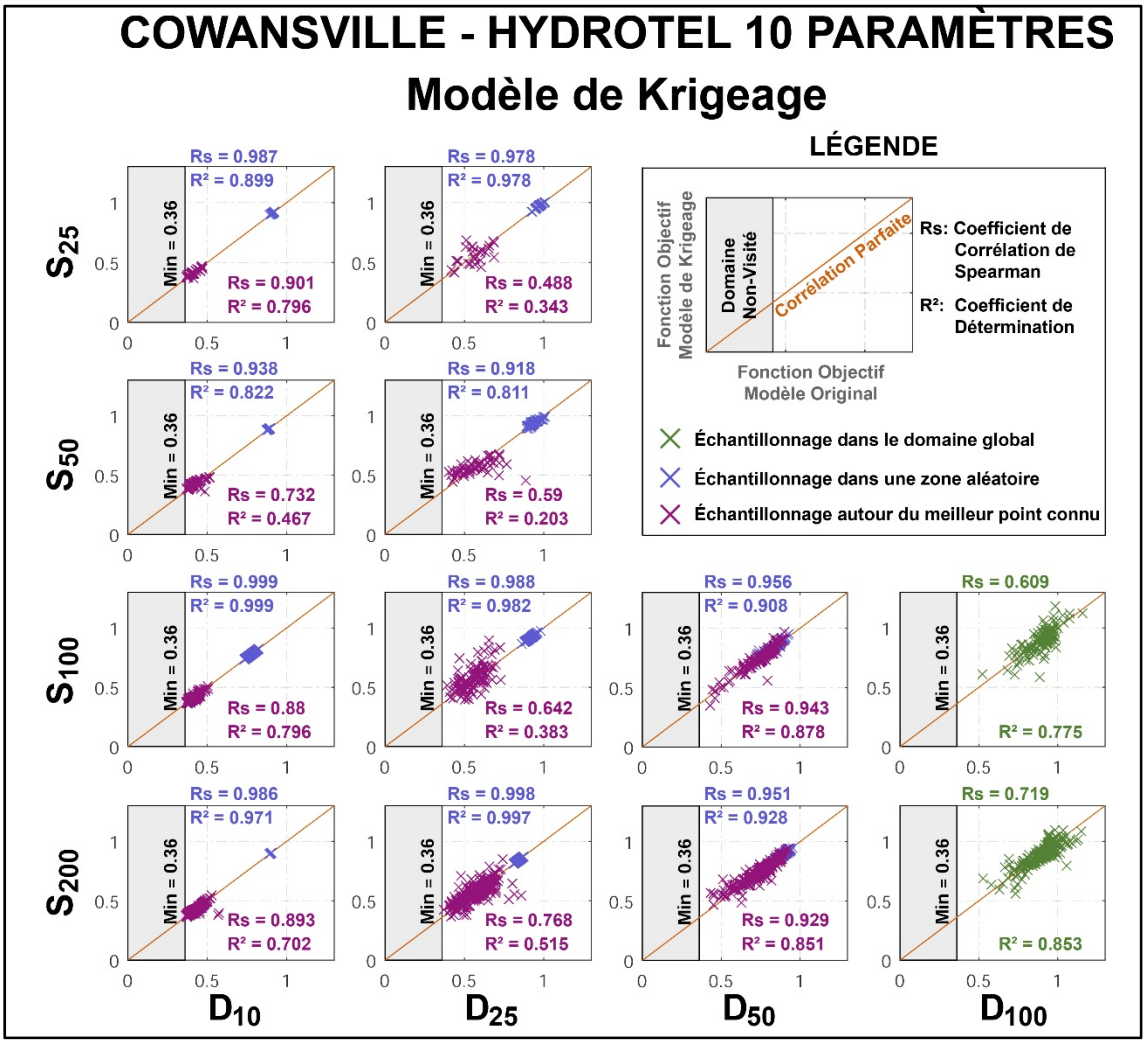


Figure-A V-3 Évaluation de la représentativité des modèles de krigeage selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 10 pour le bassin versant Cowansville. Les coefficients R_s et R^2 sont présentés pour chacun des deux scénarios d'échantillonnage

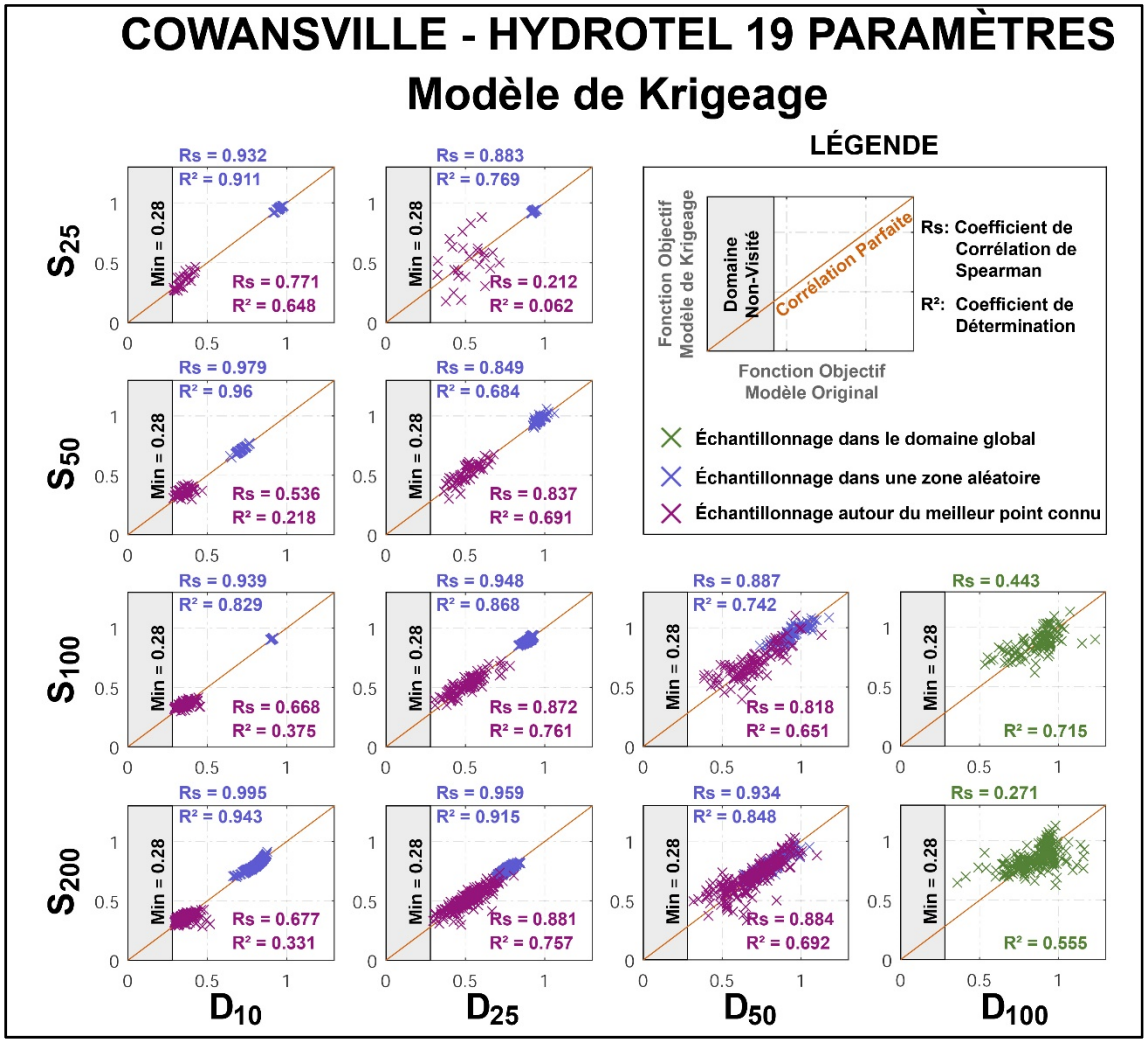


Figure-A V-4 Évaluation de la représentativité des modèles de krigeage selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 19 pour le bassin versant Cowansville. Les coefficients R_s et R^2 sont présentés pour chacun des deux scénarios d'échantillonnage

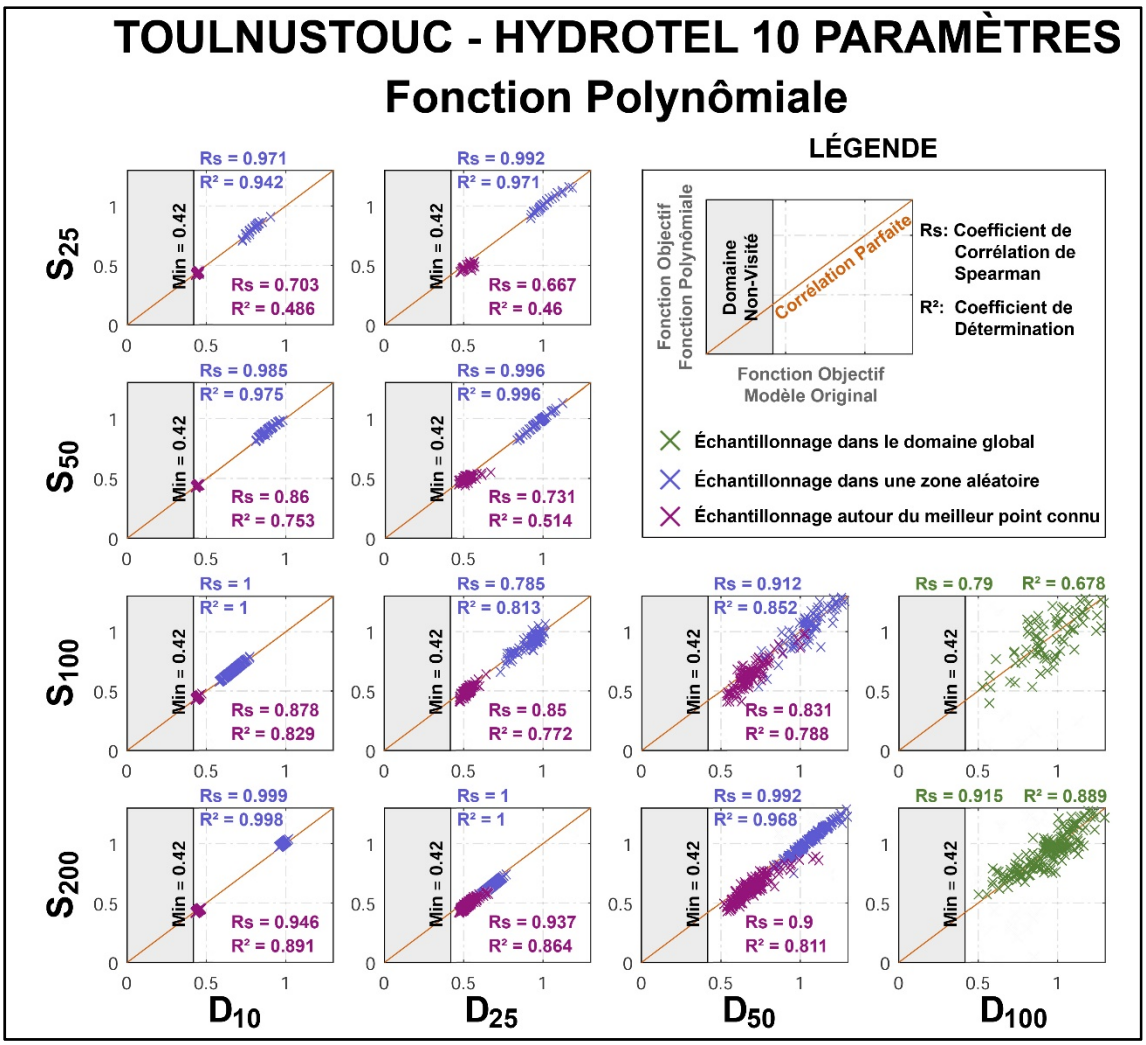


Figure-A V-5 Évaluation de la représentativité des fonctions polynomiales selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 10 pour le bassin versant Toulnustouc. Les coefficients R_s et R^2 sont présentés pour chacun des deux scénarios d'échantillonnage

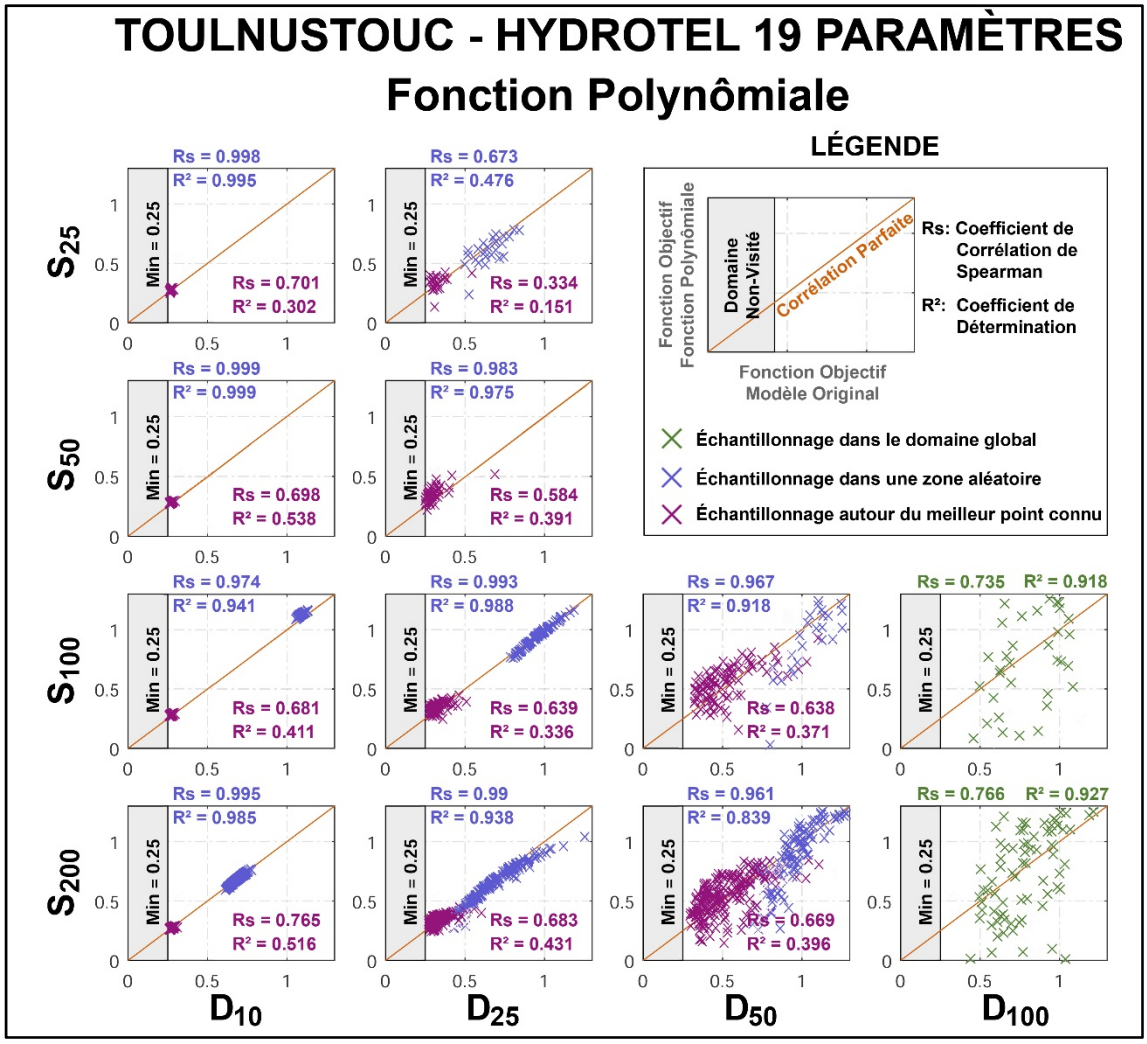


Figure-A V-6 Évaluation de la représentativité des fonctions polynomiales selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 19 pour le bassin versant Toulnustouc. Les coefficients R_s et R^2 sont présentés pour chacun des deux scénarios d'échantillonnage

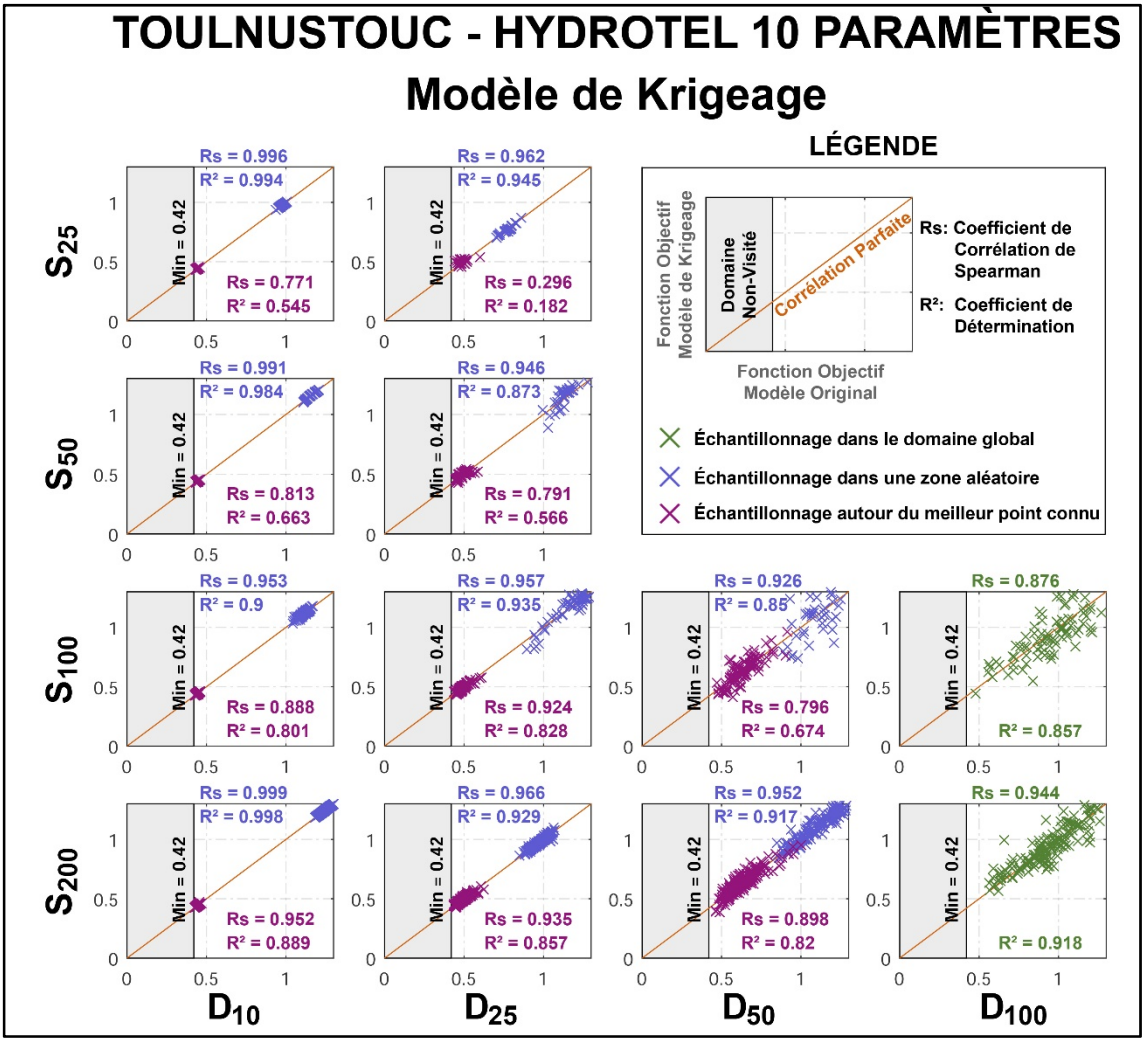


Figure-A I-7 Évaluation de la représentativité des modèles de krigeage selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 10 pour le bassin versant Toulnustouc. Les coefficients R_s et R^2 sont présentés pour chacun des deux scénarios d'échantillonnage

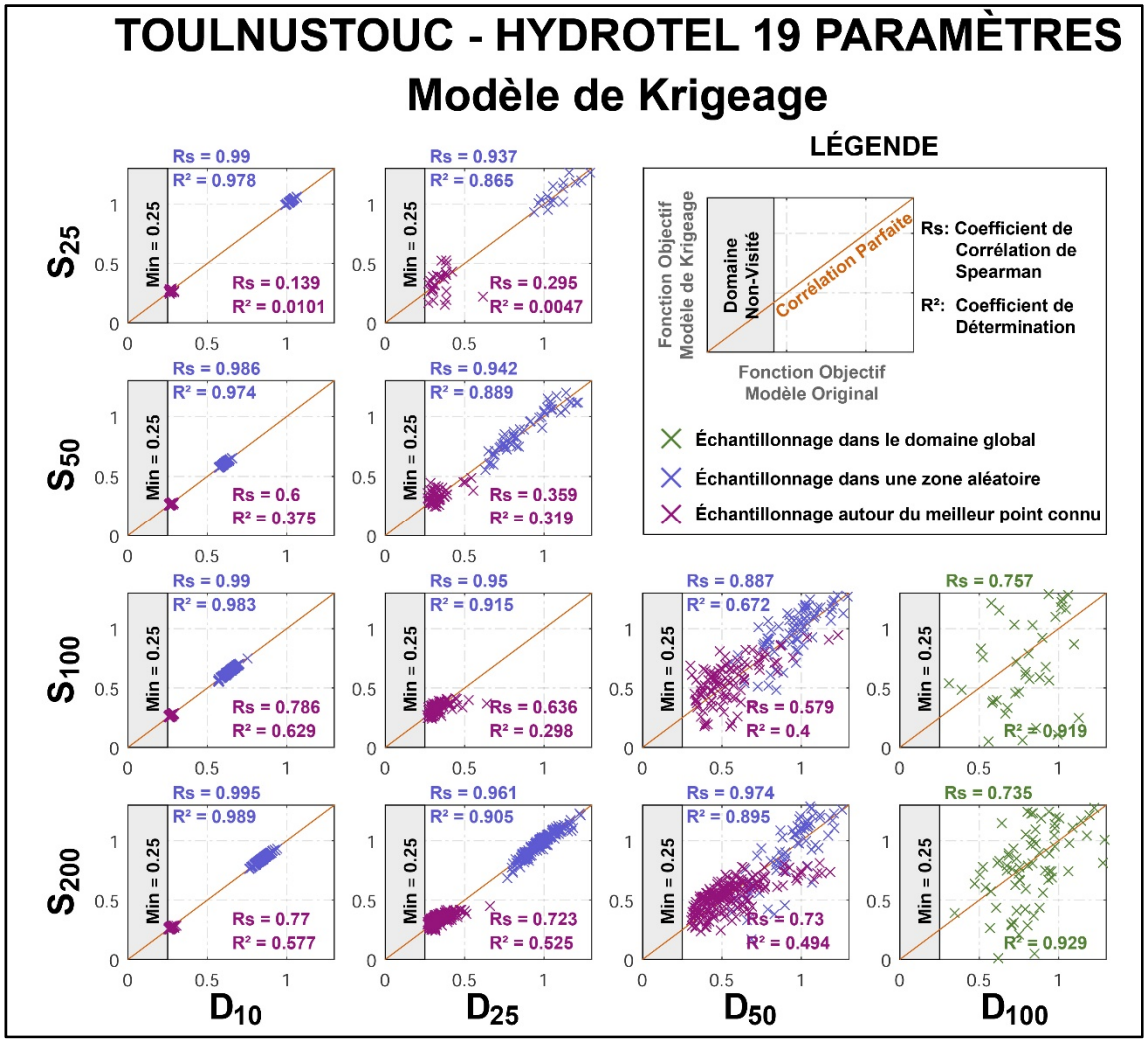


Figure-A I-8 Évaluation de la représentativité des modèles de krigeage selon les valeurs de la fonction objectif $1 - NSE$ obtenues avec HYDROTEL 19 pour le bassin versant Toulnustouc. Les coefficients R_s et R^2 sont présentés pour chacun des deux scénarios d'échantillonnage

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

Abramson, M.A., Audet, C. (2006). « Convergence of mesh adaptive direct search to second-order stationary points ». *SIAM Journal on Optimization*, Volume 17, n°2, p. 606-619. doi : 10.1137/050638382.

Abramson, M.A., Audet, C., Dennis Jr., J.E., Le Digabel, S. (2009). « A deterministic MADS instance with orthogonal directions ». *SIAM Journal on Optimization*, Volume 20 n°2, p. 948-966.

Anderson, E. (2002). « Calibration of Conceptual Hydrologic Models for Use in River Forecasting [online] ». The National Oceanic and Atmospheric Administration, Available from: http://www.nws.noaa.gov/oh/hrl/calb/calibration1102/main.htm.

Arsenault, R., Poulin, A., Côté, P., Brissette, F. (2014). « Comparison of Stochastic Optimization Algorithms in Hydrological Model Calibration ». *Journal of Hydrologic Engineering*, Volume 19 n°7, p. 1374-1384.
doi : 10.1061/(ASCE)HE.1943-5584.0000938.

Arsenault, R., Brissette, F., Martel, J.-L. (2018). « The hazards of split-sample validation in hydrological model calibration ». *Journal of hydrology*, Volume 566, p. 346-362.
doi : 10.1016/j.jhydrol.2018.09.027

Asadzadeh, M., Tolson, B.A. (2013). « Pareto Archived Dynamically Dimensioned Search with Hypervolume Based Selection for Multi-Objective Optimization ». *Engineering Optimization*, Volume 45 n°12, p. 1489-1509.

Audet, C., Dennis Jr., J.E. (2006). « Mesh adaptive direct search algorithms for constrained optimisation ». *SIAM Journal on Optimization*, Volume 17, n°1, p. 188-217.
doi : 10.1137/040603371.

Audet, C., Orban, D. (2006). « Finding Optimal Algorithmic Parameters Using Derivative-Free Optimization ». *SIAM Journal on Optimization*, Volume 17 n°3, p. 642-664.
doi : 10.1137/040620886.

Audet, C., Dennis Jr., J.E., Le Digabel, S. (2008). « Parallel Space Decomposition of the Mesh Adaptive Direct Search algorithm ». *SIAM Journal on Optimization*, Volume 10 n°3, p. 1150-1170. doi : 10.1137/070707518.

Audet, C., Dennis Jr., J.E. (2009). « A progressive barrier for derivative-free nonlinear programming ». *Siam Journal of Optimization*, Volume 20 n°1, p. 455-472.
doi : 10.1137/070692662.

Audet, C., Savard, G., Zghal, W. (2010). « A mesh adaptive direct search algorithm for multiobjective optimization ». *European Journal of Operational Research*, Volume 204 n°3, p. 545-556.

Audet, C., Dennis Jr., J.E., Le Digabel, S. (2012). « Trade-off studies in blackbox optimization ». *Optimization Methods and Software*, Volume 27, n°4-5, p. 613-624. doi : 10.1080/10556788.2011.571687.

Audet, C., Ianni, A., Le Digabel, S., Tribes, C. (2014). « Reducing the Number of Function Evaluations in Mesh Adaptive Direct Search Algorithms ». *SIAM Journal on Optimization*, Volume 24, n°2, p. 621-642.

Audet, C. (2014). « A survey on direct search methods for blackbox optimization and their applications ». In Mathematics without boundaries: Surveys in interdisciplinary research (chapter 2), sous la dir. de Pardalos, P.M., Rassias, T.M., 648p. Montréal (Canada) : Springer.

Audet, C., Hare, W. (2017). *Derivative-Free and Blackbox Optimization*. Publié par : Springer International Publishing, *Springer Series in Operations Research and Financial Engineering*. 302 p. doi: 10.1007/978-3-319-68913-5.

Audet, C., Ihaddadene, A., Le Digabel, S., Tribes, C. (2018). « Robust optimization of noisy blackbox problems using the Mesh Adaptive Direct Search algorithm ». *Optimization Letters*, Volume 12 n°4, p.675-689. doi: 10.1007/s11590-017-1226-6.

Bajamgnigni Gbambie, A.S., Poulin, A., Boucher, M.-A., Arsenault, R. (2017). « Added Value of Alternative Information in Interpolated Precipitation Datasets for Hydrology ». *Journal of Hydrometeorology*, Volume 18 n°1. doi : 10.1175/JHM-D-16-0032.1.

Beven, K.J. (2001). *Rainfall-Runoff Modelling: The Primer*. Chichester: Wiley, United Kingdom. 488 p.

Booker, A.J., Dennis, Jr., J.E., Frank, P.D., Serafini, D.B., Torczon, V., Trosset, M.W. (1999). « A rigorous framework for optimization of expensive functions by surrogates ». *Structural and Multidisciplinary Optimization*, Volume 17 n°1, p. 1-13.

Bouda, M., Rousseau, A.N., Konan, B., Gagnon, P., Gumiere, S.J. (2012). « Case study: Bayesian uncertainty analysis of the distributed hydrological model HYDROTEL ». *Journal of Hydrologic Engineering*, Volume 17 n°9, p. 1021-1032.

Bouda, M., Rousseau, A.N., Gumiere, S.J., Gagnon, P., Konan, B., Moussa, R. (2014). « Implementation of an automatic calibration procedure for HYDROTEL based on prior OAT sensitivity and complementary identifiability analysis ». *Hydrological Processes*, Volume 28 n°12, p. 3947-3961.

Boukouvala, F., Misener, R., Floudas, C.A. (2016). « Global optimization advances in mixed-integer nonlinear programming, MINLP, and constrained derivative-free optimization, CDFO ». *European Journal of Operational Research*, Volume 252 n°3, p. 701-727.

Brunetti, G., Simunek, J., Turco, M., Piro., P. (2017). « On the use of surrogate-based modeling for the numerical analysis of Low Impact Development techniques ». *Journal of Hydrology*, Volume 548, p.263-277. doi : 10.1016/j.jhydrol.2017.03.013.

Castelleti, A., Pianosi, F., Soncini-Sessa, R., Antenucci, J.P. (2010). « A multiobjective response surface approach for improved water quality planning in lakes and reservoirs ». *Water Resources Research*, Volume 46, W06502. doi : 10.1029/2009WR008389.

Castelleti, A., Galelli, S., Ratto, M., Soncini-Sessa, R., Young, P.C. (2012). « A general framework for Dynamic Emulation Modelling in environmental problems ». *Environmental Modelling & Software*, Volume 34, p. 5-18. doi : 10.1016/j.envsoft.2012.01.002.

Conn, A.R., Scheinberg, K., Vicente, L.N. (2009). *Introduction to Derivative-Free Optimization*, MPS-SIAM Book Series on Optimization: SIAM (Philidelphia), 277 p.

Conn, A.R., Le Digabel, S. (2013). « Use of quadratic models with mesh-adaptive direct search for constrained black box optimization ». *Optimization Methods and Software*, Volume 28 n°1, p. 139-158. doi : 10.1080/10556788.2011.623162.

Dan Lu, M.X., Gui, D., Qi, Z., Zhang, G. (2017). « Calibration of an agricultural-hydrological model (RZWQM2) using surrogate global optimization ». *Journal of Hydrology*, Volume 544, p. 456-466. doi : 10.1016/j.jhydrol.2016.11.051.

Das, T., Bardossy, A., Zehe, E., He, Y. (2008). « Comparaison of conceptual model performance using different representations of spatial variability ». *Journal of Hydrology*, Volume 356, p. 106-118. doi : 10.1016/j.jhydrol.2008.04.008.

Duan, Q.Y., Gupta, V.K., Sorooshian, S. (1993). « Shuffled complex evolution approach for effective and efficient global minimization ». *Journal of Optimization Theory and Applications*, Volume 76, n°3, p. 501-521.

Duan, Q.Y., Sorooshian, S., Gupta, V.K. (1994). « Optimal use of the SCE-UA global optimization method for calibrating watershed models ». *Journal of Hydrology,* Volume 158, n°3-4, p. 265-284. doi : 10.1016/0022-1694(94)90057-4.

Eberhart, R.C., Kennedy, J. (1995). « A new optimizer using particle swarm theory ». Présenté dans la conférence : *Sixth International Symposium on Micro Machine and Human Science (IEEE)*, Nagoya, Japan. p. 39-43.

Fen, C.S., Chan, C.C., Cheng, H.C. (2009). « Assessing a response surface-based optimization approach for soil vapor extraction system design ». *Journal of Water Resources Planning and Management*, Volume 135 n°3, p. 198-207. doi : 10.1061/(ASCE)0733-9496(2009)135:3(198).

Forrester, A.I.J., Keane, A.J. (2009). « Recent advances in surrogate-based optimization ». *Progress in Aerospace Sciences*, Volume 45 n°1, p. 50-79. doi : 10.1016/j.paerosci.2008.11.001.

Fieller, E.C., Hartley, H.O., Pearson, E.S. (1957). « Test for rank correlation coefficients. I ». *Biometrika*, Volume 44 n°3-4, p. 470-481. doi : 10.1093/biomet/44.3-4.470.

Fortin, J.P., Turcotte, R., Massicotte, S., Moussa, R., Fiztback, J. (2001a). « A Distributed Watershed Model Compatible with Remote Sensing and GIS Data, Part 1: Description of the model ». *Journal of Hydrologic Engineering, American Society of Civil Engineering*, Volume 6, n°2, p. 91-99.

Fortin, J.P., Turcotte, R., Massicotte, S., Moussa, R., Fiztback, J. (2001b). « A Distributed Watershed Model Compatible with Remote Sensing and GIS Data, Part 2: Application to the Chaudière watershed ». *Journal of Hydrologic Engineering, American Society of Civil Engineering*, Volume 6, n°2, p. 100-108.

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F. (2009). « Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling ». *Journal of Hydrology*, Volume 377 n°1-2, p. 80-91. doi : 10.1016/j.jhydrol.2009.08.003.

Haghnegahdar, A., Tolson, B.A., Craig, J.R., Paya, K.T. (2015). « Assessing the performance of a semi-distributed hydrological model under various watershed discretization schemes ». *Hydrological Processes*, Volume 29, p. 4018-4031. doi : 10.1002/hyp.10550.

Hansen, N. (2006). « The CMA evolution strategy: a comparing review ». In: Towards a New Evolutionary Computation: Advances on Estimation of Distribution Algorithms, Édité par : Jose A. Lozano, Pedro Larranaga, Inaki Inza et Endika Bengoetxea, p. 75-102. Berlin (Germany): Springer-Heidelberg.

Huang, X., Liao, W., Lei, X., Jia,Y., Wang, Y., Wang, X., Jiang, Y., Wang, H. (2014). « Parameter optimization of distributed hydrological model with a modified dynamically dimensioned search algorithm ». *Environmental Modelling & Software*, Volume 52, p. 98-110.

Hussain, M.F., Barton, R.R., Joshi, S.B. (2002). « Metamodeling: Radial basis functions, versus polynomials ». *European Journal of Operational Research*, Volume 138 n°1, p. 142-154. doi : 10.1016/S0377-2217(01)00076-5.

Huot, Pierre-Luc (2014). *Évaluation de méthodes d'optimisation pour le calage efficace de modèles hydrologiques coûteux en temps de calcul*. (Mémoire, École de technologie supérieure, Montréal, Canada).

Huot, P.-L., Poulin, A., Alarie, S. (2014). « Assessment of blackbox optimization methods for efficient calibration of computationally intensive hydrological models ». In: CUNY Academic works, New-York City, USA.

Huot, P.-L., Poulin, A., Audet, C., Alarie, S. (2017). « A Hybrid Optimization Approach for Efficient Calibration of Computationally Intensive Hydrological Models ». *Les Cahiers du GERAD*, G-2017-56, 20 p.

Ingber, L. (1993). « Simulated annealing: Practice versus theory ». *Mathematical and Computer Modelling*, Volume 18 n°1, p. 29-57.

Krige, D.G. (1951). « A statistical approach to some mine valuation and allied problems on the Witwatersrand ». Master's thesis of the University of Witwatersrand, 272 p.

Khu, S.T., Savic, D., Liu, Y., Madsen, H. (2004). « A fast Evolutionary-based Meta-Modelling Approach for the Calibration of a Rainfall-Runoff Model ». Presenté à la conférence : *The First Biennial Meeting of the International Environmental Modelling Software Society*, Osnabruck, Germany.

Koziel, S., Leifsson, L. (2013). « Multi-level CFD-based Airfoil Shape Optimization with Automated Low-fidelity Model Selection ». *Procedia Computer Science*, Volume 18, p. 889-898. doi : 10.106/j.procs.2013.05.254.

Le Digabel, S. (2011). « Algorithm 909: NOMAD: Nonlinear Optimization with the MADS algorithm ». *ACM Transactions on Mathematical Software*, Volume 37, n°4, 44:1-44:15. doi : 10.1145/1916461.1916468.

Lei, X.H., Wang, Y.H., Liao, W.H., Jiang, Y.Z., Tian, Y., Wang, H. (2011). « Development of efficient and cost-effective distributed hydrological modeling tool MWEasyDHM based on open-source MapWindow GIS ». *Computer & Geosciences*, Volume 37 n°9, p. 1476-1489. doi : 10.1016/j.cageo.2011.03.016.

Leifsson, L., Koziel, S., Bekasiewicz, A. (2014). « Fast Low-Fidelity Wing Aerodynamics Model for Surrogate-Based Shape Optimization ». *Procedia Computer Science*, Volume 29, p. 811-820. doi : 10.1016/j.procs.2014.05.073.

Leifsson, L., Koziel, S. (2015). *Simulation-Driven Aerodynamic Design Using Variable-Fidelity Models*, London (UK): Imperial College Press, 444 p.

Lophaven, S.N., Nielsen, H.B., Sondergaard, J. (2002). DACE - A Matlab Kriging Toolbox, Version 2.0. Report IMM-REP-2002-02, Informatics and mathematical Modelling, DTU. 36 p.

Madsen, H. (2003). « Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives ». *Advances in Water Resources*, Volume 26, n°2, p. 205-216. doi : 10.1016/S0309-1708(02)00092-1.

Maier, H.R., Kapelan, Z., Kasprzyk, J., Kollat, J., Matott, L.S., Cunha, M.C., Dandy, G.C., Gibbs, M.S., Keedwell, E., Marchi, A., Ostfeld, A., Savic, D., Solomatine, D.P., Vrugt, J.A., Zecchin, A.C., Minsker, B.S., Barbour, E.J., Kuczera, G., Pasha, F., Castelletti, A., Giuliani, M., Reed, P.M. (2014). « Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions ». *Environmental Modelling & Software*, Volume 62, p. 271-299.

Marsden, A.L., Wang, M., Dennis Jr., J.E., Moin, P. (2004). « Optimal Aeroacoustic Shape Design Using the Surrogate Management Framework ». *Optimization and Engineering*, Volume 5 n°2, p. 235-262.
doi : 10.1023/B:OPTE.0000033376.89159.65.

McKay, M.D., Backman, R.J., Conover, W.J. (1979). « A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code ». *Technometrics*, Volume 21, n°2, p. 239-245.

Ministère de l'Environnement et de la Lutte contre les Changements Climatiques (MELCC) (2018). « Prévisions hydrologiques et hydrauliques ». Disponible au : www.cehq.gouv.ac.ca/prevision/. (accédé le : 26-02-2018).

Ministère de l'Environnement et de la Lutte contre les Changements Climatiques (MELCC) (2019). « Gestion du domaine hydrique de l'État ». Available at: http://www.cehq.gouv.qc.ca/Domaine-hydrique/gestion. (accédé le : 10-04-2019).

Minville, M., Cartier, D., Guay, C., Leclaire, L.-A., Audet, C., Le Digabel, S., Merleau, J. (2014). « Improving process representation in conceptual hydrological model calibration using climate simulations ». *Water Resources Research*, Volume 50, n°6, p. 5044-5073. doi : 10.1002/2013WR013857.

Moradkhani, H., Sorooshian, S. (2008). « General Review of Rainfall-Runoff Modeling: Model Calibration, Data Assimilation, and Uncertainty Analysis ». In: *Hydrological Modeling and Water Cycle: Coupling of the Atmospheric and Hydrological Models*, sous la dir. de Sorooshian Soroosh, Hsu Kuo-Lin, Coppola Erika, Tomassetti Barbara,

Verdecchia Marco et Visconti Guido, p. 1-23. « Water Science and Technology Library », Volume 63. Berlin (Heidelberg): Springer.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L.n Harmel, R.D., Veith, T.L. (2007). « Model evaluation guidelines for systematic quantification of accuracy in watershed simulations ». *Transactions of the ASABE*, Volume 50, n°3, p. 885-900.

Morris, M.D. (1991). « Factorial Sampling Plans for Preliminary Computational Experiments ». *Technometrics*, Volume 33, n°2, p. 161-174. doi : 10.2307/1269043.

Mousavi, S.J., Shourian, M. (2010). « Adaptive sequentially space-filling metamodeling applied in optimal water quantity allocation at basin scale ». *Water Resources Research*, Volume 46, W03520. doi : 10.1029/2008WR007076.

Mugunthan, P., Shoemaker, C.A., Regis, R.G. (2005). « Comparison of function approximation, heuristic, and derivative-based methods for automatic calibration of computationally expensive groundwater bioremediation models ». *Water Resources Research*, Volume 41, n°11, W11427. DOI: 10.1029/2005WR004134.

Nash, J.E., Sutcliffe, J.V. (1970). « River flow forecasting through conceptual models part I - A discussion of principles ». *Journal of Hydrology*, Volume 10 n°3, p. 282-290.

Neelin, J.D., Bracco, A., Luo, H., McWilliams, J.C., Meyerson, J.E. (2010). « Considerations for parameter optimization and sensitivity in climate models ». *Proceedings of the National Academy of Sciences of the United States of America*, Volume 107 n°50, p. 21349-21354. doi : 10.1073/pnas.1015473107.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., King, K.W. (2002). « Soil and water assessment tool theoretical documentation. Version 2000 ». Édité par : *Agricultural Research Service*, U.S. Department of Agricultural. Temple, Texas. 506 p.

Nelder, J., Mead, R. (1965). « A simplex method for function minimization ». *Computer Journal*, Volume 7, n°4. 308-313.

Papadrakakis, M., Lagaros, N.D., Tsompanakis, Y. (1998). « Structural optimization using evolution strategies and neural networks ». *Computer Methods in Applied Mechanics and Engineering*, Volume 156 n°1-4, p. 309-333.
doi : 10.1016/S0045-7825(97)00215-6.

Pechlivanidis, I.G., Jackson, B., McIntyre, N., Wheater, H.S. (2011). « Catchment scale hydrological modelling: A review of model types, calibration approaches and uncertainty analyssi methods in the context of recent developments in technology and applications ». *Global NEST Journal,* Volume 13, n°3, p. 193-214.

Poissant, D., Arsenault, R., Brissette, F. (2017). « Impact of parameter set dimensionality and calibration procedures on streamflow prediction at ungauged catchments ». *Journal of Hydrology: Regional Studies*, Volume 12, p. 220-237. doi: 10.1016/j.ejrh.2017.05.005.

Polak, E., Wetter, M. (2006). « Precision Control for Generalized Pattern Search Algorithms with Adaptive Precision Function Evaluations ». *SIAM Journal on Optimization*, Volume 16 n°3, p. 650-669. doi : 10.1137/040605527.

Poulin, A., Brissette, F., Leconte, R., Arsenault, R., Malo, J.S. (2011). « Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river bassin ». *Journal of Hydrology*, Volume 409, n°3-4, p. 626-636. doi : 10.1016/j.jhydrol.2011.08.057.

Pushpalatha, R., Perrin, C., Le Moine, N., Andréassian, V. (2012). « A review of efficiency criteria suitable for evaluating low-flow simulations ». *Journal of Hydrology*, Volume 420-421, p. 171-182. DOI: 10.1016/j.jhydrol.2011.11.055.

Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R., Tucker, P.K. (2005). « Surrogate-based analysis and optimization ». *Progress in Aerospace Sciences*, Volume 41 n°1, p. 1-28. doi : 10.1016/j.paerosci.2005.02.001.

Rardin, R.L., Uzsoy, R. (2001). « Experimental Evaluation of Heuristic Optimization Algorithms: A Tutorial ». *Journal of Heuristics*, Volume **7**, p. 261-304.

Razavi, S., Tolson, B.A., Matott, L.S., Thomson, N.R., Maclean, A., Seglenieks, F.R. (2010). « Reducing the computationally cost of automatic calibration through model pre-emption ». *Water Resources Research*, Volume 46, W11523. doi : 10.1029/2009WR008957.

Razavi, S., Tolson, B.A., Burn, D.H. (2012a). « Review of surrogate modeling in water resources ». *Water Resources Research*, Volume 48, W07401. doi : 10.1029/2011WR011527.

Razavi, S., Tolson, B. A., Burn, D.H. (2012b). « Numerical assessment of metamodelling strategies in computationally intensive optimization ». *Environmental Modelling & Software*, Volume 34, p. 67-86. doi : 10.1016/j.envsoft.2011.09.010.

Razavi, S., Tolson, B.A. (2013). « An efficient framework for hydrologic model calibration on long data periods ». *Water Resources Research*, Volume 49, p. 8418-8431. doi : 10.1002/2012WR013442.

Regis, R.G., Shoemaker, C.A. (2004). « Local function approximation in evolutionary algorithms for the optimization of costly functions ». *IEEE Transaction on Evolutionary Computation*, Volume 8 n°5, p. 490-505. doi : 10.1109/TEVC.2004.835247.

Regis, R.G., Shoemaker, C. (2007). « A Stochastic Radial Basis Function Method for the Global Optimization of Expensive Functions ». *INFORMS Journal on Computing*, Volume 19 n°4, p. 497-509. doi : 10.1287/ijoc.1060.0182.

Regis, R.G., Shoemaker, C.A. (2009). « Parallel stochastic global optimization using radial basis functions ». *INFORMS - Journal on Computing*, Volume 21 n°3, p. 411-426. doi : 10.1287/ijoc.1090.0325.

Regis, R.G., Shoemaker, C.A. (2013). « A quasi-multistart framework for global optimization of expensive functions using response surface models ». *Journal of Global Optimization*, Volume 56 n°4, p.1719-1753. doi : 10.1007/s10898-012-9940-1.

Ricard, S., Bourdillon, R., Roussel, D., Turcotte, R. (2012). « Global Calibration of Distributed Hydrological Models for Large-Scale Applications ». *Journal of Hydrologic Engineering*, Volume 18 n°6. doi: 10.1061/(ASCE)HE.1943-5584.0000665

Rios, L.M., Sahinidis, N.V. (2013). « Derivative-free optimization: a review of algorithms and comparison of software implementations ». *Journal of Global Optimization*, Volume 56, n°3, p. 1247-1293. doi : 10.1007/s10898-012-9951-y.

Roos, G.N. (2009). « Development of the Dipole Flow and Reactive Tracer Test (DFRTT) for aquifer parameter estimation ». Publié par : *Civil and Environmental Engineering*, University of Waterloo. Waterloo, Ontario, Canada.

Servat, E., Dezetter, A. (1991). « Selection of calibration objective functions in the context of rainfall-runoff modelling in a sudanese savannah area ». *Hydrological Sciences Journal*, Volume 36 n°4, p. 307-300. doi : 10.1080/02626669109492517.

Shan, S., Wang, G.G. (2010). « Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions ». *Structural and Multidisciplinary Optimization*, Volume 41, n°2, p. 219-241. doi : 10.1007/s00158-009-0420-2.

Shoemaker, C.A., Regis, R.G., Fleming, R.C. (2007). « Watershed calibration using multistart local optimization and evolutionary optimization with radial basis functions approximation ». *Hydrological Sciences Journal*, Volume 52 n°3, p. 450-465.

Singh, V., Woolhiser, D. (2002). « Mathematical Modeling of Watershed Hydrology ». *Journal of Hydrology Engineering*, Volume 7, n°4, p. 270-292. doi : 10.1061/(ASCE)1084-0699(2002)7:4(270).

Talgorn, B., Audet, C., Le Digabel, S., Kokkolaras, M. (2017). « Locally weighted regression models for surrogate-assisted design optimization ». *Optimization and Engineering*, Volume 19 n°1, p. 213-238. doi : 10.1007/s11081-017-9370-5.

Toal, D.J.J., (2015). « Some considerations regarding the use of multi-fidelity Kriging in the construction of surrogate models ». *Structural and Multidisciplinary Optimization*, Volume 51 n°6, p. 1223-1245. doi : 10.1007/s00158-014-1209-5.

Tolson, B.A., Shoemaker, C.A. (2007). « Dynamically dimensioned search algorithm for computationally efficient watershed model calibration ». *Water Resources Research*, Volume 43 n°1, W01413. doi : 10.1029/2005WR004723.

Tolson, B.A., Sharma, V., Swayne, D., Fan, L. (2007). « A parallel implementation of the Dynamically Dimensioned Search (DDS) algorithm ». In: *International Symposium on Environmental Software Systems (ISESS)*, Prague, Czech Republic. May 22-25.

Tolson, B.A., Shoemaker, C.A. (2008). « Efficient prediction uncertainty approximation in the calibration of environmental simulation models ». *Water Resources Research*, Volume 44 n°4, WR005869.

Tolson, B.A., Matott, L.S., Gaffoor, T., Asadzadeh, M., Shafii, M., Xu, X., Pomorski, P., Razavi, S., Hagnegahdar, A., Jahanpour, M., Craig, J.R. (2015). « Parallel and Preemptable Dynamically Dimensioned Search Algorithms for Single and Multi-Objective Optimization ». In: *Water Resource Research*, AGU Fall Meeting, San Francisco, California, USA, December 14-18.

Torczon, V. (1997). « On the convergence of pattern search algorithms ». *SIAM Journal on Optimization*, Volume 7 n°1, p. 1-25. doi :10.113/S1052623493250780.

Tsoukala, I., Kossieris, P., Efstratiadis, A., Makropoulos, C. (2016). « Surrogate-enhanced evolutionary annealing simplex algorithm for effective and efficient optimization of water resources problems on a budget ». *Environmental Modelling and Software*, Volume 77, p.122–142, doi : 10.1016/j.envsoft.2015.12.008.

Turcotte, R., Rousseau, A.N., Fortin, J.-P., Villeneuve, J.-P. (2003). « Development of a process-oriented, multi-objective, hydrological calibration strategy accounting for model structure ». In: Advances in Calibration of Watershed Models, Édité par : Q. Duan, S. Sorooshian, H. Gupta, A.N. Rousseau and R. Turcotte, p. 153-163, American Geophysical Union (AGU), Washington, USA.

Turcotte, R., Fortin, L.G., Fortin, V., Fortin, J.P., Villeneuve, J.P. (2007). « Operational analysis of the spatial distribution and the temporal evolution of the snowpack water equivalent in southern Quebec, Canada ». *Nordic Hydroogy.*, Volume 38, n°1, p. 211-234. doi : 10.2166/nh.2007.009.

Verbunt, M., Gurtz, J., Jasper, K., Lang, H., Warmerdam, P., Zappa, M. (2003). « The hydrological role of snow and glaciers in alpine river basins and their distributed modeling ». *Journal of Hydrology*, Volume 282 n°1-4, p. 36-55.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., Sorooshian, S. (2003). « Effective and efficient algorithm for multiobjective optimization of hydrologic models ». *Water Resources Research*, Volume 39 n°8. doi : 10.1029/2002WR001746.

Wang, C., Duan, Q., Gong, W., Ye, A., Di, Z., Miao, C. (2014). « An evaluation of adaptive surrogate modeling based on optimization with two benchmark problems ». *Environmental Modelling and Software*, Volume 60, p. 197-179. doi : 10.1016/j.envsoft.2014.05.026.

Willers Moore, W.O., Mala-Jetmarova, H., Gebreslassie, M., Tabor, G.R., Belmont, M.R., Savic, D.A. (2016). « Comparison of multiple surrogates for 3D CFD model in tidal farm optimisation », *Procedia Engineering*, Volume 154, p. 1132-1139. doi : 10.1016/j.proeng.2016.07.523.

Yapo, P. O., Gupta, H. V., Sorooshian, S. (1996). « Automatic calibration of conceptual rainfall-runoff models: Sensitivity to calibration data ». *Journal of Hydrology*, Volume 181 n°1-4, p. 23-48. doi : 10.1016/0022-1694(95)02918-4.

Yazdi, J., Salehi Neyshabouri, S.A.A. (2014). « Adaptive surrogate modeling for optimization of flood control detention dams ». *Environmental Modelling & Software*, Volume 61, p. 106-120. doi : 10.1016/j.envsoft.2014.07.007.

Zhang, X., Srinivasan, R., Van Liew, M. (2009). « Approximating SWAT Model Using Artificial Neural Network and Support Vector Machine ». *JAWRA Journal of the American Water Resources Association*, Volume 45, n°2, p. 460-474. doi : 10.1111/j.1752-1688.2009.00302.x.